Ф. БАУЭР

РАСШИФРОВАННЫЕ СЕКРЕТЫ

Методы и принципы криптологии

Издательство «Мир»

Расшифрованные секреты

Методы и принципы криптологии

Friedrich L. Bauer

Decrypted Secrets

Methods and Maxims of Cryptology

Third, Revised and Updated Edition

With 167 Figures, 26 Tables,
and 16 Color Plates

Springer

Ф. Бауэр

Расшифрованные секреты

Методы и принципы криптологии

Перевод с третьего английского издания
В. И. Ахмолина и В. И. Петрова

под редакцией А. В. Чашкина

Москва «Мир»
2007

УДК 519.6
ББК 22.1
Б29

Бауэр Ф.

Б29 Расшифрованные секреты. Методы и принципы криптологии: Пер. с англ. — М.: Мир, 2007. — 550 с.: ил., 16 с. цв. ил.

ISBN 5-03-003551-6

Книга представляет собой вводный курс в криптологию, основанный на лекциях автора в Мюнхенском технологическом институте и материалах семинара по методам криптоанализа. Первая часть посвящена классической области криптографии. Подробно описываются шифры замены и перестановочные шифры — основные кирпичики симметричных криптосистем. Разбор этого класса шифров доведен до современных реализаций (шифры DES и IDEA). С меньшим уровнем подробности рассмотрено шифрование с открытым ключом. Во второй части на классических примерах раскрытия известных шифров описываются основные приемы криптоатак — это первое систематическое изложение криптоанализа на русском языке.

Изложение сопровождается многочисленными примерами. Включенные автором исторические экскурсы делают чтение книги увлекательным.

Книга не имеет аналогов на русском языке. Она будет интересна специалистам и может быть рекомендована любителям математики, криптографии и истории, студентам, преподавателям и учащимся старших классов.

УДК 519.6
ББК 22.1

РФФИ

Издание осуществлено при финансовой поддержке
Российского фонда фундаментальных исследований
по проекту № 01-01-14083

Редакция литературы по математическим наукам

ISBN 5-03-003551-6 (русск.)
ISBN 3-540-42674-4 (англ.)

От редактора перевода

Книга «Расшифрованные секреты» написана Ф. Л. Бауэром, известным немецким специалистом в области вычислительной техники и программирования, удостоенным за свою деятельность многочисленных наград, в том числе и медали Computer Pioneer. Она не похожа на другие книги по криптографии и занимает в современной криптографической литературе особое место. Ее нельзя назвать ни учебником, ни монографией. Скорее, это «книга для чтения», в которой подробно и с большим количеством технических деталей рассказывается о криптографии и ее истории. Книга состоит из двух частей. Первая часть называется «Криптография» и содержит описания различных систем шифрования, начиная со средних веков и кончая двадцатым веком. Вторая часть «Криптоанализ» посвящена методам анализа и взлома систем шифрования, описанных в первой части. В книге можно выделить две самостоятельные линии — криптографическую, посвященную непосредственно криптографии и криптоанализу, и математическую, в которой автор проводит математический анализ криптографических методов и объектов. Эти линии, переплетаясь, сменяя и дополняя одна другую, создают любопытную картину криптологии — довольно специфической области человеческой деятельности, лежащей на границе науки, техники и искусства.

В отличие от большинства авторов современных книг по криптографии Ф. Л. Бауэр не ограничивается техническим описанием тех или иных методов шифрования. Задача предотвращения несанкционированного доступа к информации рассматривается им комплексно — помимо методов шифрования и анализа их стойкости обсуждаются также способы предварительной подготовки информации перед шифрованием и вопросы криптографической дисциплины. Большим достоинством изложения является обилие фактического материала, иллюстрирующего создание и применение криптографических методов в реальных исторических условиях. Особенно интересны и полезны в этом смысле многочисленные описания успешных взломов различных шифров. Знание причин, по которым тот или иной шифр был взломан, знание условий, при которых это произошло, несомненно, поможет избежать многих ошибок при использовании криптографических методов тем читателям, чья деятельность будет связана с защитой информации. Неподготовленному читателю трудно представить, какие мелочи могут привести к взлому даже теоретически достаточно надежного шифра. Ярким примером такого

невнимания к мелочам является не единожды упоминаемый в тексте американский дипломат Мёрфи, настаивавший на том, чтобы его корреспонденция начиналась словами «к Мёрфи» и «от Мёрфи». Результаты этой настойчивости были таковы, что немецкий криптограф г-жа Фридрихс, увидев его после войны, говорила, что хотела «пожать ему руку — так много он для нас сделал». Следует также отметить, что на русском языке практически нет литературы, рассматривающей лингвистические аспекты криптоанализа. Вторая часть «Расшифрованных секретов» восполняет этот пробел.

Хотелось бы также отметить ряд особенностей изложения, с которыми, по моему мнению, не всегда можно согласиться. Так, например, наряду с излишне подробным описанием классов вычетов приводится и требующее от читателя значительной математической культуры краткое дополнение с аксиоматической теорией информации; ничего не говорится о криптоанализе современных блочных систем шифрования, хотя такие системы описаны в первой части, а в современной криптографической литературе можно легко найти многочисленные методы анализа таких систем. Следует также отметить, что автор слишком прямолинеен в оценке некоторых исторических событий, переоценивая влияние криптографии и криптоанализа на ход истории.

В целом книга достаточно удачно описывает состояние и развитие криптографии и криптоанализа вплоть до середины двадцатого века, она не имеет аналогов на русском языке и может быть рекомендована любителям математики, криптографии и истории, студентам, преподавателям и учащимся старших классов.

А. В. Чашкин

Предисловие

К концу шестидесятых годов прошлого века под влиянием бурного развития микроэлектроники электромеханические криптографические машины начали заменять электронными устройствами шифрования данных, построенными на основе больших интегральных схем. Это обещало более безопасное шифрование по более низким ценам. Позже, в 1976 г., Диффи и Хеллман открыли новое направление в криптологии — системы шифрования с открытыми ключами. Криптография, над которой до этого царил покров таинственности, стала превращаться в открытую область. Раскрытие тайны ЭНИГМЫ привлекло к этой науке дополнительный общественный интерес.

Информатика была новой развивающейся областью, и специалисты-компьютерщики стали интересоваться различными аспектами криптологии. Но большинство из них были не очень хорошо знакомы с ее многовековой историей и с высоким уровнем, которого она достигла. Я сталкивался с несколькими «изобретателями велосипеда» в криптографии, попадались и люди, наивно убежденные в надежности шифрования. В результате предметом моих забот стали коммерческая и научная разработка и продвижение профессиональной криптологии в среде компьютерных специалистов, и неустойчивая ситуация, связанная с оказанием официальных услуг в области безопасности. Я решил прочитать курс лекций по данному вопросу в Мюнхенском технологическом институте. Первая серия лекций в зимнем семестре 1977/78 гг. опиралась на исчерпывающую и капитальную книгу «*Взломщики кодов*» (1967 г.) Дэвида Кана и прошла под нейтральным названием «Специальные проблемы теории информации», в силу чего не привлекла большого внимания.

В следующий раз, в летнем семестре 1981 г., мои лекции были анонсированы под открытым названием «Криптология». Это был, по-видимому, первый открыто заявленный университетский курс лекций под таким названием в Германии, а может быть и в континентальной Европе.

Курс лекций был повторен несколько раз, и в 1986/87 гг. был напечатан конспект лекций, который соответствует первой части этой книги. Активный интерес со стороны студентов привел к семинару по методам криптоанализа в летнем семестре 1988 г., из которого родилась вторая часть настоящей книги.

Появившееся в 1993 г. первое издание моей книги «*Kryptologie*», написанной главным образом для студентов, изучающих информатику, вызвало силь-

ный интерес к ней также за пределами этой области. Она попала в поле зрения некоторых ведущих научных обозревателей и была ими благосклонно принята. Последовало издание в виде монографии в твердом переплете в 1995 г. под названием «*Entzifferte Geheimnisse*» («Секреты расшифровывания»), давшее мне возможность обобщить некоторые вещи. Обзоры в американских журналах, которые можно считать английской версией монографии, в конце концов, привели к появлению предлагаемой книги.

В среде криптологов стало принятым объяснять, каким образом они пришли в эту область. В моем случае это не было связано со Второй мировой войной. Фактически я никогда не был членом какой-либо официальной службы — и я рассматриваю это как мое величайшее преимущество, поскольку я не связан какими-либо обязательствами секретности. С другой стороны, имея открытые глаза и уши, я много узнавал, читая между строк, и из бесед (хорошим отправным пунктом для этого была моя научная деятельность), хотя я точно никогда не знаю, позволено ли мне знать то, что мне случалось узнать.

Все началось в 1951 г., когда я сообщил моему учителю, профессору формальной логики Мюнхенского университета, Вильгельму Бритзельмейру, о моем изобретении кода с исправлением ошибок для телетайпных линий[1]. Это навело его на неправильные ассоциации, и он дал мне копию книги Сакко[2], которая была им только что получена. Мне посчастливилось: для меня это была наилучшая книга из всех, какие бы я мог получить в это время — хотя тогда я не знал этого. Тем не менее, я с интересом прочитал книгу. Обратив внимание на это, мой дорогой друг и коллега Пауль Август Манн, который был осведомлен о моем знакомстве с Шенноновским кодированием, дал мне копию ныне знаменитой статьи Клода Шеннона «*Теория связи в секретных системах*»[3] (которая в те дни была почти недоступным в Германии техническим отчетом лаборатории Bell Systems). Я увлекся этим подходом Шеннона к теории информации, с которой я был уже знаком. Это пробудило мой интерес к криптологии, как разделу теории кодирования и формальной теории языков, которые в течение многих лет были областями моих академических интересов.

Странные случайности — или, может быть, просто острая наблюдательность — затем сводили меня все больше и больше с людьми, близкими к криптологии, начиная с Вилли Йенсена (Фленсбург) и Карла Штейна (Мюнхен) в 1955 г., с Гансом Рорбахом, моим коллегой по Манцскому университету в 1959 г., а также с Хельмутом Грюнски, Гизбертом Хесенжегером и Эрнстом Виттом. В 1957 г. я познакомился с Эрихом Гуттенхейном (Бад Гёдесберг), но наши дискуссии о применимости компьютеров для криптологических работ проходили в условиях определенных ограничений. Среди американ-

[1] DBP No. 892767, application date January 21, 1951.

[2] Général Luigi Sacco, *Manuel de Cryptographie*. Payot, Paris 1951.

[3] Bell Systems Technical Journal **28**, Oct. 1949, pp. 656–715. [Имеется русский перевод в сб. Шеннон К. Работы по теории информации и кибернетике. Пер. с англ. — М.: ИЛ, 1963, с. 333–402.]

ских и британских коллег в области численного анализа и информатики, с которыми я имел тесные контакты, были люди, связанные с криптологией во время Второй мировой войны; но никто не говорил об этом вплоть до 1974 г., когда вышла книга Винтербозема (Winterbotham) «*Ultra Secret*». В 1976 г. на симпозиуме в Лос-Аламосе я слышал сообщение Б. Рэндалла и И. Д. Гуда, раскрывающее некоторые детали проекта Колосс (Colossi). Мои интересы в криптологии связаны с компьютерным криптоанализом. Другие аспекты распознавания сигналов («SIGINT»), например анализ трафика и обнаружение направления, выходят за рамки книги; по тем же причинам в книге не рассматриваются физические устройства экранирования электромеханического излучения шифровальных машин.

В первой части этой книги представлены методы шифрования. Вторая часть вводит в криптоанализ, в ней рассматриваются в первую очередь факты, которые важны для аналитиков методов шифрования и знание которых может уберечь пользователя от неожиданных ловушек. Это следует из принципа Керкхоффса: только криптоаналитик может судить о безопасности криптосистемы. Теоретический курс лишь по методам шифрования кажется мне бескровным. Но курс по криптанализу проблематичен: либо он не достаточно полон, и в этом случае — бесполезен, либо он исчерпывающ, но соприкасается с деликатной областью. Здесь есть небольшой зазор. Я попытался, по крайней мере, охватить все существенные факты, которые встречаются в открытой литературе или могут быть выведены из них. Какая-либо цензура при работе над книгой отсутствовала.

Криптология распространена повсеместно и обладает специфической терминологией. Поэтому я счел полезным иногда давать в книге ссылки на иноязычные термины.

Мое интеллектуальное восхищение криптологией можно обнаружить в использовании в книге экспонатов из коллекции «Информатика и автоматика» («Informatik und Automatik») Немецкого музея в Мюнхене, которую я составлял в 1984–1988 гг., и где есть раздел, посвященный криптографическим устройствам и машинам. Я благодарен Немецкому музею, предоставившему возможность поместить цветные вклейки с изображениями некоторых музейных экспонатов.

Также я благодарен моим прежним студентам и сотрудникам в Мюнхене, Манфреду Брою, Герберту Эхлеру и Энтони Джеролду за продолжающуюся годами поддержку, кроме того, школе «Hugh Casement» за лингвистические лакомства и моему последнему свояку Алстону С. Хаусхолдеру за улучшение моего английского. Карл Штейн и Отто Лейберих снабдили меня подробностями истории ЭНИГМЫ, у меня были плодотворные дискуссии и обмен письмами с Ральфом Эрскином, Гейнцом Ульбрихтом, Тони Сале, Фродом Вейрудом, Кжел-Ове Видманом, Отто Д. Хораком и Фрицем-Рудольфом Гюнтшем. Большую помощь оказали мне Кирк Г. Киршхофер из «Crypto AG», Цуг (Швейцария). Хильдегард Бауэр-Вогг предоставил переводы трудных латинских текстов, Мартин Бауэр, Ульрих Бауэр и Бернгард Бауэр выполнили вычисления и подготовили рисунки. Благодарю их всех.

Английская версия существенно улучшена Д. Эндрю Россом, работа с которым была приятной. Особая, моя искренняя благодарность обращена к Дэвиду Кану, который поспособствовал мне («книга — отличная и заслуживает самого широкого распространения») и предложил большое количество улучшений текста. Наконец, я должен поблагодарить еще раз Ганса Восснера из Springer-Verlag за хорошо организованное сотрудничество в течение длительного времени. Издательство необходимо поблагодарить за прекрасное оформление книги. И я заранее благодарю читателей, которые будут настолько любезны, чтобы сообщить мне об ошибках и упущениях.

Графрас, осень 2001 г.

Ф. Л. Бауэр

Часть I

Криптография

Ars ipsi secreta magistro.
[Секрет в искусстве есть даже для мастера.]

Жан Робер дю Карле, 1644 *г.*

Защита важной информации является желанием, восходящим к истокам человеческой культуры.

Отто Хорак, 1994 *г.*

Введение: действующие лица

В. Ф. Фридман

М. Режевски

А. М. Тьюринг

Лишь несколько лет тому назад стало возможным открыто говорить о *криптологии*. Изучение шифрованных записей и их негласное дешифрование были областью, которая всегда процветала под покровом тайны и была с любовью взлелеяна профессионалами. Криптология — истинная наука: она имеет дело со знанием (*scientia — лат.*), обучением и практическими приемами. По своей сути она не только соприкасается с келейностью, но и сама остается под покровом тайны, находясь иногда в полном мраке. Она, можно сказать, является почти секретной наукой. Доступная классическая литература скудна и скупа для проникновения в глубины криптологии: под всемогущим авторитетом власти профессиональные криптологи из дипломатических и военных служб были обязаны носить мантию анонимности или, по крайней мере, соглашаться с цензурой своих публикаций. В результате доступная литература никогда полностью не отражала истинного состояния искусства, которое мы можем предполагать; положение вещей не изменилось в этом отношении и сейчас. Страны различаются по степени своей скрытности: так, США достаточно щедро открыли информацию о состоянии своей криптографии в период Второй мировой войны, а Советский Союз в данной области сохранил полную тишину. Это не удивляло; но странно, что и Великобритания также следовала политике скрытности, которая в некоторых случаях (как в истории с проектом COLOSSUS) представляется чрезмерной. По поводу Германии можно сказать лишь одно: состояние ее криптологии было обнародовано после краха Рейха в 1945 г.[4)]

[4)]Hans Rohcbach (1948 г.). Mathematische und maschinelle Methoden beim Chiffrieren und Dechiffrieren. В: ФИАТ Обзор немецкой науки 1939–1941 гг.: Прикладная Математика, часть I, Висбаден.

Криптология как наука имеет тысячелетнюю историю. Ее развитие шло рука об руку с развитием математики, по крайней мере со времен таких математиков, как Франсуа Виет (1540–1603 гг.) и Джон Валлис (1616–1703 гг.). С точки зрения современной математики криптология демонстрирует черты статистики (Уильям Ф. Фридман, 1920 г.), комбинаторной алгебры (Лейстер С. Хилл, 1929 г.) и теории вероятностей (Клод К. Шеннон, 1941 г.). Вторая мировая война свела, наконец, математиков с криптологией напрямую: в эту область были привлечены такие математики, как Ганс Рорбах (1903–1993 гг.) в Германии и Алан Мэтисон Тьюринг (1912–1954 гг.) в Англии; в США в эту область были привлечены А. Адриан Алберт (1905–1972 гг.) и Маршал Холл (1910–1990 гг.), а также Баркли Россер, Виллард ван Орман Куин, Эндрю М. Глейзон и математики-прикладники Ваннивар Буш (1890–1974 гг.) и Уорен Вейвер (1894–1978 гг.). Упомянем также шведа Арне Бьюрлинга (1905–1986 гг.), поляка Мариана Режевски (1905–1980 гг.), голландца Маурица де Вриза и норвежца Эрнста С. Селмера (р. 1920 г.).

Математические дисциплины, которые играют важную роль в современной криптологии, включают теорию чисел, теорию групп, логику, комбинаторику, теорию сложности, эргодическую теорию и теорию информации. Сама криптология уже может практически рассматриваться как раздел прикладной математики и информатики. И, наоборот, для специалиста в области компьютерных технологий криптология приобретает повышенное практическое значение в связи с проблемами разграничения доступа в операционных системах и базах данных, а также при передаче данных в компьютерных сетях.

Можно было бы также упомянуть современных математиков, которые некоторое время были заняты в официальной криптологии. Некоторые из них предпочитают оставаться инкогнито. Обычной практикой является, когда разведывательные службы не раскрывают даже имена своих ведущих криптологов. Адмирал сэр Хью П. Ф. Синклер, который в 1923 г. стал шефом британской секретной службы (МИ-6), имел прозвище «Quex». Полуофициально он и его преемник генерал сэр Стюарт Г. Мензис (1890–1968 гг.), были известны в то время только как «С». В их подчинении находились многочисленные «чиновники паспортного контроля» в посольствах, а также криптоаналитическая служба в Блетчли Парк. А имя Эрнста С. Феттерлейна (ум. 1944 г.), который вплоть до Октябрьской революции был главой Русского криптоаналитического бюро (с псевдонимом Попов), а с июня 1918 г. обслуживал правительственную школу кодов и шифров (ПШКШ) британского министерства иностранных дел, лишь случайно было упомянуто в открытой криптологической литературе Кристофером Эндрю в 1985 г. и Нигелем Вестом в 1986 г.

Профессиональный криптолог находится в слишком большой опасности из-за посягательств иностранных секретных служб. Важно оставить потенциального противника в полном неведении как относительно собственного выбора методов шифрования («философия шифрования»), так и относительно способности расшифровывания («философия криптоанализа»), чтобы тот считал

свои сообщения непонятными для других. Это удалось в операции по тайному чтению зашифрованных с помощью машины ЭНИГМА сообщений, которое Англия производила в 1940–1945 гг., поскольку факт взлома шифра удалось удержать в секрете от ее противника и не обнаружить его своими ответными действиями. В результате проницательности британской стороны у компетентных немецких служб хотя и возникали время от времени подозрения, но до конца войны (а у некоторых наиболее упертых лиц — до 1974 г.) сохранилось убеждение в непробиваемости шифра, генерируемого их ЭНИГМАМИ.

Меры предосторожности, принятые Союзниками, доходили даже до дезинформации своих собственных людей: капитан Лоуренс Ф. Саффорд из секции криптографии отдела морской связи ВМФ США записал во внутреннем отчете от 18 марта 1942 г. после возвращения капитана Абрахама Синкова и лейтенанта Лео Розена из командировки в феврале 1941 г. в Блетчли Парк: «Наши перспективы когда-либо взломать немецкую шифровальную машину ЭНИГМА довольно призрачны». Это не было его личным мнением. Запись была адресована «читателям» в ВМФ США.

Во время войны материальные ресурсы и даже человеческие жизни часто приносились в жертву, чтобы избежать бóльших потерь в другом месте. В 1974 г. полковник авиации Винтерботхэм сообщил, что Черчилль допустил бомбежку Ковентри из-за того, что боялся обнаружения противником факта чтения англичанами немецких сообщений, зашифрованных ЭНИГМОЙ. Этот рассказ был полной ложью: цели бомбежек указывались изменяемыми кодовыми словами, поэтому узнать их фактически было невозможно. Тем не менее, англичане были первоначально очень обеспокоены, когда в 1943/1944 гг. американцы начали систематически уничтожать все танкеры немецких подводных лодок, чьи позиции они стали узнавать в результате взлома 4-х роторной ЭНИГМЫ, используемой на немецких субмаринах. Англичане обоснованно беспокоились, что немцы должны заподозрить неладное и существенно модифицировать свою систему ЭНИГМА. Немцы же не сделали этого, приписывая (ошибочно) потери предательству. Насколько обоснованными были опасения, стало видно, когда Союзники обнаружили, что на 1 мая 1945 г. планировалось изменение в процедурах генерирования ключей ЭНИГМЫ, что делало бы все существующие методы криптоанализа бесполезными. Это изменение «могло быть, вероятно, осуществлено значительно раньше» (Ральф Эрскин), если бы оно казалось заслуживающим внимания.

Этот шедевр работы секретной службы официально включал в себя «информацию, полученную путем взламывания высококачественных кодов и шифров». Она называлась англичанами кратко «специальная информация» и имела кодовое имя ULTRA, которое также означало его гриф секретности. Американцы аналогично дали кодовое имя MAGIC информации, получаемой в результате взлома японских шифровальных машин; его они дублировали также кодовым именем PURPLE. Как ULTRA, так и MAGIC остались скрытыми от шпионов Оси.

Криптология также имеет точки соприкосновения с криминологией. Ссылки на методы шифрования можно обнаружить в различных учебниках по

криминологии, обычно сопровождаемые описаниями успешного криптоанализа шифрованных сообщений преступников, начиная от бутлегеров до крупных контрабандистов, наркоторговцев, торговцев оружием, шантажистов или мошенников, находящихся уже за решеткой, обычно в связи с попытками их освобождения или подкупа важных свидетелей. В судах экспертная оценка криптолога может быть решающей в вынесении обвинительного приговора. Во времена сухого закона в США Элизабет С. Фридман (урожд. Смит) (1892–1980 гг.), жена знаменитого Уильяма Ф. Фридмана (1891–1969 гг.)[5] и сама профессиональный криптолог, оказала значительную услугу по этой линии. В суде ей иногда было нелегко: защитник развил теорию, что из шифрованного сообщения можно прочитать все что угодно, и поэтому ее криптоанализ есть всего лишь «мнение». Шведский криптолог Ивес Гильден (1895–1963 гг.), племянник астронома Хьюго Гильдена, в 1934 г. помогал полиции в ловле контрабандистов. Известно лишь несколько криминальных криптологов, например Абрахам П. Чесс из Нью-Йорка в начале 50-х.

Бок о бок со штатной криптологией в дипломатических и военных службах сосуществовали любители, особенно начиная с XIX века. Начиная с раскрытия исторических событий отставными профессионалами, такими, как, например, Этьен Базерье[6], и послеобеденных развлечений, практиковавшихся Вейтстоуном[7] и Бэббиджем[8], с журнальными криптоаналитическими публикациями, начиная от Эдгара Аллана По до современной рубрики *Cryptoquip* в *Лос-Анжелес Таймс*, сопровождаемая экскурсами в оккультизм, посещением марсиан и терроризмом, криптология смотрится богатым гобеленом, увитым рассказами из одной из старейшей ветвей криптологии — обменом сообщениями между любовниками.

Учебники по написанию писем, которые появились около 1750 г., вскоре стали предлагать помощь в шифровании. Примерами являются *De geheime brievenschryver, angetoond met verscheydene voorbeelden* некоего G. v. K., Амстердам, 1780, и *Dem Magiske skrivekunstner*, Копенгаген, 1796 г. В последующем столетии мы находим *Sicherster Schutz des Briefgeheimnisses* Эмиля Каца,

[5] Фридман, вероятно наиболее известный современный американский криптолог, введший в 1920 г. *индекс совпадения*, лучшее средство современного криптоанализа.

[6] Этьен Базерье (Etienne Bazeries, 1846–1931 гг.), вероятно наиболее известный французский криптолог современности, автор книги «*Les chiffres secrets devoiles*» (1901 г.).

[7] Сэр Чарльз Вейтстоун (Charles Wheatstone, 1802–1875 гг.), английский физик, профессор королевского колледжа (Лондон), более известный по мосту Wheatstone (не им построенному).

[8] Чарльз Бэббидж (Charles Babbage, 1791–1871 гг.), профессор математики в Кембриджском университете, наиболее известен по его разностной и аналитической машинам.

1901 г., и *Amor als geheimer Bote. Geheimsprache fur Liebende zu Ansichts-Postkarten*, возможно, Карла Петерса, 1904 г.

Восхитительное изображение криптологии в компактной, насыщенной форме, перемешанное с сенсационными деталями из двух мировых войн, впервые открылось широкой публике в 1967 г. в шедевре журналистики и исторической науки, книге Дэвида Кана «*Взломщики кодов*». В последующих 70-х годах картина пополнилась некоторыми существенными материалами из английских источников, чьи документы военного времени были наконец (более или менее) рассекречены; среди них «*Тайная война*» Брайана Джонсона, позже «*История шестого корпуса*» Гордона Уэлчмана. Личности многих криптологов превращают ее историю в чрезвычайно занимательную область.

Коммерческий интерес в криптологии после изобретения телеграфа сконцентрировался на изготовлении кодовых книг, и к концу столетия — в разработке и изготовлении механических и электромеханических шифровальных машин. Позже для вскрытия криптограмм стали использоваться компьютеры, первые успешные попытки применения которых имели место в период Второй мировой войны. Но широкое использование компьютеров началось только после того, как в середине 70-х стал быстро расти коммерческий интерес к шифрованию в частной связи («Криптология идет в народ», Кан, 1979 г.); возможности, открываемые интегральными схемами, совпали с потребностями передачи и хранения компьютерной информации. В дальнейшем росту роли криптологии способствовали требования секретности и опасения подслушивания телефонов, хакерство и промышленный шпионаж. Возрастающая потребность в информационной безопасности придала криптологии значение, ранее неведомое. Частные коммерческие применения криптологии вдруг вырвались вперед, что привело к несколько неортодоксальному управлению ключами, в частности к асимметричным открытым ключам, первоначально предложенным открыто в 1976 г. Диффи и Хеллманом. Более широко, отсутствие адекватной защиты авторского права на компьютерные программы способствовало использованию методов шифрования в программном обеспечении, предназначенном для коммерческого использования.

Однако спрос на «криптологию для всех» порождает противоречия и ведет к столкновению интересов между государством и учеными. Когда использование криптологии становится широко распространенным и многочисленная армия ученых открыто работает в этой области, возникают проблемы для национальной безопасности. Например, власти в Соединенных Штатах начали рассматривать вопрос, должны ли быть запрещены частные исследования в криптологии, подобно запрету на частные исследования в области ядерных вооружений. Через два года после революционной статьи Диффи и Хеллмана, 11 мая 1978 г. Джон М. Хармон, помощник генерального прокурора, из юридического совета министерства юстиции написал доктору Франку Прессу, научному консультанту Президента: «... криптографические исследования ученых и математиков в частном секторе известны как «открытая криптография». Как Вы знаете, серьезное беспокойство, выраженное академическими кругами по поводу правительственного контроля над открытой криптографией

привело специальный комитет сената по разведке к необходимости провести анализ определенных аспектов этой области». Эти аспекты концентрировались вокруг вопроса о применении Правил международной торговли оружием (ITAR) «к распространению криптографической информации, разработанной независимыми от государственного контроля или поддержки учеными и математиками в частном секторе» неконституционным в силу первой поправки к конституции, которая гарантирует свободу слова и прессы. Отмечалось: «Криптография является очень специализированной областью с аудиторией, ограниченной достаточно избранной группой ученых и математиков... следовательно, временная задержка в передаче результатов или идей криптографического исследования, вероятно, не должна лишить последующую публикацию своего полного значения».

Криптологическая информация в высшей степени жизненно важна, она и уязвима в той же степени. Если криптологическая информация раскрыта, усилиям правительства в защите национальной безопасности наносится непоправимый ущерб. Таким образом, как писал Хармон в 1978 г., «схема лицензирования, требующая предоставления криптографической информации до публикации», могла бы преодолеть презумпцию неконституционности. Такая схема наложила бы «требование цензуры криптографической информации, однако с соблюдением строго обозначенных инструкций», поскольку «произвольные ограничения на раскрытие криптографических идей и информации, разработанных учеными и математиками в частном секторе неконституционны».

Кроме того, в 1980 г. министерство юстиции предупредило, что экспортный контроль в криптографии затрагивает «важные конституционные проблемы».

Будем считаться с фактами: криптосистемы рассматриваются правительством США (и не только правительством США) как оружие; оружие для защиты и оружие для атаки. Вторая мировая война преподала нам этот урок.

Хармон кроме того писал: «Исследование в области атомной энергетики аналогично во многих отношениях криптографическим исследованиям. Развитие в обеих областях определяется правительством. Решения правительства о создании или субсидировании исследований в этих областях автоматически приводит к установлению режима секретности из-за неминуемых угроз безопасности, возникающих вследствие утечки информации. Все еще значимое исследование может обойтись без доступа к государственной информации. Результаты исследований, как в атомной энергетике, так и в криптографии кроме военного имеют также значительное гражданское значение. Главное различие между областями в том, что многие исследователи в атомной энергетике зависимы от правительства при получении радиоактивных исходных материалов, необходимых в их исследованиях. Криптографам же нужен только доступ к компьютеру». Другими словами, криптология привлекательна для опасных махинаций даже более, чем атомная энергетика. Хотя криптооружие не убивает непосредственно, оно может скрыть преступления.

Ответственность правительства и ученых в области криптологической деятельности отражена в Акте о компьютерной безопасности Конгресса США

от 1987 г. (Public Law, 100-235). Этим актом был учрежден консультативный совет по безопасности и секретности в компьютерных системах (CSSPAB), сформированный из представителей федерального правительства и компьютерной индустрии. Хотя в США существовал скрытный конфликт, вплоть до 1993 г. казалось, что открытого противостояния удалось избежать благодаря добровольным ограничениям со стороны криптологов (осуществляемым общественной группой анализа криптографии).

В 1993 г., тем не менее, разразилась *криптовойна* между правительством и группами по гражданским правам, которые болезненно восприняли анонсированные в апреле 1993 г. и опубликованные в феврале 1994 г. стандарт шифрования с депонированием ключа (EES) и федеральный стандарт обработки информации (FIPS 185), которые явились также сюрпризом и для CSSPAB. Стандарт делает обязательным депонирование ключей в системах шифрования для частного использования. Этот все еще существующий конфликт не научный, а скорее политический и он все еще представляет угрозу для свободы науки. Лучше положение в либеральной, демократической Европе; власти там имеют менее радужные перспективы успеха в ограничении научной криптологии. В объединенной Европе дискуссии начались в 1994 г. под ключевым словом «Евро-Шифрование», и они могут также привести в итоге к урегулированию неизбежного столкновения интересов государства и научного сообщества. Система депонирования ключей пала в 1999 г. во Франции. В прежнем Советском Союзе проблема была бы конечно легко решена в рамках системы, но и в сегодняшней России, в Китае и в Израиле продолжает сохраняться сильный правительственный контроль.

* * *

Криптография и криптоанализ — два лица криптологии: каждое зависит от другого и каждое влияет на другое, усиливая криптоаналитическую защиту на одной стороне и проводя более эффективную атаку с другой. Успехи довольно редки, неудачи более типичны. Атмосфера молчания, сохраняемая усилиями разведывательных служб, конечно, помогает скрывать затруднения. Все крупные успехи вскрытия криптосистем неприятеля во Второй мировой войне удавались, чаще всего случайно, но все в свою очередь иногда терпели поражения, по крайней мере, частичные. Полагаю, положение не сильно изменится и в XXI столетии — благодаря человеческой глупости и беззаботности.

Глава 1

Вводный конспект

En cryptographie, aucune régle n'est absolue.
[В криптографии никакое правило не абсолютно.]
Этьен Базерье, 1901 *г.*

1.1. Криптография и стеганография

Мы должны различать криптографию (греч. *kryptos*, скрытый) и стеганографию (греч. *steganos*, накрытый). Термин *криптография*, в значении *секретность при записи*, использовался в 1641 г. Джоном Уилкинсом, основателем (вместе с Джоном Валлисом) Королевского научного общества; слово «криптография» было введено в оборот в 1658 г. Томасом Брауном, известным английским врачом и писателем. Цель криптографии — сделать сообщение непонятным для непосвященного читателя: *ars occulte scribendi* (искусство тайной записи). Речь идет об *открытой секретной записи*, т. е. явно распознаваемой как секретная запись.

Термин *стеганография* также использовался в этом смысле Каспаром Шоттом, учеником Афанасия Кирхера в заголовке его книги «*Schola Steganographia*», изданной в Нюрнберге в 1665 г. Однако уже Тритемий использовал его в своем первом труде «*Steganographia*», начатом в 1499 г., в значении «скрытная запись». Ее методы имеют целью сокрытие самого факта *существования* сообщения (которое, однако, может быть неискаженным) — передается нечто, не вызывающее подозрение (Фрэнсис Бэкон, 1623: *ars sine secreti latentis suspicione scribendi* — искусство тайного письма не вызывающего подозрений). По аналогии, мы можем назвать это *скрытой секретной записью* или истинной «*стеганографией*».

Криптографические методы подходят для ведения личного дневника или записной книжки: примеры многочисленны — от Самюэля Пеписа (1633–1703 гг.) до Альфреда К. Кинси (1894–1956 гг.); эти методы годятся и в качестве средства против излишнего любопытства посыльного, несущего важное сообщение. Стеганографические же методы более подходят для передачи сообщения из тюрьмы: здесь список примеров также обширен — от сэра Джона Тревеньона (см. рис. 13), заключенного в тюрьму во вре-

мена гражданской войны в Англии (в XVII в.) до грабителя Французского банка Пастура, чье осуждение описано Андре Ланже, и Клауса Круассана, адвоката и сотрудника Штази, который защищал террористическую группу Бадера—Мейнхофа. Террорист Кристиан Клэр использовал книжный шифр (помечая буквы в книге).

Стеганография распадается на две ветви, лингвистическую и техническую стеганографии. Только первая близко связана с криптографией. Технические виды стеганографии могут быть обрисованы достаточно быстро: невидимые чернила использовались начиная со времен Плиния. Сок лука и молоко доказали свою эффективность и были популярными на протяжении веков (изначально невидимые на бумаге, они приобретают коричневый цвет под действием тепла или ультрафиолетового излучения). Другие классические примеры — полые каблуки и чемоданы с двойным дном.

Среди современных методов технической стеганографии, заслуживающих внимание — высокоскоростная телеграфная связь, передача залпом предварительно записанной на магнитную ленту последовательности кодов Морзе по 20 символов в секунду, и перестановка частотных поддиапазонов («скремблирование») спектра речевого сигнала в случае телефонной связи, широко используемая сегодня в коммерческих целях. Во время Второй мировой войны научное подразделение *Forschungsstelle* немецкой Рейхспочты, возглавляемое Куртом Е. Веттерлейном подслушивало, предположительно, с марта 1942 г. секретные радиотелефонные разговоры между Франклином Делано Рузвельтом и Уинстоном Черчиллем, в том числе и разговор от 29 июля 1943 г., непосредственно перед перемирием с Италией. Через Шелленберга об этих разговорах сообщалось Гиммлеру.

Передача секретных письменных сообщений была революционизирована микрофотографией: *микроточка* размером с пятнышко грязи может содержать целую страницу текста. Микроточка — это экстраординарное развитие «макроточки» Хистиаеса[1)], который обривал голову своего раба, записывал сообщение на ней, и затем ждал, когда волосы отрастут снова. Микроточки были изобретены в 1920-х годах Эммануэлем Голдбергом. Советский разведчик Рудольф Абель наносил свои микроточки на пленку для спектроскопа, имея возможность покупать ее, не привлекая излишнего внимания. Другой советский разведчик, Гордон Арнольд Лонсдейл, скрывал свои микроточки в переплетенных складских накладных. Микроточки, используемые немцами во Второй мировой войне, имели размер, позволяющий использовать их в качестве точки в машинописном документе.

1.2. Семаграммы

В лингвистической стеганографии различают два метода: секретное сообщение либо написано так, чтобы казаться невинным в *открытом сообщении*,

[1)] Кан в своей книге пишет имя Хистиаеса как Histiaeus на стр. 82, Histiaeius на стр. 780 и Histiaieus в указателе имен. Воистину, вот пример *ars occulte scribendi* в остальном очень достоверной книге!

либо оно выражено в форме видимых (хотя, часто, мельчайших) графических деталей в документе или рисунке, в *семаграммах*. Второй метод особенно популярен среди любителей, но большинство быстро охладевает к нему, так как эти детали слишком очевидны для тренированного и бдительного глаза. В молодости Фрэнсис Бэкон (1561–1626 гг.) предложил использовать два шрифта для передачи секретных сообщений (рис. 1); шрифты описаны в латинском переводе «*De dignitate et augmentis scientiarum*» (1623 г.) его книги 1605 г. «*Prolicience and Advancement*». Этот способ никогда не имел большого практического значения (но см. разд. 3.3.3, где он введен для двоичного кода).

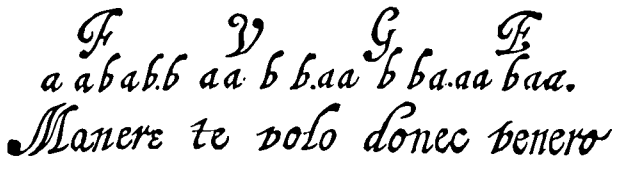


Рис. 1. Фрэнсис Бэкон: видимая маскировка двоичного кода («билитеральный шифр») посредством различных типов шрифтов. Отметим различные формы буквы /e/ в слове *Manere*

В то же самое время данный способ стеганографии, по-видимому, был известен в Париже и упоминается Виженером в 1586 г. Несмотря на свою неуклюжесть, он продолжает использоваться: самый последний пример, известный мне — использование А. ван Вайнгарденом римских цифр и точек в конце предложений в отчете по АЛГОЛ'у 68.

In Königsberg i. Pr. gabelt sich der Pregel und umfließt eine Insel, die *Kneiphof* heißt. In den dreißiger Jahren des achtzehnten Jahrhunderts wurde das Problem gestellt, ob es wohl möglich wäre, in einem Spaziergang jede der sieben Königsberger Brücken genau einmal zu überschreiten.

Daß ein solcher Spaziergang unmöglich ist, war für L. EULER der Anlaß, mit seiner anno 1735 der Akademie der Wissenschaften in St. Petersburg vorgelegten Abhandlung *Solutio problematis ad geometriam situs pertinentis* (Commentarii Academiae Petropolitanae 8 (1741) 128-140) einen der ersten Beiträge zur Topologie zu liefern.

Das Problem besteht darin, im nachfolgend gezeichneten Graphen einen einfachen Kantenzug zu finden, der alle Kanten enthält. Dabei repräsentiert die Ecke vom Grad 5 den Kneiphof und die beiden Ecken vom Grad 2 die Krämerbrücke sowie die Grüne Brücke.

Рис. 2. Семаграмма в учебнике 1976 г. по комбинаторной логике (фрагмент описывает известную задачу о Кенигсбергских мостах). Пониженные символы дают сообщение «*flieder mit dem sowjetimpenalismus*» [вниз с советским империализмом]

Второй способ стеганографии состоит в маркировке отдельных букв в книге или газете, например, точками или черточками. Это значительно заметнее, чем в первом способе (если не используются невидимые чернила), но проще в реализации. Выше показан его вариант (в книге по комбинаторной логике), где в качестве маркера используется почти незаметное понижение выбранных символов (рис. 2).

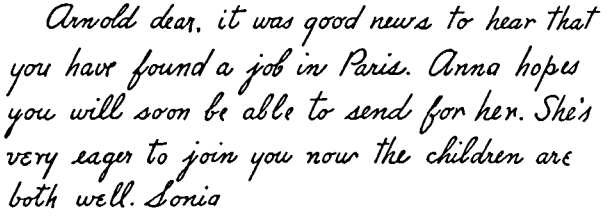


Рис. 3. Видимое сокрытие числового кода с помощью расположения букв

В третьем способе используются расстояния между определенными буквами в пределах слова (рис. 3). В приведенном примере важны не сами буквы, а количество букв между последовательными буквами, заканчивающимися ходом вверх. Подсчитав, получаем последовательность 335151412343334145..., которая, собственно, и несет в себе информацию. В 1895 г. А. Бётцель и Чарльз О'Кинан демонстрировали этот стеганографический прием (также используя числовой код) французским специалистам, которых, однако, не удалось убедить в его полноценности, и не без причины. По-видимому, он был известен еще ранее в российских анархистских кругах, в комбинации с «шифром нигилиста» (см. разд. 3.3.1). Им также пользовались попадавшие в плен немецкие моряки с субмарин, которые в письмах домой пытались сообщать о противолодочной тактике Союзников.

Все это — примеры семаграмм (явное сокрытие секретной записи). Этот список можно продолжить. В античности Эней использовал астрагал (*αρχιτ.*, ободок вокруг колонны — *прим. перев.*), в котором шнур проходил через отверстия, символизирующие буквы. Поле домино может скрывать сообщение (позициями ячеек), как и партия карманных часов (положениями стрелок). Пляшущие человечки Шерлока Холмса (рис. 4) несут собой сообщение также как и невидимый код Морзе (рис. 5): «поздравление от коллектива CPSA MA нашему шефу полковнику Гарольду Р. Шоу в честь его посещения 11 мая

Рис. 4. Секретное сообщение, прочитанное Шерлоком Холмсом (AM HERE ABE SLANEY), из повести Артура Конан Дойла «Пляшущие человечки»

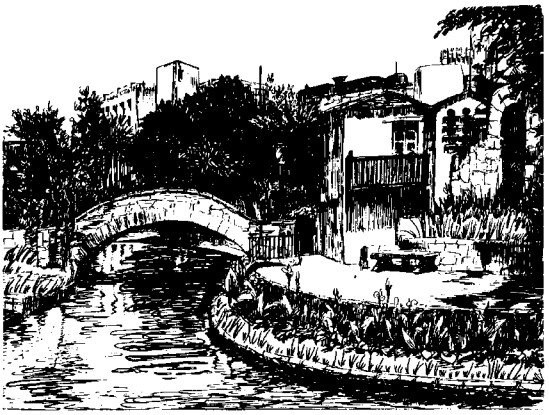


Рис. 5. Семаграмма. Сообщение передается кодом Морзе, сформированным короткими и длинными стеблями травы налево от моста, по береговому откосу и садовой ограде

Где приземлится воздушный шар, в A или в B?

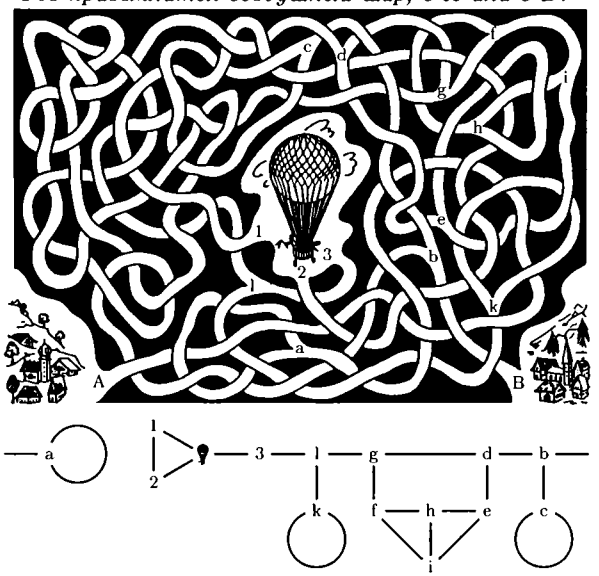


Рис. 6. Лабиринт и диаграмма, связанная с ним

1945 г. Сан Антонио» (Шоу был начальником отдела технических операций департамента цензуры американского правительства, начиная с 1943 г.).

Лабиринт — хороший пример понятной картины, скрытой в большом количестве несущественных деталей: извилистые пути на рис. 6 изображают диа-

грамму, которая может быть понята с первого взгляда. Стереограммы, просмотр которых возможен только при помощи специальных средств (рис. 7), также чрезвычайно подходят для сокрытия изображений, по крайней мере, на некоторое время.

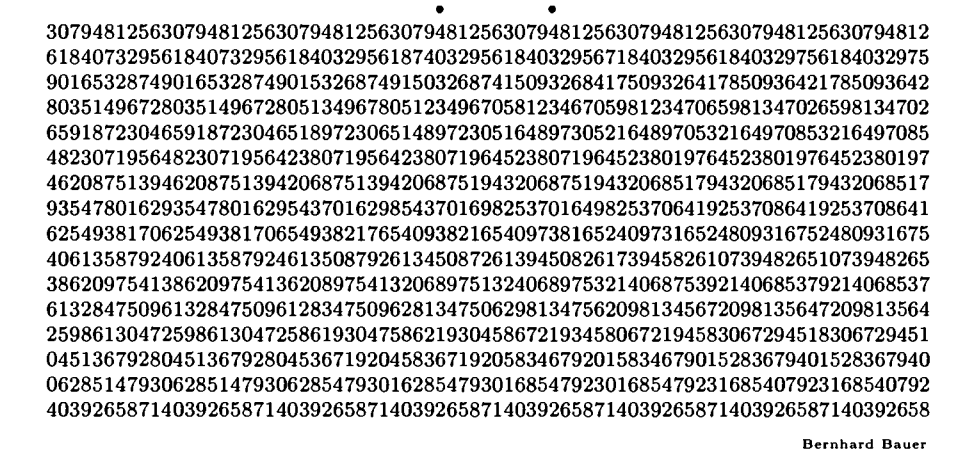


Рис. 7. Стереограмма

Наибольший интерес представляют те методы лингвистической стеганографии, которые преобразуют секретное сообщение в такое, которое является очевидно безопасным и легко понятым, хотя и ненужным (открытое сообщение). Принцип — скрыть сам факт шифрования. Здесь также имеются две подкатегории: маскировка и вуалирование.

1.3. Открытое сообщение: маскировка

Записанное секретное сообщение, замаскированное под открытое сообщение, требует предварительного соглашения относительно истинного значения передаваемых безопасных фраз. Это, вероятно, самая старая форма секретной техники, которую можно найти во всех культурах. Восточные и дальневосточные крупье и картежники (а также некоторые западные) считаются мастерами в использовании условных знаков и выражений. Следующая система, как считают, общепринята среди американских шулеров. Манера держать сигарету или почесывание головы указывает масть или значение карт, имеющихся на руках. Рука на груди с выставленным большим пальцем означает: «Я собираюсь взять эту игру. Есть желающие поучаствовать?». Правая рука ладонью вниз на столе означает «Да», сжатый кулак — «Нет, я работаю один, и я открыл эту игру первым, так что проваливай!». Предполагают, что французский фокусник Роберт Хоудин (1805–1871 гг.) использовал подобную систему около 1845 г., обозначая буквами I, M, S, V масти карт *черви*, *бубны*, *трефы*, *пики*: *il fait chaud* или *il y a du monde* означает «я имею черви», поскольку эти фразы

начинаются с буквы /I/. Подобные вещи были не в диковинку и в английских вист-клубах в Викторианскую эпоху: фраза «вы видели старого Джонса две недели назад?» («Have you seen old Jones in the past fortnight?») означала черви, поскольку фраза начинается с буквы /H/. Британская команда бриджистов подозревалась в обмене сигналами на мировом первенстве в Буэнос-Айресе в 1965 г. — разумеется, ничего нельзя было доказать.

Иногда секретное сообщение может быть передано замаскированным невинным способом, при помощи обстоятельств, известных только отправителю и получателю. Подобные случаи происходят постоянно. Известный пример был рассказан женой Томаса Манна Катей Манн: В марте 1933 г. она позвонила из Ароса в Швейцарии своей дочери Эрике в Мюнхен и сказала: «*Ich weiß nicht, es mußdoch jetzt bei uns gestöbert werden, es ist doch jetzt die Zeit*». [Я не знаю, у нас теперь начнутся метели, это лишь вопрос времени.] Но Эрика ответила: «*Nein, nein, außerdem ist das Wetter so abscheulich. Bleibt ruhig noch ein bissel dort, ihr versäumt ja nichts*». [Нет, нет, в любом случае, погода здесь стоит самая зверская. Задержитесь немного пока, вы ничего здесь не пропустите.] После этого разговора для Кати и Томаса Маннов стало ясно, что им рискованно возвращаться в Германию.

Рис. 8. Секретные метки бродяг («вилки»), предупреждающие о полицейском участке и агрессивном домовладельце (Центральная Европа, около 1930 г.)

Секретная маркировка использовалась в течение столетий, начиная от странствующих школяров средневековья до современных бродяг и бездельников. На рис. 8 изображена пара секретных меток, какие можно было видеть в провинциальных городках Центральной Европы в 1930-х гг., а на рис. 9 показано несколько аналогичных меток, применявшихся на среднем Западе Соединенных Штатов в первой половине XX столетия. Крохотная секретная маркировка также используется при изготовлении штампов или денежных

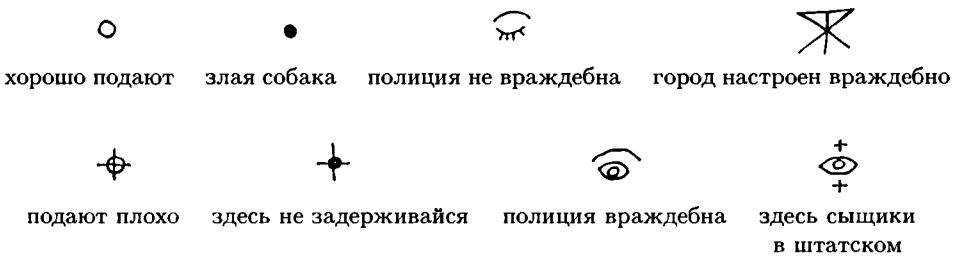


Рис. 9. Секретная маркировка бродяг для «полиция не враждебна» и других сообщений (средний Запад, Соединенные Штаты, первая половина XX столетия)

банкнот в качестве идентифицирующей метки определенного гравера или печатающего устройства.

Специфические языки, характерные для некоторых родов занятий или социальных групп, известные как жаргонные, именуются *argot* во Франции и США, *cant* или *thieves' Latin* (Англия), *rotwelsch* (Германия), *fourbesque* (Италия), *alemania* (Испания) или *calao* (Португалия). Жаргон чаще всего используется нищими, бродягами и некоторыми мошенниками, чтобы оградить (и сохранить в неприкосновенном виде) социальную группу. Они часто используют маскировку. Поэтому замаскированная тайная запись называется *жаргонным кодом.*

Самый старый папский код XIV столетия использовал кодовые слова *Египтяне* для гибеллинов, *Сыновья Израиля* для гвельфов (гибеллины и гвельфы — политические группировки противников и сторонников папского престола, действовавшие в Италии XII–XV вв. — *прим. перев.*). Один французский код XVII столетия использовал исключительно жаргон: *Jardin* означало Рим, *La Roze* — папу Римского, *Prunier* — кардинала Реца, *La Fenestre* — брата короля, *L'Ecurie* (в значении или конюшня, или дворянство) означало Германию, *Le Roussin* — герцога Баварского и т. д. Простая маскировка имен использовалась в Бонапартистском заговоре 1831 года.

Определенный стеганографический интерес представляют также и языки уголовного мира. Французский жаргон предлагает много примеров, некоторые из которых вошли в обычную разговорную речь: синоним *rossignol* (соловей) для отмычки известен с 1406 г.; *mouche* (насекомое) для информатора («стукач» в британском сленге) — с 1389 г. Распространено аллитерирующее повторение: *rebecca* для *rebellion* (бунт), *limace* (железка) для *lime* (напильник), которое, в свою очередь, есть жаргонное обозначение (в итальянском языке) рубашки; *marquise* (маркиза) для *marque* (родинка или шрам), которое, в свою очередь, есть жаргонное название (в испанском языке) девушки; *frise* (кудрявый) для Fritz (популярное название для немца). Не столь устойчивы метафоры: *chateau* (замок) для больницы, *mitraille* (пуля) для мелочи, или живописное, но уничижительное *marmite* (кухонный горшок) для подруги сутенера, *sac a charbon* (угольный мешок) для священника. Саркастические метафоры типа *mouthpiece* (болтун) для адвоката не ограничиваются преступным миром.

Некоторые жаргонные слова имеют интернациональный характер: «hole» — *trou* — *Loch* для тюрьмы; «snow» — *neige-Schnee* или «sugar» — *sukre* для кокаина; «hot» — *heiβ* для недавно украденного товара; «clean out» — *nettoyer* — *abstauben* для «грабить»; «rock» — *galette* (*galet* = гравий) для денег. Все виды игр слов и игр со словами находят здесь свое место. Во Второй мировой войне британский «Двадцатый комитет», который специализировался на двойных агентах, взял свое название от римского числа XX как обозначения двойного креста.

Хорошо замаскированные секретные коды для более или менее *универсального* использования трудно изобрести и еще труднее использовать должным образом — опытный цензор быстро определит неестественный

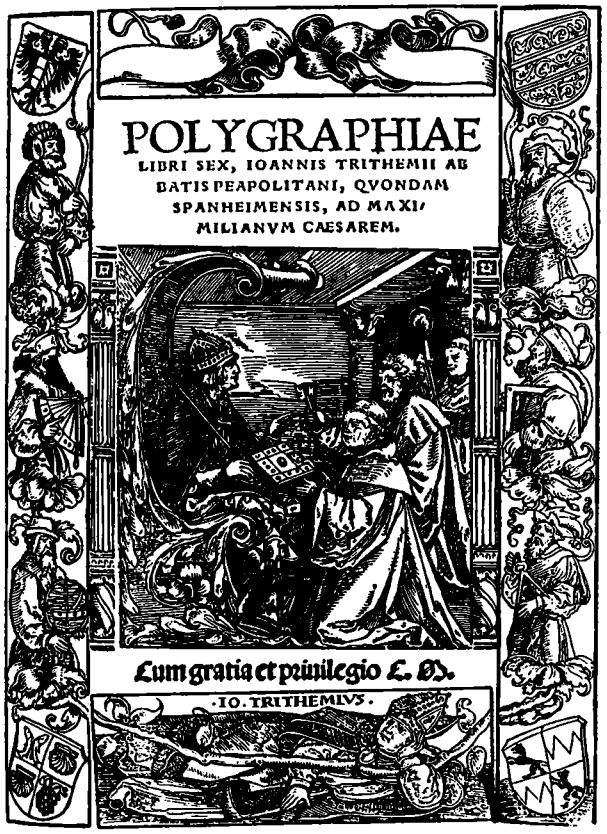


Рис. 10. Титульный лист (гравюра на дереве) первой печатной книги по криптографии (1508 г.)

язык. Аббат Иоганнес Тритемий (1462–1516 гг.) в своей книге «Полиграфия» (*Polygraphiae*), шесть книг которой были напечатаны в 1508–1518 гг. (рис. 10), в шифре *Ave Maria* предложил набор латинских слов в качестве кодов для отдельных букв (рис. 11); например слово «Head» могло быть замаскировано как «ARBITER MAGNUS DEUS PIJSSIMUS». Фактически, имелось 384 таких алфавита в первой книге, используемых последовательно — замечательный и один из наиболее ранних случаев многоалфавитного шифрования (разд. 2.3.3).

Возможно, современные цензоры не достаточно хорошо знают латынь, чтобы справиться с подобным кодом. Любимой уловкой цензуры является переформулирование сообщения с сохранением семантики. В Первой мировой войне цензор изменил сообщение, отправив вместо «Father is dead (отец умер)» сообщение «Father is deceased (отец скончался)». Назад прибыло сообщение «Is father dead or deceased? (отец умер или скончался?)».

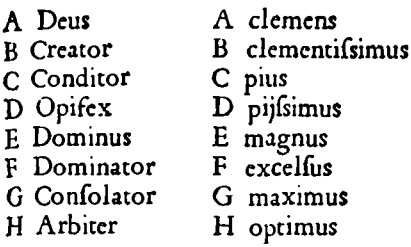


Рис. 11. Первые коды шифра Тритемия *Ave Maria*

Дадим небольшую справку об аллегорическом языке. В 1755 г. дипломатической службой короля Людовика XV в Россию был послан с секретной миссией шевалье Дуглас (Chevalier Douglas) в качестве торговца мехом. Он имел набор аллегорических сообщений для передачи донесений, начиная с *renard noir était cher,* означающего «влияние Английской партии при дворе увеличивается» до *le loup-cervier avait son prix,* означающего «Австрийская партия (руководимая Бестужевым) сохраняет свое доминирующее влияние». Бестужев, питавший к Пруссии симпатию, кодировался как *loup-cervier,* в то время как *une peau de petit-gris* означало 3000 наемников на службе Британии.

Можно надеяться, что шевалье был более искусен в использовании своего аллегорического кода, чем немецкие шпионы, которые под видом голландских торговцев — как рассказал генерал-майор Кирк — каждый день заказывали сигары партиями в тысячах поочередно из Плимута, затем из Портсмута, далее из Гравезенда и так далее — 1000 «коронас» означали один линкор. Их надуманная система преждевременно оборвала их жизни 30 июля 1915 г. Более удачлива была Вельвали Дикинсон, женщина из японского квартала Нью-Йорка, которая в 1944 г. вела оживленную переписку о сломанных куклах. Дело вскрылось, когда письмо, отправленное из Портленда, штат Орегон, было возвращено и обнаружилось, что имя отправителя на нем было вымышленным. Леди действительно занималась продажей изящных куколок из магазина на Мэдисон авеню. Отдел технических операций, подразделение в структуре ФБР трудился особенно интенсивно, чтобы найти скрытые сообщения в ее переписке и собрать доказательства для обвинения в шпионаже, но удалось лишь добиться того, что она получила десять лет тюрьмы и была оштрафована на $ 10 000. В фильме «*Завтрак у Тиффани*» Одри Хепберн в роли мисс Холли Голайтли провела ночь за решеткой за то, что помогала гангстеру руководить сетью торговцев кокаином из его тюремной камеры посредством «сообщений о погоде» — эта связь через нее закончилась, когда она сообщила о «снеге в Новом Орлеане», что звучало несколько невероятно.

1.4. Сигналы

Наиболее важный частный случай маскировки, а именно жаргонное кодирование сообщений, связан с использованием сигнальных сообщений (*cue* —

от французского *mot convenu* — условное слово), заранее подготовленных фраз или стихов, которые означают определенные сообщения. Смысл сообщения может быть связан со временем передачи; сообщение может играть роль сигнала тревоги или подтверждения. Во время Второй мировой войны радиостанция Би-Би-Си передавала в больших количествах сообщения для Французского *Сопротивления*. Поначалу они не привлекали большого внимания, пока ряд замаскированных сообщений, несущих несколько важных распоряжений, не стали чаще других появляться в эфире. Например, 1 июня 1944 г. 9-часовые новости сопровождались чтением стихов «по заявкам», включающим первую половину первого стиха поэмы *Chanson d'Automne* Поля Верлена («Долгие рыдания осенних скрипок»). Вторая половина стиха («Ранят мое сердце своим монотонным томлением») последовала 5-го июня. Германское командование, информированное *Абвером* адмирала Канариса, уже в январе 1944 г. имело коды сообщений и их значения. Когда 15-я армия приняла ожидаемый сигнал (рис. 12), командные пункты немцев были предупреждены, но по причинам, которые не известны до настоящего вре-

Tag Uhrzeit Ort und Art der Unterkunft	Darstellung der Ereignisse (Dabei wichtig: Beurteilung der Lage (Feind- und eigene), Eingangs- und Abgangszeiten von Meldungen und Befehlen)
5.6.44	Am 1., 2. und 3.6.44 ist durch die Nast innerhalb der
	"Messages personelles" der französischen Sendungen des
	britischen Rundfunks folgende Meldung abgehört worden :
	"Les sanglots longs des violons de l'automne ".
	Nach vorhandenen Unterlagen soll dieser Spruch am 1. oder
	15. eines Monats durchgegeben werden, nur die erste Hälfte
	eines ganzen Spruches darstellen und ankündigen, dass
	binnen 48 Stunden nach Durchgabe der zweiten Hälfte des
	Spruches, gerechnet von 00.00 Uhr des auf die Durchsage
	folgenden Tages ab, die anglo-amerikanische Invasion be-
	ginnt.
21.15 Uhr	Zweite Hälfte des Spruches "Blessent mon coeur d'une longeur
	monotone" wird durch Nast abgehört.
21.20 Uhr	Spruch an Ic-AO durchgegeben. Danach mit Invasionsbeginn
	ab 6.6. 00.00 Uhr innerhalb 48 Stunden zu rechnen.
	Überprüfung der Meldung durch Rückfrage beim Militärbe-
	fehlshaber Belgien/Nordfrankreich in Brüssel (Major von
	Wangenheim).
22.00 Uhr	Meldung an O.B. und Chef des Generalstabes.
22.15 Uhr	Weitergabe gemäss Fernschreiben (Anlage 1) an General-
	kommandos. Mündliche Weitergabe an 16. Flak-Division.

Рис. 12. Извлечения из журнала регистрации отдела радиоразведки 15-й армии (подполковник Хельмут Майер, сержант Вальтер Рейхлинг), *longeur* следует читать как *longueur*

мени, предупреждение не достигло 7-й армии, на часть побережья которой происходила высадка союзников в течение 48 часов 6 июня 1944 г.

Японцы использовали аналогичную систему в 1941 г. Например, «ХИГАШИ НО КАЗЕ АМЕ» (ветер восточный, дождь), вставленное в сообщение о погоде в зарубежных новостях и повторенное дважды, сообщало об объявлении «войны с США». ВМФ США перехватил дипломатическую радиограмму с этим содержанием 19 ноября 1941 г. и сумел ее расшифровать 28-го. Поскольку напряженность в отношениях повысилась, многочисленные станции прослушивания в США вели мониторинг передач японского радио в поисках сигнального сообщения. Оно прошло 7-го декабря, спустя несколько часов после атаки на Перл-Харбор в виде «ХИШИ НО КАЗЕ ХАРЕ» (ветер западный, ясно), обозначая начало военных действий с Великобританией, что для американцев было несколько неожиданным. Возможно, все это было, в свою очередь, японской хитростью.

В техническом плане замаскированная секретная запись демонстрирует некоторое сходство с зашифрованной секретной записью (разд. 3.2), особенно с применением замен (гл. 3) и кодов (разд. 4.4).

Другой стеганографической подкатегорией являются секретные записи или сообщения, завуалированные под открытые (незаметно скрывающие в себе секретную запись). Здесь сообщение, предназначенное для передачи, тем или иным способом преобразуется в открытое, невинно выглядящее сообщение путем добавления несущественной информации — *пустышек*.

Для того чтобы было возможным реконструировать реальное сообщение, место, где оно скрыто, должно быть оговорено заранее (*шифрование сокрытием*). Имеются две очевидных возможности в использовании *пустышек:* путем фиксации правил (*шифрование пустышками*, *шифрование с известным символом*), либо с использованием *трафарета* (в английском языке используется французский термин *grille* (решетка, трафарет)).

1.5. Открытый код: маскировка пустышками

Правила для скрытых сообщений имеют очень часто тип «*n*-й символ после соответствующего символа», например, следующий символ после пробела (на нем построен «семейный код», который был популярен среди солдат Второй мировой войны и вызывал большое неудовольствие цензоров); лучше был бы третий символ после пробела, или третий символ после знака препинания. Такие секретные сообщения называются акростихами. Опытный цензор обычно немедленно распознает их по неестественности языка, кое-каким неправильностям, и его наметанный глаз сразу обнаружит, что означает сообщение

ŘRESIDENT'S ĚMBARGO ŘULING ŠHOULD ĤAVE ĬMMEDIATE ŇOTICE. ĞRAVE ŠITUATION ĂFFECTING ĬNTERNATIONAL ĽAW. ŠTATEMENT ŘORESHADOWS ŘUIN ŎF ŇANY ŇEUTRALS. ŸELLOW ĴOURNALS ŬNIFYING ŇATIONAL ĚXCITEMENT ĬMMENSELY,

перехваченное в Первую мировую войну. Иногда может помочь выписывание слов одно под другим:

```
    ↓
I N S P E C T
D E T A I L S
F O R
T R I G L E T H
A C K N W L E D G E
T H E
B O N D S
F R O M
F E W E L L
```

Маскировка отпадает: открытый текст (STRIKE NOW — немедленно нанести удар) тут же бросается в глаза.

Сэр Джон Тревеньон, который жил во времена Оливера Кромвеля (1599–1658 гг.), избежал казни благодаря своему воображению. В письме от своего друга R. T. он обнаружил сообщение «panel at east end of chapel slides» (панель в конце восточного крыла часовни сдвигается) и, воспользовавшись им, сумел выбраться из плена (рис. 13).

> Worthie Sir John: — Hoṗe, thăt is ye beste comfort of ye afflicted, caṅnot much, I fĕar me, heĺp you now. Thăt I would saye to you, is t̆his only: if ĕver I may be able to requite that I do owe you, stănd not upon asking me. 'Tiš not much that I can do: buṫ what I can do, beĕ ye verie sure I wille. I k̆nowe that, if d̆ethe comes, if ŏrdinary men fear it, it f̆rights not you, acc̆ounting it for a high honour, to h̆ave such a rewarde of your loyalty. Prăy yet that you may be spared this soe bitter, cuṗ. I fĕar not that you will grudge any sufferings; onl̆y if bie submission you can turn them away, 'tiš the part of a wise man. Teĺl me, an f̆ you can, to d̆o for you anythinge that you wolde have done. Thĕ general goes back on Wednesday. Reštinge your servant to command. — R. T.

Рис. 13. Сообщение сэру Джону Тревеньону: panel at east end of chapel slides (по третьим символам после знаков препинания)

Существует история о солдате армии США, который договорился со своими родителями, что он сообщит им название места своей службы, используя начальные буквы первых слов (после приветствия) в последовательных письмах домой — с криптографической и стеганографической точек зрения не такая уж плохая идея. Однако его уловка была раскрыта, когда его родители написали ему: «Где находится Нутси? Мы не можем найти его в нашем атласе». Бедняга забыл проставить даты в своих письмах!

Техника акростиха нашла применение даже в беллетристике. В классическом акростихе использовались начальные символы, слоги или слова последовательных строк, стихов, параграфов или глав. Таким способом зашифровывались слова, имена авторов или предложения (рис. 14). Акростихи также служили защитой против изъятий и вставок в текст: первые примеры современных проверок на четность или кодов с обнаружением ошибок.

Fast writing method

He must have had a special trick, said Robert K. Merton, for he wrote such an amazing quantity of material that his friends were simply astonished at his prodigious output of long manuscripts, the contents of which were remarkable and fascinating, from the first simple lines, over fluently written pages where word after word flowed relentlessly onward, where ideas tumbled in a riot of colorful and creative imagery, to ends that stopped abruptly, each script more curiously charming than its predecessors, each line more whimsically apposite, yet unexpected, than the lines on which it built, ever onward, striving toward a resolution in a wonderland of playful verbosity. Fuller could write page after page so fluently as to excite the envy of any writers less gifted and creative than he. At last, one day, he revealed his secret, then died a few days later. He collected a group of acolytes and filled their glasses, then wrote some words on a sheet of paper, in flowing script. He invited his friends to puzzle a while over the words and departed. One companion took a pen and told the rest to watch. Fuller returned to find the page filled with words of no less charm than those that graced his own writings. Thus the secret was revealed, and Fuller got drunk. He died, yet still a space remains in the library for his collected works.

Л. Фишер и Д. Эндрю Росс

Рис. 14. Самоописание акростиха

Подобным способом, хронограмма скрывает (римскую) цифру в надписи; обычно это — дата, например год, установки мемориальной доски. Над южной дверью причудливой церкви на улице Гартльберг в Пфаркирхене в нижней Баварии имеется надпись:

LVDoVICVs seVerVs DVX baVarVs aC paLatInVs,
hIC In sanCta paCe qVIesCIt.

(Ludwig the Severe, Duke of Bavaria and Count Palatine, rests here in holy peace.)

Если хронограмма состоит из стихов, то используется термин хроностих, а для двустиший — хронодистих.

Композиторы скрывали сообщения в своих композициях, либо в нотных записях музыкальных тем (известный пример[2)] — B A C H), или косвенно с помощью цифрового алфавита: если i-я нота встречается k раз, то k-й символ алфавита размещается в i-й позиции. Иоганн Себастиан Бах любил этот шифр; в органном хорале «*Vor deinen Thron*», написанном в 1750 г., в тональности соль-мажор, нота g встречается дважды (B), a — один раз (A), b — три раза (C) и c — восемь раз (H).

[2)] В немецком языке b используется для обозначения си-бемоля, h для ноты си. В соль-мажоре g есть первая нота, a — вторая, h — третья и т. д.

Пустышки также используются во многих жаргонах: простое добавление концевого слога (паразитное суффицирование) — самая простая и самая старая система. Например, по-французски,

floutiere для *fiou* (жаргонное «уходи!»);
girolle для *gie* (жаргонное «да»);
mezis для *me* (жаргонное «мне?»);
icicaille для *ici* (жаргонное «здесь»)

и имеются сотни подобных форм. Картуш (XVIII век) писал

vousierge trouvaille bonorgue ce gigotmouche

(здесь пустышечные слоги подчеркнуты).

В Tut Latin, языке школяров, слог *tut* вставляется между всеми слогами. Подобные школьные жаргоны, по-видимому, очень стары: известен отчет 1670 г. Меца (Лорейнского) о системе «заикания», в которой, например, *undreque foudreque* обозначает *un fou*. «Яванский» язык — также из этого класса:

jave вместо *je*;
layeblavanc вместо *le blanc*;
navon вместо *non*;
chayaussayuraye вместо *chaussure*.

Другие системы используют фиктивные слоги с дублированными гласными, типа В-говора в немецком языке:

GABARTEBENLAUBAUBEBE для gartenlaube (беседка)

или Cadogan во французском языке:

CADGADODGOGADGAN для cadogan.

Иоахим Рингельнатц (1883–1934 гг.) написал целую поэму на языке *Bi* (рис. 15). Простое инвертирование порядка букв, называемое обратный сленг, встречается в жаргонных языках: OCCABOT для «tobacco», KOOL для «look», YOB для «boy», SLOP для «полиции».

Gedicht in Bi-Sprache

Ibich hibibebi dibich,
Lobirrebi, sobi liebib.
Habist aubich dubi mibich
Liebibi Neibin, vebirgibib.
Nabih obidebir febirn,
Gobitt seibi dibir gubit.
Meibin Hebirz habit gebirn
Abin dibir gebirubiht.

Рис. 15. Поэма Иоахима Рингельнатца на языке *Bi*

Перестановки слогов можно найти во французских *Verlan* (от *l'envers*): NIBERQUE от *bernique* («ничегонеделание», не сказал бы, что оно было связанно с *bernicles*, маленькими скорлупками); LONTOU от *Toulon*, LIBRECA от *calibre* (в смысле калибр огнестрельного оружия); DREAUPER для *perdreau* (куропатка, означает полисмена); RIPOU от *pourri* (гнилой); BEUR от *rebeu* (араб). Более свежие примеры — FÉCA от *café*, TÉCI от *cité*. Более сложные системы включают перемешивание букв, т. е. транспозиции (см. разд. 6.1). Преступная среда была родоначальником языка Largonji:

leudé от *deux* [франки];
linvé от *vingt* [су];
laranqué от *quarante* [су];

с фонетическими вариантами

linspré от *prince* (Видок, 1837 г.);
lorcefée от *La Force* (тюрьма в Париже);

и языка Largonjem:

lombem для *bon* (1821 г.);
loucherbem для *boucher*;
olrapem для *opéra* (1883 г.).

Название Largonji само сформировано таким же образом из слова «jargon». Вариант с подавлением начального согласного — язык Largondu:

lavedu для *cave*;
loquedu для *toque*;
ligodu для *gigo*(*t*).

Подобные правила образования видны в конструкции следующих слов:

locromuche для *maquerau* (сутенер);
leaubiche для *beau*;
nebdutac для *tabac* (1866 г.);
licelargu для *cigare* (1915 г.).

Эти системы также имеют параллели в Восточной Азии (Ханой, Хайфон). Pig Latin, другой школярский язык, добавляет AY в конце циклически переставляемого слова: third превращается в IRDTHAY. Кокни используют рифмующий сленг с опущенными словами: TWIST И TWIRL для *girl*, JAR OF JAM для tram, STORM AND STRIFE для *wife*, BOWL OF CHALK для *talk*, FLEAS AND ANTS для *pants*, APPLES AND PEARS для *stairs*, BULL AND COW для *row*, CAIN AND ABEL для *table*, FRANCE AND SPAINE для *rain*, PLATES AND MEAT для *feet*, LOAF OF BREAD для *head*. Фактическое рифмующее слово обычно опускается — посвященный может его вспомнить. Некоторые из этих выражений проникли в обычный язык (лексикализация): немногие знают происхождение выражений «use your loaf» или «mind your plates».

Джонатан Свифт (1667–1745 гг.) не слишком осторожничал в своем *Journal to Stella* (за этим именем на самом деле скрывалась Эстер Джонсон

(1681–1728 гг.)): в письме от 24 февраля 1711 г. он просто вставил пустышки после каждой второй буквы.

1.6. Открытый код: маскировка трафаретом

Метод трафарета, который восходит к Джеронимо Кардано (в «*Subtilitate*», 1550 г.), является весьма простым для понимания, но имеет тот недостаток, что обе стороны должны иметь и хранить трафарет, что затруднительно для солдата, находящегося в чистом поле, или же узника, заключенного в тюрьме. Кроме того, составить письмо, используя трафарет, чрезвычайно трудно. Если бы лорд Байрон (1788–1824 гг.) — по общему признанию, отнюдь не простой солдат — использовал бы этот метод, потребовался бы весь его весьма искусный талант для создания поэмы, подобной приведенной на рис. 16. Вероятно, он был бы также способен расположить ее настолько естественно, что открытый текст, расположенный в окнах решетки не привлек бы излишнего внимания.

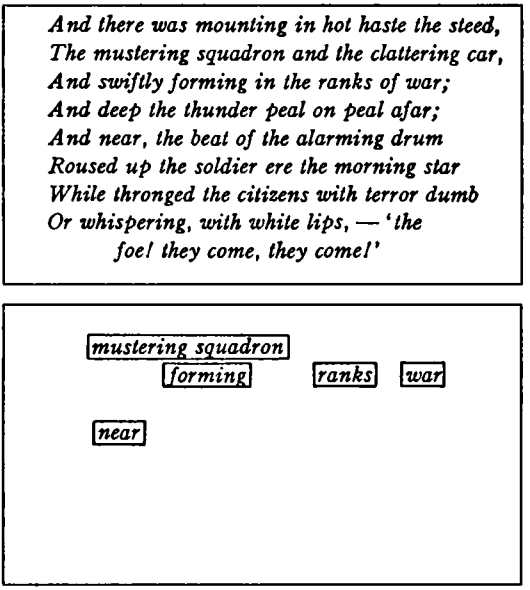


Рис. 16. Гипотетическое сообщение лорда Байрона

Кардано, кстати, настаивал на дублировании сообщения три раза, чтобы устранить любые нерегулярности в размерах букв или интервалах между ними. Этот метод иногда использовался в дипломатической переписке в XVI–XVII вв. Кардинал Ришелье, как предполагают, пользовался им. В современной литературе также упоминаются несколько более хитрых правил, например, предлагается передавать двоичные числа (которыми предварительно

зашифровывается сообщение) словами — слово с четным числом гласных представляет цифру 0, а с нечетным числом — цифру 1.

1.7. Классификация криптографических методов

На рис. 17 показана диаграмма, суммирующая классификацию методов стеганографии и их криптографические свойства, описанные в этой и следующей главах.

Маскировка и вуалирование были подробно рассмотрены здесь, поскольку они обеспечивают нас методическим ориентиром: маскировка ведет к подстановкам, вуалирование — к перестановкам, которые являются двумя основными элементами собственно криптографии. Мы рассмотрим их в следующей главе.

Стеганография также обнаруживает важный принцип: естественный язык — разговорный, письменный или знаковый — имеет свои специфические правила, которые гораздо труднее имитировать (как в стеганографии), нежели подавить их (как в криптографии).

Рис. 17. Классификация стеганографических и криптографических методов

Поэтому профессиональные криптографы относятся к лингвистической стеганографии с осторожностью: она, скорее, относится к епархии цензоров. По своей сути любительская стеганография не представляет столь большой опасности, чтобы подавлять или искать ее. Для цензуры выявление ее приме-

нения часто имеет небольшое значение (кроме, возможно, случая обеспечения доказательствами последующего суда).

Профессиональное использование лингвистической стеганографии может быть оправдано только в частных случаях, когда необходимо скрыть криптографический метод.

Стеганография и криптография относятся к криптологии. Термин *криптология* использовался (как и *криптография*) Джоном Уилкинсом в 1641 г. в значении *секретность в речи*. В 1645 г. слово «криптология» было отчеканено Джеймсом Хауэллом в надписи «криптология или переписка тайным способом, очень древняя». Использование слов «cryptography» (англ.), «*cryptographie*» (фр.), «*criptografia*» (исп.), «*Kryptographie*» (нем.) до недавнего времени доминировало в этой области, даже, когда в нее был включен криптоанализ.

Шеннон в 1945 г. назвал свой секретный доклад, посвященный защите от незаконного дешифрования, «Математической теорией криптографии».

В качестве заглавия книги, термин «*cryptologue*» был использован Ивом Жильденом в 1932 г. и криптологом Уильямом Ф. Фридманом в 1961 г.; «криптология» появляется в заголовке статьи Дэвида Кана в 1963 г.; термин использовался во внутренних отчетах Фридманом и Ламброзо Д. Каллимахосом в 1950-х гг. После книги Кана «*Взломщики кодов*» («Codebreakers», 1967 г.), термин «криптология» окончательно утвердился, включая в себя и криптографию и криптоанализ, и принят теперь повсеместно.

Клод Шеннон (1916—2001)

В последнее время стали доступны мощные программы обработки изображений, и это вернуло угасший было интерес к стеганографии. С помощью хитрого алгоритма сообщение можно скрыть в рисунке или фотографии.

Глава 2

Цели и методы криптографии

> Почти каждый изобретатель системы шифрования убежден в невозможности взлома своего детища.
>
> *Дэвид Кан*

В этой главе мы дадим обзор известных криптографических методов с точки зрения обеспечения секретности[1)] имеющихся каналов связи в условиях (пассивного) перехвата и (активной) фальсификации сообщений (стандарт ISO 7498). Защита от раскрытия секретной информации в смысле конфиденциальности и приватности является классической целью криптографии, тогда как защита против подделки и подмены сообщений, т. е. аутентификация (установление подлинности) отправителя приобрели большое значение лишь недавно.

В криптологии помимо математических вопросов заметное место занимают филологические аспекты. Родственной дисциплиной в этом плане для нее является расшифровка древних текстов на вымерших языках[2)], область, граничащая с археологией и лингвистикой. На вклейке A в качестве примера изображен Фестский диск — один из самых ранних примеров зашифрованного текста.

2.1. Природа криптографии

Цель криптографии состоит в том, чтобы сделать сообщение или запись непонятными посторонним лицам. В этом легко можно перегнуть палку, приведя сообщение в вид, не поддающийся расшифровке и адресатом — кто не испытывал бессилие в попытках прочитать через несколько недель (или даже дней) свои же второпях написанные заметки?

[1)] Со времени открытий Шенноном и Хеммингом приблизительно в 1950 г. кодов с обнаружением и исправлением ошибок, эти коды являются техническими средствами борьбы с физическимим ошибками в каналах связи, и они не должны рассматриваться здесь.

[2)] Johannes Friedrich, *Extinct Languages*, New York, 1957.

Говоря серьезно, ошибка при шифровании или искажение (или порча) сообщения в сеансе радиосвязи могут привести к фатальным последствиям. Любая попытка перешифровать и передать заново то же самое сообщение правильно представляет серьезную угрозу безопасности по причинам, которые мы обсудим в гл. 10 и во второй части книги. Поэтому строгая дисциплина шифрования запрещает это; текст должен быть переписан, конечно же, без изменения содержащейся в нем информации. Это проще сказать, чем выполнить — дорога в ад обычно мостится хорошими намерениями.

2.1.1. Поэтому методы шифрования и расшифрования также не должны быть излишне сложными: они должны соответствовать обстоятельствам и интеллекту людей, которым приходится пользоваться этими методами. Самые слабые стандарты шифрования применяются на поле боя. В области дипломатии подразумевается, что посол в состоянии сам произвести шифрование и расшифрование. Когда Вейтстоун в 1854 г. демонстрировал в британском МИДе метод шифрования, известный ныне как Плейфейр (разд. 4.2), он сказал, что трое из четырех мальчиков из ближайшей школы в состоянии освоить его, на что заместитель госсекретаря сухо заметил: «весьма возможно, что это так, но вы никогда не сможете научить этому наших атташе!».

Нужно также иметь в виду, что многие сообщения должны сохраняться секретными только до того момента, пока события, к которым они относятся, не станут достоянием гласности. По общему мнению считается разумным сохранять тайну дипломатической переписки в течение нескольких десятилетий. Англичан — не говоря уже о русских — невозможно превзойти в этом отношении, настолько плотна завеса тайны, которой они полностью накрыли свои криптографические системы. Во всяком случае, нам достаточно только знать, как много времени потребуется криптоаналитику поработать с сообщением, чтобы прочитать его (взломав шифр), и тогда становится бессмысленно использовать этот конкретный метод шифрования, если мы стремимся к абсолютной безопасности. Шифр, использованный немцами на Западном фронте в 1917 г. (известный французам как KRUSA, потому что все кодовые группы начались с одного из этих пяти символов), базировался на «запланированном устаревании». Кодовая последовательность менялась каждый месяц, но французы обычно справлялись с нею за две недели — часто после всего лишь двух дней работы.

Однако количественная оценка методов шифрования стала возможной только благодаря пионерским идеям Клода Е. Шеннона (разд. 16.5). Пригодность различных методов шифрования все еще плохо понималась в годы Первой мировой войны, что показывает пример с Лемантеном, разоблаченным в 1914 г. и выдавшим, что французы читали сообщения немцев. Немецкий генштаб 18 ноября сделал внезапную и радикальную перемену в системе шифрования. Однако смена двойной столбцовой перестановки (разд. 6.2.4) на шифр Виженера с ключом ABC (разд. 8.1.2) и последующую простую столбцовую перестановку имела *иллюзорную сложность* для французских дешифровальщиков.

Очевидно, любовники, которые использовали для объяснения в любви шифрованные сообщения в «персональных» колонках британских газет примерно в 1850 г., испытывали полное доверие к своей системе шифрования. «Подсматривание» этих сообщений доставляло удовольствие для определенной части лондонского светского общества; она включала Чарльза Бэббиджа, а также Чарльза Вейтстоуна и Леона Плейфейра, который внедрил в одну такую переписку сообщение, зашифрованное тем же шифром, тем самым подсказывая реакцию одного из корреспондентов, молодой леди: «Дорогой Чарли: не пишите больше. Наш шифр раскрыт». Между прочим, несмотря на похвальное разъяснение Шеннона (а, возможно, и благодаря нему), шифрованные сообщения до сих пор, как можно заметить, являются регулярной принадлежностью «колонок объявлений о розыске родных».

Другая дама очень сильно была увлечена человеком, которому она дала зашифрованный рецепт изготовления золота; причем ключ к шифру знала она одна. Человек этот не только сообщил ей, что он дешифровал рецепт, но также и назвал ей ключевое слово. Должно быть, он волшебник, подумала она. Поскольку, очевидно, он мог читать ее мысли, было бы лучше отдать ему и ключ от ее сердца. Шел 1757 г., она была богатой мадам д'Юрфе, а кавалер (который от нее вскоре отказался) был Джакомо Джироламо Казанова, шевалье Сейнголт, чье усердие на ниве криптоанализа, очевидно, не достаточно хорошо известно.

Казанова

Мария Антуанетта тоже знала, как совмещать любовь с криптографией, знал это и король Эдуард VIII (позже герцог Виндзорский). Используют криптографию не только дипломаты и военные. Она применяется в частной и гражданской сферах, не говоря уже о коммерческой сфере с примерами, вроде ценового шифра книготорговца, упаковочной даты на масле (разд. 3.1.1), или маркировки на автомобильных шинах (рис. 18).

Существуют несколько забавных историй о предполагаемой невскрываемости шифров. Переоценка своей искусности — регулярный источник преимущества для противника. Такое преувеличение свойственно не только авторам шифров. Следующая история была рассказана о Пауле Шиллинге фон Канштатте, одном из изобретателей электромагнитного телеграфа (1832 г.): «для Русского министерства он составил такой секретный алфавит (так называемый *chiffre*), что в этом случае даже такой изобретательный черный кабинет[3)], как австрийский не мог бы взломать его и за пятьдесят лет» (Ф. П. Фонтон, позже А. В. Яроцкий). И уже в 1917 г. уважаемый журнал *Scientific American* объявил шифрование методом Виженера (разд. 7.4) невскрываемым.

Ирония состоит в том, что Этьен Базерье (1846–1931 гг.), крупный французский криптограф, который вскрыл целый ряд «невзламываемых» систем

[3)]Коллектив криптографов, занимавшихся перлюстрацией и расшифровкой дипломатической почты. До середины XIX века черные кабинеты существовали во многих европейских столицах.

(1) безкамерная;
(2) 175 — ширина шины в мм;
S — скорость (до 180 км/ч для летних шин);
R — радиальную усадку (опускается для шины с перекрещенными слоями);
14 — диаметр обода колеса в дюймах;
(3) TWI = индикатор износа протектора (шесть ребер, которые появятся в рисунке протектора, когда он сотрется на 1/16 дюйма);
(4,5) дополнительные маркировки для Европы:
88 — (кодированная) максимальная нагрузка на колесо;
S — вновь означает 180 км/ч;
(6) боковая сторона имеет два слоя фибры из вискозы;
(7) протектор имеет два стальных слоя и два слоя из вискозы;
(8) максимальное давление камеры для холодного времени (применяется только в США);
(9) максимальная нагрузка на колесо (применяется только в США);
(10) фирменный знак;
(11,12) проверено по Европейским стандартам;
4 — страна где происходило тестирование (в нашем случае это — Голландия);
(13) DOT = Department of Transportation (министерство транспорта США);
(14) коды изготовителя:
LM = фабрика; J3 = размер; MEB = тип;
344 = дата (34-я рабочая неделя 1974 г.).

Рис. 18. Система маркировки для автомобильных шин

шифрования, предлагаемых французской секретной службе, в свою очередь был достаточно самонадеян, чтобы поверить, что он изобрел абсолютно надежный метод шифрования (*je suis indéchiffrable*, см. рис. 19). Его антагонист маркиз Виари — кстати, первый современный криптолог, использующий математику — не получил ни малейшего удовольствия от реванша при взломе нескольких шифров, которые прислал ему Базерье (разд. 7.5.3, 14.3.1).

2.1.2. Изобретение и финансовая эксплуатация шифраторов и дешифраторов — прибыльная отрасль криптографии. До XIX в. это были механические устройства; с начала XX в. появились автоматизированные устройства, в середине столетия до них добралась электроника и позже микроэлектроника. В конце 1939 г. Конрад Цузе, 30-летний компьютерный пионер старой Гер-

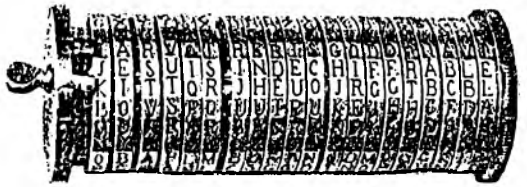


Рис. 19. Цилиндр Базерье с сообщением «je suis indéchiffrable»

мании, служивший в пехотной части на Западном валу, изобрел шифратор, подключаемый к телетайпу. Он не смог убедить немецкое военное министерство в преимуществах своего изобретения, в котором использовался принцип Вернама (разд. 8.3); ему дали понять, что ведомство уже обладает хорошими устройствами аналогичной конструкции. Имелись в виду машины Lorenz SZ 40 и Siemens T 52, но не ENIGMA, как Цузе недавно неправильно предположил.

Сегодняшний микрокомпьютер[4)], имея размеры, вес и цену карманного калькулятора, обладает характеристиками, сопоставимыми с лучшими шифраторами времен Второй мировой войны. Это восстанавливает значение некоторых использовавшихся ранее хороших методов, роль которых значительно понизилось после появления «больших» компьютеров в центрах дешифровки. Более того, обычный коммерческий микрокомпьютер себестоимостью порядка $100 (не имеется в виду персональный компьютер) может выполнять шифрование с применением более сложных алгоритмов по сравнению с теми, которые были реализованы в классических шифраторах.

При оценке метода, основанного на какой-либо документации или шифровальной аппаратуре, необходимо помнить, что эти объекты могут попасть в руки противника (принцип Шеннона, разд. 11.2.3). Микрокомпьютер, в который загружена программа или данные на магнитной карте, не обладает никакой собственной криптографической структурой, кроме, возможно, клавиатуры и дисплея.

В случае открытых ключей, распространяемых для связи по коммерческим каналам связи, известны даже методы шифрования и расшифровывания. Секретным остается только ключ расшифровки. Здесь, таким образом, принцип Шеннона «противник знает систему» доведен до крайности. В то же самое время, увеличивающееся использование открытых каналов связи приводит к проблеме аутотентификации, которая становится такой же декларируемой целью криптографии, как и секретность (разд. 10.5, 10.6).

2.1.3. Криптологические методы время от времени используются в литературе. Прихотливо сотканные литературные труды типа «*Zettels Traum*» Арно Шмидта (1970 г.) так и просятся, чтобы их «расшифровали». Особую проблему представляют якобы секретные сообщения. В романе «*Физиология брака*» (1829 г.) Оноре де Бальзак приводит такой пассаж:

«*La Bruyère a dit très spirituellement: «C'est trop contre un mari que la dévotion et la galanterie: une femme devrait opter». L'auteur pense que La Bruyère s'est trompé. En effet:* — — —.»

Затем следует беспорядочный набор символов, словно на странице была рассыпана наборная касса. Четыре издания книги, три из которых вышли при жизни Бальзака, фактически содержат четыре различных варианта этого текста. Автор, по-видимому, совершил розыгрыш над читателем. Однако, Базерье в 1901 г. исследовал такую криптограмму и обнаружил, что она не соответствует никакой известной схеме шифрования; это был *une facétie de l'auteur* — шутка творца.

[4)] Вторая половина 90-х годов XX века.

В 1878 г. много шума наделал Игнатиус Донелли. Это был американский провинциальный политический деятель и наделенный богатым воображением псевдоученый, который уже в то время спекулировал по поводу Атлантиды и столкновения Земли с метеоритом. Он приступил к поиску в работах Шекспира стеганографического доказательства того, что их автором был сэр Фрэнсис Бэкон (Георг Кантор, создатель современной теории множеств, также много лет охотился на эту химеру). Так, если вы возьмете достаточно длинный текст и объявите достаточное количество символов несоответствующими некоему вашему критерию и удалите их (возможно, также произведете перестановку оставшихся символов), тогда вы вполне сможете прочитать что-нибудь в оставшейся выжимке — например, это может быть гипотетическое сообщение лорда Байрона из разд. 1. Итак, Донелли, очевидно, преуспел на этом поприще, так как поток любителей хлынул в эти изыскания. Ни об одном из них нечего было бы сказать, не будь среди них некоего Уильяма Фредерика Фридмана[5)], который изучал генетику и был нанят богатым торговцем текстилем из Женевы близ Чикаго, полковником Джорджем Фабианом. Помимо лабораторий по биологии, химии и акустике Фабиан финансировал криптологов, которые, как предполагалось, должны были доказать, что Бэкон был Шекспиром. Фридман увлекся криптологией, а также молодой Элизабет Смит, криптологом, работающей у того же Фабиана. Он связал свою жизнь с обеими и стал наиболее успешным американским криптологом.

2.1.4. Официальные криптологические службы XX века имеют таинственно звучащие названия, пронизанные духом времени. Они обычно входят в состав секретных служб, занимающихся контрразведкой и внешней разведкой. Наиболее известны среди них MI-6, британская секретная разведывательная служба (S.I.S.), которая непосредственно подчиняется Министерству иностранных дел и Центральное разведывательное управление (C.I.A.) в США (с 1947 г.), которое вместе с Разведывательным управлением Министерства обороны (D.I.A) подчинено Разведывательному комитету США (U.S.I.B.) и, следовательно, контролируется законодательными и исполнительными органами. Послевоенная Германия имеет *Bundesnachnchtendienst* (BND), подчиняющуюся непосредственно ведомству канцлера.

Существующие криптологические службы, особенно отделы криптоанализа часто разделены по дипломатическим и военным сферам применения. В основе этого деления могут лежать веские организационные причины, но часто оно же препятствует обмену опытом. В Англии военного времени Адмиралтейство (O.I.C., Центр оперативной разведки) и Министерство иностранных дел (Отдел связи) вынуждены были тесно взаимодействовать ввиду отчаянного положения в 1940 г.; не обошлось без вмешательства Уинстона Черчилля, который создал мощный координирующий орган в виде Лондонской секции управления (L.C.S.), возглавляемой полковником Джоном Генри Бивaном и подчинявшейся непосредственно премьер-министру. В пределах MI-6

[5)] Родился Фридман в Кишиневе (Молдавия) 24 сентября 1891 г. и при рождении получил имя Вольф. На следующий год семейство эмигрировало в США. Он умер 2 ноября 1969 г. и был захоронен на Арлингтонском мемориальном кладбище в Вашингтоне.

ответственность за дешифровку возлагалась на G.C.H.Q. (штаб правительственной связи), имевшего различные вывески. Вот некоторые из них с историческими значениями: G.C. & C.S. (Government Code and Cyper[6]) School — Правительственная школа кодов и шифров), War Station (Военная станция), Station X (станция X), Room 47 Foreign Office (комната 47 Министерства иностранных дел); он также часто назывался B. P. (Блетчли Парк), по названию места, где он размещался с 1939 г. Даже в пределах B. P. существовало некоторое традиционное разделение между AI (воздушная разведка) и MI (военная разведка), с одной стороны, и Navy (военно-морскими силами). Обе стороны опирались на свои успехи в Первой мировой войне, достигнутые M.I.1 (b) (отделом военной разведки Военного министерства) и комнатой 40 Адмиралтейства. В послевоенное время G.C.H.Q. располагался в Чельтенхэме.

После того, как в 1917 г. Соединенные Штаты вступили в Первую мировую войну, стало необходимым быстрое расширение американской военной криптологии. Как часть A.E.F. (американские экспедиционные силы) создаются отделы G.2 A.6 (генеральный штаб, разведывательное управление, военно-информационный отдел, отдел радиоразведки) и отдел формирования кодов службы связи под контролем MI-8 (криптологический отдел разведуправления военного министерства), возглавляемый Гербертом Осборном Ярдли (1889–1958 гг.). Конкуренция между армией и военно-морским флотом продолжалась и во время Второй мировой войны: OP-20-G было флотской криптологической службой с криптоаналитическим отделом OP-20-GY, в то время как SIS (разведывательная служба связи) была аналогичной структурой в армии, которую построил Ярдли и которая начиная с 1929 г. возглавлялась Уильямом Фридманом. Опыт, полученный во Второй мировой войне привел к концентрации ресурсов в рамках G.2: Агентство безопасности связи армии в 1945 г. было слито с криптоаналитическим отделом службы связи, чтобы породить A.S.A. (Агентство безопасности армии), которое в 1949 г. преобразовалось в A.F.S.A. (Агентство безопасности вооруженных сил) и в 1952 г. — в N.S.A. (Агентство национальной безопасности), которое подчинялось министру обороны и возглавлялось в 1977–1981 гг. легендарным Бобби Рей Инмэном. Важными его подразделениями являются Агентство безопасности министерства обороны и I.D.A., институт исследований министерства обороны, который является более открытой организацией и может поддерживать отношения с некоторыми университетами. N.S.A. расположен в форте Джорджа Г. Меада в Мэриленде.

В немецком *Рейхе* криптологические службы также были разделены: *Auswärtiges Amt* (министерство иностранных дел), с одной стороны, и армия и военно-морской флот, с другой. Дальнейшее соревнование во время Второй мировой войны продолжили *Reichsluftfahrtministerium u Sicherheitsdienst* (СД, «служба безопасности»). Будущий генерал Эрих Фельгебель (1886–1944 гг.) принял на вооружение в 1934 г. единую шифровальную машину ENIGMA; но координация всей военной криптографии, которую постоянно требо-

[6] Cyper — устаревшая форма слова chiper, все еще употребляемая в Великобритании.

вал OKW/Chi (шифровальное управление высшего командования Вермахта), вплоть до осени 1943 г. блокировалась Риббентропом, Герингом и Гиммлером. Когда координация WNV/Chi (*Wehrmachtnachrichtenverbindungen Chiffrierwesen*) по указанию фюрера была, наконец, достигнута в ноябре 1944 г., все нити твердо держал в своих руках Вальтер Шелленберг (1910–1952 гг.), который был честолюбив, всегда симулировал преданность своим лидерам Гиммлеру и Гитлеру, и дослужился до звания генерала СС. Он умер от болезни печени, отсидев только шесть лет своего тюремного срока, полученного от Нюрнбергского трибунала; затраты на его похороны оплатила дизайнер Коко Шанель.

В послевоенной Германии в 1953 г. авторитеты криптологии собрались в Бад Гёдесберге вблизи Бонна, в организации, называвшейся *Bundesstelle für Fernmeldestatistik* (федеральное бюро статистики связи), что было некоторым преуменьшением. На самом деле это было криптоаналитическое подразделение BND; и истинное название его было *Zentralstelle für das Chiffrierwesen* (центральное бюро криптологии). В 1990 г. было произведено преобразование («*Amt für Militarkunde*»), выделившее BSI, которое стало иметь дело с вопросами открытой криптографии.

Во Франции 2^{bis} (номер улицы на Авеню Трувиль) было кодовым именем (*nom de guerre*) для S.R. (*Service de Renseignement*) с криптоаналитическим бюро (*section de transmission et decryptement*). Шведское криптоаналитическое агентство было известно под аббревиатурой FRA (*Forsvarets Radioanstalt*), а в Италии аналогичная служба называлась SIM (*Servizio dell'Informazione Militare*). В Японии *Tokumu Han* (отдел разведки) было названием группы криптоанализа отдела информации адмиралтейства, созданной в 1925 г., а *Ango Kenkyu Han* (отдел исследования шифров) именовало такую же группу в министерстве иностранных дел.

Криптоаналитическая и криптографическая служба России была организована по приказу В. И. Ленина в 1921 г. и некоторое время возглавлялась Л. Д. Троцким.

2.2. Шифрование

Подведем итог: криптология является наукой о (явном) секретном письме (криптография), негласном расшифровании (криптоанализ) и правилах, которые в свою очередь должны сделать этот криптоанализ более трудным (безопасность шифрования).

2.2.1. Словарь, алфавит. Множество V символов, используемых для формирования открытого текста (plain text)[7] образует словарь открытого текста или алфавит открытого текста. Множество W символов, используемых для формирования шифротекста (используется также термин кодотекст) образует словарь шифротекста или алфавит шифротекста. Отдельные символы в W могут также быть знаками, т. е. специальными символами, представляющими слово или фразу, типа &, %, \$, £, ©; возможны также стенографические сим-

[7] В отличие от cleartext, который означает текст, переданный без шифрования.

волы. Множества V и W могут быть различными, частично пересекающимися или совпадающими множествами.

2.2.1.1. Пусть V^* и W^* обозначают множества конечных слов, созданных из символов множеств V и W соответственно (слова открытого текста и слова шифротекста), ε означает пустое слово (в обоих множествах). $Z^n \subset Z^*$ обозначает множество всех слов длины n; $Z^{(n)}$ обозначает $\{\varepsilon\} \cup Z^1 \cup Z^2 \cup Z^3 \cup \ldots \cup Z^n$. Множество V^* называется пространством открытых текстов, W^* — пространством шифротекстов.

2.2.1.2. Во всех практических случаях V и W являются непустыми конечными множествами. Теоретически, однако, мы могли бы рассматривать счетные множества V и W; тогда V^n и W^m также будут счетными.

2.2.2. Шифрование и расшифрование. Шифрование определяется как отображение $\mathbf{X}\colon V^* \to W^*$. Обратное отображение $\mathbf{X}^{-1}\colon V^* \leftarrow W^*$ (определяемое как $x \leftarrow y$, тогда и только тогда, когда $x \to y$) в этом случае называется расшифрованием.

2.2.2.1. Законный получатель шифрованного сообщения должен быть в состоянии восстановить исходное сообщение однозначным образом. Шифрование поэтому, как правило, *инъективно*, т. е. однозначно справа налево (левооднолистно):

$$(x \mapsto z) \wedge (y \mapsto z) \Rightarrow (x = y).$$

Для слова $x \in V^*$ определим *слой* как множество $\mathcal{H}_x = \{y \in W^*\colon x \overset{\mathbf{X}}{\mapsto} y\}$.

Как правило, к шифрованию $\mathbf{X}$ также предъявляется требование, чтобы оно было *полным*; иными словами множество $\mathcal{H}_x$ должно быть непустым для всех слов $x \in V^*$.

2.2.2.2. Шифрование $\mathbf{X}\colon V^* \to W^*$ реализуется посредством многозначного Гильбертова недетерминированного «оператора выбора» η, где $\mathbf{X}(x) = \eta\mathcal{H}_x$. Элементы $\mathcal{H}_x$ (предполагается, что существует более одного элемента) называются вариантами или омофонами x. Таким образом, варианты — это различные слова шифротекста, получаемые из одного и того же слова открытого текста при шифровании $\mathbf{X}\colon V^* \to W^*$.

Если отображение $V^* \to W^*$ также однозначно слева направо (правооднолистно), т. е. $\mathcal{H}_x$ содержит не более одного элемента для каждого $x \in V^*$, тогда *оператор* шифрования $V^* \to W^*$ является функцией $V^* \to W^*$, а если он, кроме того, еще и сюръективен, то он становится взаимно однозначной функцией $V^* \leftrightarrow W^*$.

2.2.2.3. Как правило, $\varepsilon \overset{\mathbf{X}}{\mapsto} \varepsilon$. Если множество $\mathcal{H}_\varepsilon$ содержит элементы, отличные от ε, которые, очевидно, являются омофонами для $\varepsilon \in V^*$, то эти элементы называются пустыми или фиктивными текстами.

Заметим, что множество всех шифрований $V^* \to W^*$ (в случае фиксированных непустых алфавитов V и W) является несчетным.

2.2.3. Индуктивные определения. Будем говорить, что шифрование $\mathbf{X}\colon V^* \to W^*$ конечно, если множество всех пар слов $\{(x, \mathbf{X}(x))\colon x \in V^*\}$ конечно. Тогда для каждого натурального числа n найдется подходящее натуральное m такое, что $\mathbf{X}\colon V^{(n)} \to W^{(m)}$.

Но как определить и специфицировать отображение $V^* \to W^*$? Даже если оно и конечно, нереально перечислить все его пары. Поэтому часто используются индуктивные правила. Их мы изучим в следующем параграфе.

2.3. Системы шифрования

Определим систему шифрования M, как непустое *конечное* (как правило) множество $\{\chi_0, \chi_1, \chi_2, \dots, \chi_{\theta-1}\}$ (инъективных) отображений χ_i: $V^{(n_i)} \to W^{(m_i)}$. Каждое χ_i называется шагом шифрования. Систему шифрования вместе с соответствующей системой расшифрования будем называть *криптосистемой*.

Шифрование $\mathbf{X} = [\chi_{i1}, \chi_{i2}, \chi_{i3}, \dots]$ называется конечно порожденным (посредством системы шифрования M), если оно *индуцировано* конкатенацией $.\star.$ элементов последовательности (конечной или бесконечной) $(\chi_{i1}, \chi_{i2}, \chi_{i3}, \dots)$, состоящей из шагов шифрования $\chi_i \in M$, т.е. $x \overset{\mathbf{X}}{\mapsto} y$ для $x \in V^*$, $y \in W^*$, тогда и только тогда, когда существуют разложения $x = x_1 \star x_2 \star x_3 \star \dots \star x_k$ и $y = y_1 \star y_2 \star y_3 \star \dots \star y_k$ с отображениями[8)]

$$x_j \overset{\chi_{i_j}}{\longmapsto} y_j \quad \text{для} \quad j = 1, 2, \dots, k.$$

Пример:

χ_i: $V^{(n_i)} \to V^{(n_i)}$: циклическая перестановка n_i элементов ($\theta = 4$); $n_1 = 3$, $n_2 = 5$, $n_0 = 2$, $n_3 = 6$

$$\begin{array}{c} \underline{\mathrm{n\,e\,a}}\,\underline{\mathrm{r\,l\,y}}\,\underline{\mathrm{e\,v\,e}}\,\underline{\mathrm{r}}\,\underline{\mathrm{y}}\,\underline{\mathrm{i\,n\,v\,e\,n}} \\ \mathrm{e\,a\,n\,l\,y\,e\,v\,r\,r\,e\,i\,n\,v\,e\,n\,y} \end{array} \quad (\chi_1, \chi_2, \chi_0, \chi_3)$$

2.3.1. Основные понятия. Пусть $\theta = |M|$ обозначает число шагов системы шифрования. Шаг шифрования χ_i: $V^{(n_i)} \to W^{(m_i)}$ является порождающим отображением; число n_i называется (максимальной) шириной открытого текста шифрования, а число m_i — (максимальной) шириной шифра χ_i. Отображение χ_i может быть недетерминированным. Шаг шифрования называется эндоморфным, если $V = W$.

Если говорить об омофонах и вариантах (эквивалентные термины: произвольные замены, множественные замены) и пустых текстах (или фиктивных; фр. *nonvaleurs*, нем. *Blender*, *Blindsignale*), то в большинстве случаев предполагается их наличие на шагах шифрования. Если словарь шифротекста шага шифрования содержит слова различной длины, то он именуется шагом шифрования «с разбросом» (англ. *straddling*, нем. *gespreizt*).

Шифрование не обязательно инъективно, даже если порождающие шаги шифрования являются таковыми. Пусть

$$\begin{aligned} \mathrm{a} &\mapsto \cdot - \\ \mathrm{i} &\mapsto \cdot\cdot \\ \mathrm{l} &\mapsto \cdot - \cdot\cdot \end{aligned}$$

[8)] Каждое $x \in V^*$ берется с соответствующим заполнением бессмысленными символами.

принадлежат инъективному отображению $V^1 \to W^{(4)}$, тогда в порожденном отображении $V^* \to W^*$ имеем

$$\text{ai} \mapsto \cdot - \cdot \cdot \quad \text{и} \quad \text{l} \mapsto \cdot - \cdot \cdot ,$$

что означает нарушение инъективности (вызванное, например, неряшливостью радиооператоров).

2.3.2. Шифрование и кодирование. Шаг шифрования $\chi_i\colon V^{(n_i)} \to W^{(m_i)}$ является конечным, если только V и W конечны; в принципе его можно определить перечислением (в виде таблицы шифрования). Такое перечисление часто называется шифром (фр. *chiffre*) или кодом, в последнем случае шаг шифрования называется шагом кодирования. Границы между терминами «шифр», «шифровать», «дешифровать» и «код», «кодировать», «декодировать» расплывчаты и по существу определяются историческими аспектами их использования (см. также разд. 4.4). Термины «шифр», «код» и, в более общем смысле, «крипто» также используются для элементов $W^{(m_i)}$.

2.3.2.1. Шифрование $\mathbf{X} = [\chi_{i_1}, \chi_{i_2}, \chi_{i_3}, \ldots]$, конечно порожденное системой шифрования M, является *одноалфавитным*, если оно включает или использует единственный шаг шифрования («алфавит»). В противном случае оно называется *многоалфавитным*. (Если M использует единственный алфавит ($\theta = 1$), то каждое шифрование, порожденное посредством M, одноалфавитно.)

2.3.2.2. Конечнопорожденное шифрование называется *односимвольным*, если на всех использованных шагах шифрования $n_i = 1$ и *многосимвольным* в противном случае. В частном случае, представляющем интерес для автоматического шифрования, все шаги шифрования из M имеют максимальную ширину n, которая равна максимальной ширине шифра m. В этом случае система шифрования M обязательно конечна. Если для всех $\chi_i \in M$ выполняется

$$\chi_i\colon V^{(n)} \to W^{(m)}$$

(т. е. отсутствуют шаги шифрования с разбросом), то $[\chi_{i_1}, \chi_{i_2}, \chi_{i_3}, \ldots]$ называется *блочным* шифрованием; слова из V^n именуются блоками шифрования. В соответствующем словаре символьных блоков длины n (используют также термин n-кортежи) блочное шифрование теоретически можно интерпретировать как однобуквенное шифрование. Системы шифрования с $\chi_i\colon V^{(n)} \to W^{(m)}$ для $n = 2, 3, 4$ определяют шифрования биграммами, триграммами, тетраграммами, которые для $m = 1, 2, 3$ называются однодольными, двудольными, трехдольными (фр. *bifide*, *trifide*) шифрованиями. Часто берутся $V = W$ и $m = n$, что дает нам блочное шифрование (в узком смысле).

2.3.3. Текстовые потоки. Поток $(z_1, z_2, z_3, \ldots)$ представляет собой бесконечную последовательность блоков, состоящих из символов.

Поток открытого текста является бесконечной последовательностью блоков $(p_1, p_2, p_3, \ldots)$, где $p_j \in V^n$; соответственно бесконечная последовательность блоков $(c_1, c_2, c_3, \ldots)$, где $c_j \in W_m$, является потоком шифротекста.

Поточное шифрование является блочным шифрованием сегментов потока открытого текста в сегменты аналогичного потока шифротекста.

Шифрование $\mathbf{X} = [\chi_{i_1}, \chi_{i_2}, \chi_{i_3}, \ldots]$, конечно порожденное системой шифрования M, называется *периодическим* (с повторяющимся ключом) или *непериодическим* («апериодическим», с бегущим ключом) в зависимости от того, является ли бесконечная последовательность $(\chi_{i_1}, \chi_{i_2}, \chi_{i_3}, \ldots)$ периодической или нет.

Односимвольное шифрование является, очевидно, периодическим. Поэтому всякое непериодическое (с бегущим ключом) шифрование обязательно имеет многосимвольный алфавит. Ему мы уделим большое внимание в разд. 8.7. Каждое периодическое блочное шифрование с периодом r можно теоретически интерпретировать как односимвольное шифрование с единственным шагом шифрования

$$\chi_0 \colon V^{n \cdot r} \to W^{m \cdot r}.$$

Иная ситуация имеет место для шифрований с бегущим ключом. Такие шифрования принадлежат к категории существенно более мощных методов. Имеется взаимно однозначное соответствие между последовательностью $(\chi_{i_1}, \chi_{i_2}, \chi_{i_3}, \ldots)$, $\chi_i \in M$, и вещественным числом, представленным в системе счисления по основанию θ дробью $0.i_1 i_2 i_3 \ldots$ Для фиксированной системы шифрования M счетное подмножество рациональных чисел соответствует множеству периодических блочных шифров; множество непериодических (с бегущим ключом) шифрований, таким образом, соответствует несчетному множеству иррациональных чисел, лежащих между 0 и 1.

Современным примером одноалфавитного, многосимвольного блочного шифрования является криптосистема на основе стандарта DES, метод блочного шифрования (и расшифрования), распространяемый, начиная с 1977 г., Национальным бюро стандартов США; шаг шифрования (в режиме ECB) в нем является взаимно однозначным эндоморфным шифрованием, выбранным из 2^{56} возможных (длина ключа 56, разд. 9.6.1.1) перестановок $V^8 \leftrightarrow V^8$ в словаре $V = Z_2^8$ из 256 различных 8-разрядных двоичных слов. Шаг шифрования такого объема невозможно задокументировать перечислением, но можно определить алгоритмически.

Имеется пример многоалфавитного, многосимвольного шифрования, которое не принадлежит классу блочных шифров: открытый текст зашифровывается слово за словом с использованием нескольких кодовых книг в некотором периодически или непериодически изменяющемся порядке. Этот метод не очень подходит для компьютеризированной криптографии. А вот блочное многоалфавитное мносимвольное шифрование — прерогатива современных компьютеров.

2.4. Многозначность

Использование омофонов и пустышек стало стандартным в криптографии с 1400 г. Примерно в 1500 г. криптографы начали использовать шифрование со словами шифротекста различной длины. Самое позднее около 1580 г.

папские криптографы Джованни Баттиста Ардженти и его племянник Маттео Ардженти осознавали значение левосторонней единственности для шагов шифрования с разбросом (разд. 3.4). Современное условие Фано («никакое слово шифротекста не может быть началом другого слова шифротекста») является достаточным условием этого, и оба Ардженти несомненно были знакомы с ним. Для шагов шифрования без разброса границы слов и соответствующие правосторонние разложения могут быть найдены прямым подсчетом.

2.4.1. Полифоны. Они беззастенчиво используются в английском языке, когда, например, фонемы \ā\ в \brāk\ и \ē\ в \frēk\ печатаются как «ea». В криптографии многозначные шаги шифрования, когда несколько слов открытого текста заменяются одним и тем же словом шифротекста, нарушают инъективность и крайне редки.

Шифр SA, код, использованный британским Адмиралтейством в 1918 г. (разд. 4.4.3, рис. 37) и шифр Дюшес де Берри, который использовал для подстановки алфавит LEGOUVERNEMENTPROVISOIRE (разд. 3.2.5) представляют собой такие крайне редкие примеры подобной многозначности.

На практике обычно имеющейся семантической информации достаточно, чтобы избежать неоднозначности при расшифровании, скажем, «Disel oil», «Corporal» и «Paris» (дизтопливо, капрал, Париж) или «runway», «General», и «ground fog» (взлетная полоса, генерал, наземный туман), если они являются полифонами. Идея использования полифонов, кажется, приходит любителям чаще, нежели профессионалам. В Англии любовная пара в 1853 г. подкинула Бэббиджу крепкий орешек в виде многозначного цифрового шифра, в котором, например, 1 означает буквы t и u, 8 — буквы h и i, 2 — буквы m и o, 4 — буквы e и r. Сообщение начиналось с

1821 82734 29 30 84541,

которое (с учетом двух ошибок при шифровании) означало: «Вы являетесь отражением моей души». Создается впечатление, что любовная пара получала особое удовольствие от ненужной сложности своего шифра.

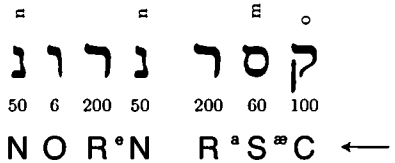


Рис. 20. Значение 666, связанное с написанием на Иврите имени Cœsar Nero (любезно предоставлено Ральфом Штейнбрюггеном)

Однако, многозначные шифры использовались в древних цивилизациях между Нилом и Евфратом. Поскольку буквы алфавита также служили и числовыми символами, было популярным развлечением сложить значения символов, представляющих секретное слово (*gematria*). Таким образом, *изопсефон* (*isopsephon*) 666, упоминаемый в Апокалипсисе (Стих. 13.18), был использован для обозначения императора Нерона (рис. 20). Говорят, имеются люди, которые отказываются принимать автомобильный номер, содержащий так называемое «число Зверя» 666.

С точки зрения привычных европейских языков, арабский алфавит (без гласных) также многозначен. Загадка значения «Pthwndxrclzp» в романе

```
G X Y Y S X D B R Z Z B G B B G S I C U
H Z Q X R V P I Y D L D L C C N O U H S
I A R V O T R E B I S G O D D F N A V T
J E S U I S I N D E C H I F F R A B L E
K I T T Q R J H E U O J R G G T B C B L
L O V S P Q U U T P U K E J H H C F D A
M U X R N P G R S R R N M K K U D G F C
N Y Z Q M N V X L O A P T L M B F J G F
O B A P L M B L F T N Q D M O C G K I B
```

Рис. 21. Несколько многозначных текстов для одной настройки цилиндра Базерье

Джеймса Джойса «*Поминки по Финнегану*» сохраняла актуальность для историков литературы (и владельцев похоронных бюро) много лет.

С технической точки зрения, цилиндр Базерье (см. рис. 19), рассматриваемый далее в разд. 7.5.3, оперирует и с омофонами и полифонами. Однако иньективность в нем эффективно поддерживается, потому что «запрещенные» многозначные тексты почти всегда бессмысленны (рис. 21). Многозначность может также быть причиной трудностей при использовании некоторых способов криптоанализа. Многозначные открытые тексты, порожденные одним и тем же шифротекстом, называются вариантами.

2.4.2. Межсловный пробел и пунктуация. Одним из основных правил классической профессиональной криптографии (т. е. «формальных шифров») является подавление межсловного пробела и пунктуации, что приводит, строго говоря, к неоднозначности. В некоторых случаях подлинная неоднозначность может иметь место, если позиция границы между словами неопределена; например, «темно там учитель» и «темнота мучитель»[9]. Предложения

«Пять пальцев имею я на каждой руке десять на всех»,
«Казнить нельзя помиловать»

также допускают различные толкования, в зависимости от пунктуации; но лишь одно толкование имеет логический смысл.

Однозначность часто нарушается, когда имеются недостаточная контекстная информация:

«the captive flies»

может означать как «пленного летчика», так и «пойманную муху». Фраза

«two thousand year old horses»

имеет даже три различных толкования: двухтысячелетние лошади, две тысячелетние лошади, две тысячи годовалых лошадей.

[9]У автора примеры приведены на английском языке: «dark ermine» и «darker mine» («темный горностай» и «более темная шахта»). Мы позволили себе привести аналогичный пример на русском языке. В дальнейшем тексте мы будем, где это уместно, придерживаться этой практики. — *Прим. перев.*

Вот другой пример, где подавление дефисов может вызывать затруднение:

«a man eating fish — a man-eating fish» («человек, едящий рыбу — рыба-людоед»).

2.5. Множества символов

Мы используем N для обозначения $|W|$, конечной мощности множества символов открытого текста или шифротекста W. Так как случай $N=1$ не несет никакой информации, будем считать, что $N \geqslant 2$. Алфавит — это линейно упорядоченное множество символов.

2.5.1. Множества символов открытого текста. Их использование при шифровании зависит от языка и эпохи. Например, для гавайского языка множество символов исчерпывается набором

$$Z_{12} = \{\mathrm{a},\ \mathrm{u},\ \mathrm{i},\ \mathrm{o},\ \mathrm{e},\ \mathrm{w},\ \mathrm{h},\ \mathrm{k},\ \mathrm{l},\ \mathrm{m},\ \mathrm{n},\ \mathrm{p}\}.$$

В средние века, согласно традиции латинского языка, 20 букв, по-видимому, было достаточно для большинства писателей, включая Джованни Баттиста Порта в 1563 г. (рис. 23 в гл. 3):

$$Z_{20} = \{\mathrm{a},\ \mathrm{b},\ \ldots\ ,\ \mathrm{i},\ \mathrm{l},\ \ldots\ ,\ \mathrm{t},\ \mathrm{v},\ \mathrm{z}\}.$$

Часто в алфавит включались буквы /k/, /x/ и /y/, или лишь /x/ и /y/ (Порта в другие годы), /w/ долгое время писалось как /vv/, таким образом резервируя место для /&/, как на диске Леона Баттиста Альберти в 1466 г. (рис. 26 в гл. 3). С 1600 г. алфавит из 24 букв стал европейским стандартом:

$$Z_{24} = Z_{20} \cup \{\mathrm{k},\ \mathrm{w},\ \mathrm{x},\ \mathrm{y}\}$$

с использованием /v/ и для /u/. Тритемий (в 1508 г.) задействовал букву /w/ (см. рис. 52). Во французской транскрипции 1561 г. (Габриель де Коланж), «немецкая» буква /w/ была заменена на /&/, согласно Эйрауду (Eyraud).

В XVIII столетии добавляется буква /u/:

$$Z_{25}^{uw} = Z_{24} \cup \{\mathrm{u}\}.$$

Однако, когда была затребована буква /j/ (в французском языке, например), то вновь пожертвовали буквой /w/ (Базерье, 1891):

$$Z_{25}^{ju} = Z_{20} \cup \{\mathrm{j},\ \mathrm{k},\ \mathrm{u},\ \mathrm{x},\ \mathrm{y}\}.$$

Буквы /j/, /k/, /w/, /x/, /y/ были весьма необычны для итальянцев, а буквы /k/, /w/ — для французов. Ирландцы вполне могут обходиться без букв /j/, /k/, /q/, /v/, /w/, /x/, /y/, /z/.

Приблизительно с 1900 г. ныне существующий алфавит

$$Z_{26} = Z_{24} \cup \{\mathrm{j},\ \mathrm{u}\}$$

стал использоваться повсеместно. Но даже в Центральной Европе имеются исключения. Во время Второй мировой войны чешское правительство в изгнании использовало расширенное множество символов, состоящее из 31 буквы и 13 числовых и некоторых других символов:

$$Z_{44} = \{\text{a, b, c, č, d, e, ě, f, } \ldots \text{, r, ř, s, š, t, } \ldots \text{, z, ž, ·, ', *, 0, 1, } \ldots \text{, 9}\}.$$

Итальянский шифр *cifrario tascabile* времен Первой мировой войны использовал множество символов

$$Z_{36} = Z_{26} \cup \{0, 1, \ldots, 9\}.$$

Современный кириллический алфавит имеет 32 буквы (игнорируется буква Ё):

$$Z_{32} = \{\text{А Б В Г Д Е Ж З И Й К Л М Н О П}$$
$$\text{Р С Т У Ф Х Ц Ч Ш Щ Ъ Ь Э Ю Я}\}.$$

Для представления цифр и, в случае необходимости, знаков препинания и диакритических знаков применялось множество разных специальных соглашений.

В профессиональной криптографии пробелы между словами подавляются. Даже в немецком языке, где слова длиннее, чем в большинстве других языков, пробел встречается чаще буквы /e/.

2.5.2. Множества технических символов. Используемые алфавиты шифротекста обычно определяются техническими ограничениями; помимо упомянутых выше алфавитов, имеются следующие технические множества символов:

$Z_{256} = Z_2^8$	(байты; IBM приблизительно в 1964 г.);
$Z_{32} = Z_2^5$	(Фрэнсис Бэкон 1605, 1623 гг., Бодо 1874 г.);
$Z_{10} = \{0, 1, 2, \ldots, 9\}$	(десятеричный);
$Z_6 = \{\text{A, D, F, G, V, X}\}$	(шестеричный; эти символы соответствуют легко различимым символам в азбуке Морзе);
$Z_4 = \{1, 2, 3, 4\}$	(четверичный; Альберти 1466 г., Карамуэль (Caramuel) 1670 г., Вейгель 1673 г.);
$Z_3 = \{1, 2, 3\}$	(троичный; Тритемий 1518 г., Уилкинс 1641 г., Фридеричи 1685 г.);
$Z_2 = \{\text{O, L}\}$	(двоичный; Фрэнсис Бэкон 1605, 1623 гг.);

кроме того, встречаются и другие символы, популярные среди любителей (см. разд. 3.1.1).

Бинарные[10], троичные, пятеричные, и десятеричные шифры имеют алфавиты $W \mathrel{\hat{=}} Z_2$, $W \mathrel{\hat{=}} Z_3$, $W \mathrel{\hat{=}} Z_5$, $W \mathrel{\hat{=}} Z_{10}$ соответственно.

[10]Бинарный в смысле билитеральный (множество символов состоит из двух «битов»: Бекон 1605, 1623 гг.; с явными значениями a=1, b=2, c=4, d=8 и т. д. (abfg=99): Непер 1617 г.; позиционную систему, использующую цифры применяли Харриот не позже 1621 г., Карамье 1670 г., Лейбниц 1679 г.

1	2	3
4	5	6
7	8	9

Рис. 22. Цифровой шифр в картографической сетке

2.5.2.1. В военно-морском флоте Германии (*Kriegsmarine*) для шифрования географических координат использовались девять цифр 1, 2, 3, 4, 5, 6, 7, 8, 9 (в случае необходимости, итеративно) (рис. 22).

2.5.2.2. Шифротекст часто записывают в виде пятисимвольных групп. Эта традиция берет свое начало в тарифных правилах Международного телеграфного союза, который с 1875 г. ограничил длину слова при передаче телеграмм десятью символами (что наложило серьезные ограничения на использование кодов). В 1904 г. допускались коды до десяти символов; позже телеграммы обычно передавались пятисимвольными группами (Вайтлоу: 20000 удобопроизносимых пятисимвольных групп дают 400 миллионов десятисимвольных групп).

2.5.3. В отображении $\mathbf{X}$: $V^* \to W^*$, криптоаналитик не знает ни V, ни $\mathbf{X}$. Однако по множеству $\mathbf{X}(V^*)$ фактически встречающихся криптослов он может иногда определить использованный метод шифрования (например, квадрат Полибия, см. разд. 3.3.1).

2.5.4. Эндоморфный случай. Часто имеет место случай, когда открытый текст и шифротекст используют одинаковый алфавит ($V = W$, в этом случае $\mathbf{X}$ является эндоморфизмом). В теоретических криптологических работах нашего времени в этом случае принято записывать символы открытого текста в нижнем регистре (маленькие буквы), а символы шифротекста — заглавными буквами. Курсивные символы верхнего регистра используются для так называемых ключевых символов. Отметим, что на диске Альберти (рис. 26) использовано противоположное соглашение. Даже в книге Ланге-Судара (рис. 27, на котором изображена линейка Сан-Сира) открытый текст дан в верхнем регистре, а шифротекст в нижнем.

2.6. Ключи

Ключ (фр. *clef*, *clé*, нем. *Schlüssel*) служит для выбора шага системы шифрования M. Ключи дают возможность корреспондентам менять шифрование в соответствии с предварительно установленными правилами, например, каждый день или после каждого сообщения или после каждого символа. Часто ключи организованы таким образом, что позволяют произвести индивидуальный шаг шифрования, следуя лишь простым правилам. Комбинаторная сложность метода шифрования определяется числом ключей, доступных в этом методе. Ключевая техника очень многообразна, и будет рассматриваться отдельно для индивидуальных классов методов шифрования. Пусть K обозначает множество символов ключей или ключевой словарь. K^* называется

ключевым пространством. Пусть $k_j \in K$ означает j-й ключ, использованный в последовательности шифрований; тогда k_j определяет номер s_j и, таким образом, и шаг кодирования $\chi_{s_j} \in M = \{\chi_0, \chi_1, \chi_2, \cdots, \chi_{\theta-1}\}$.

2.6.1. Ключи должны меняться. Повторное применение одного и того же ключа эквивалентно использованию системы шифрования с единственным элементом. В профессиональной криптографии почти никогда не используется такое фиксированное шифрование. Использование одной и той же кодовой книги на протяжении нескольких лет типично для дипломатических кругов. Впрочем, дипломатов многих стран едва ли можно расценивать как профессионалов в области шифрования: в различные времена в городах, подобных Вене, интенсивно действовал подпольный рынок дипломатических кодов. Советский Союз имел специфическую репутацию похитителя кодовых книг. В 1936 г. русский агент в Харлеме (Нидерланды) использовал выкраденную кодовую книгу, чтобы дешифровать телеграммы между японским военным атташе в Берлине и его правительством в Токио. В начале Первой мировой войны, вероятно, каждая европейская держава обладала копиями одной или нескольких американских дипломатических кодовых книг. В августе 1941 г. Лорис Жерарди тайно добыл для *Servizio Informazione Militare* (итальянская военная криптоаналитическая служба) копию кода BLACK, используемого американскими военным атташе. Известна история, рассказанная Алленом Даллесом, об американском после в Румынии, изгнанном политическом деятеле, подобно многим дипломатам, который не желал сообщать о потере своей кодовой книги. Он ждал, пока накапливалось несколько сообщений, и затем садился в поезд на Вену, чтобы дешифровать их в тамошнем посольстве. Мораль: даже кодовые книги должны регулярно заменяться, ежемесячно, в случае необходимости.

Ключи, используемые для выбора метода шифрования, являются предметом взаимного соглашения. Если одна партия ключей отработана, то передача новых ключей связана с риском, трудна или даже невозможна. В таких случаях часто выбираются ключи, составленные из невинных наборов букв или цифр, взятых из популярных романов, статистических отчетов, телефонных книг и т. д. — почти все использовалось в свое время, от «*Бравого солдата Швейка*» Ярослава Гашека в 1935 г. до статистического ежегодника немецкого Рейха. Но даже эта система уязвима — если источник ключей обнаружен, то целый поток сообщений одним махом становится прозрачным.

2.6.2. Блоки. Пусть, в соответствие с обозначениями разд. 2.3, $\mathbf{X}$ означает конечно-порожденный блочный шифр, $\mathbf{X} = [\chi_{s_1}, \chi_{s_2}, \chi_{s_3}, \ldots]$, где $\chi_{s_j}: p_j \mapsto c_j$. $(p_1, p_2, p_3, \ldots)$, где $p_j \in V^n$, обозначает последовательность открытого текста; $(c_1, c_2, c_3, \ldots)$, где $c_j \in W^m$, обозначает последовательность шифротекста; $(k_1, k_2, k_3, \ldots)$, где $k_j \in K$, обозначает ключевую последовательность.

Пусть k_j означает ключ, который определяет χ_{s_j}, а S_j — оператор, означающий действие $\chi_{s_j}(\cdot)$. Тогда мы имеем три вида обозначений для криптографического уравнения:

$$c_j = \chi_{s_j}(p_j) \quad \text{или} \quad c_j = \mathrm{X}(p_j, k_j) \quad \text{или} \quad c_j = p_j S_j.$$

Заметим, что χ_i означает i-й шаг шифрования в нумерованном списке шагов, в то время как χ_{s_j} является шагом, используемым для выполнения j-го шага шифрования (из списка порождающих систему шифрования шагов). Если χ_{s_j} является *инъективной функцией*, что обычно имеет место, то существует обратная функция $\chi_{s_j}^{-1}$, для которой (с $\mathbf{Y} = [\chi_{s_1}^{-1}, \chi_{s_2}^{-1}, \chi_{s_3}^{-1}, \ldots]$) выполняется

$$p_j = \chi_{s_j}^{-1}(c_j) \quad \text{или} \quad p_j = \mathrm{Y}(c_j, k_j) \quad \text{или} \quad p_j = c_j S_j^{-1}.$$

Таким образом, $\chi_{s_j}^{-1}(\chi_{s_j}(p_j)) = p_j$. Если, кроме того, χ_{s_j} *сюръективно* и *однозначно*, то *для всех* $c_j \in W^m$ имеет место равенство

$$\chi_{s_j}(\chi_{s_j}^{-1}(c_j)) = c_j.$$

В случае двустороннего обмена сообщениями между двумя сторонами A и B, одна из сторон может использовать последовательность χ_{s_j} для шагов шифрования и расшифрования, а другая — последовательность $\chi_{s_j}^{-1}$ для своих шагов шифрования и расшифрования.

2.6.3. Изоморфизм. Пусть снова $\mathbf{X}$ означает конечно порожденный блочный шифр. Два открытых текста $(p'_1, p'_2, p'_3, \ldots)$, $(p''_1, p''_2, p''_3, \ldots)$ такие, что $p'_i = p''_i S$, где S — фиксированная подстановка, называются изоморфными. Предположим то же самое для двух шифротекстов $(c'_1, c'_2, c'_3, \ldots)$, $(c''_1, c''_2, c''_3, \ldots)$ таких, что $c'_i = c''_i T$, где T есть фиксированная подстановка. Тогда справедливо утверждение:

$$\text{если } c'_i = c''_i S''_i \text{ и } c''_i = p''_i S''_i, \text{ то } S S'_i = S''_i T.$$

Если шифрования S'_i и S''_i обладают обратными, то шифротексты изоморфных открытых текстов также изоморфны, и наоборот. Тогда

$$T = (S''_i)^{-1} S S'_i \quad \text{и} \quad S = S''_i T (S'_i)^{-1}.$$

Если фиксированные подстановки S и T являются обратимыми, то ключи могут быть преобразованы один в другой:

$$S'_i = S S''_i T^{-1} \quad \text{и} \quad S'_i = S^{-1} S''_i T.$$

2.6.4. Шеннон. Система шифрования, в которой шаг шифрования однозначно определяется парой открытый текст/шифротекст, может быть названа шенноновской системой шифрования. Этим свойством обладают многие общепринятые системы шифрования. Система шифрования, где шаг шифрования и шаг расшифрования совпадают и таким образом криптопроцедура симметрично определена, называется системой с ключевой симметрией. В этом случае каждый шаг шифрования инволютивен (см. разд. 3.2.1).

Глава 3

Шаг шифрования: простая замена

Среди шагов шифрования мы сразу находим два больших класса: замены (подстановки) и перестановки. Мы начнем наше рассмотрение с нескольких видов замен и обратимся к перестановкам в гл. 6.

Простая замена (нем. *Tauschverfahren* или *Ersatzverfahren*) является подстановкой с односимвольным шагом шифрования $\chi_i \in M$,

$$\chi_i \colon V^{(1)} \to W^{(m_i)}.$$

В одноалфавитном случае, из M выбирается произвольное χ_s и шифрование выполняется с последовательностью $X = [\chi_s, \chi_s, \chi_s, \ldots]$. Понятно, что в этом случае достаточно брать одноэлементную систему шифрования M.

Мы начинаем рассмотрение со случая $m_i = 1$ для всех i.

3.1. Случай $V^{(1)} \to W$ (односимвольные простые замены)

Случай $V^{(1)} \to W$ имеет дело с односимвольными простыми заменами или, кратко, простыми заменами (фр. *substitution simple ordinaire*).

3.1.1. $V \to W$, неоднородное шифрование без омофонов и пустышек. Это первичный случай. В качестве W часто используется алфавит из странно выглядящих, необычных графем: известные примеры таких алфавитов можно найти в Таиланде, Персии, Эфиопии (коптский язык) и в других местах. Подобную маркировку использовал Джованни Баттиста Порта на своем шифровальном диске (рис. 23, см. также рис. 30). Карл Великий, как считают, использовал следующие символы (рис. 24):

Рис. 23. Шифровальный диск Порта, 1563 г.

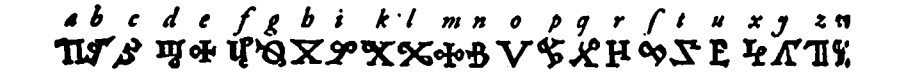


Рис. 24. Секретные символы Карла Великого

Следует упомянуть здесь и шифр масонов. Он возвращается к древнему шифру «хлев» («pigpen»), и в современной форме выглядит так:

a b c d e f g h i j k l m n o p q r s t u v w x y z

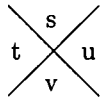
Его легко запомнить в соответствии со следующими схемами:

a	b	c
d	e	f
g	h	i

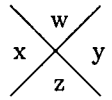
без точки

j	k	l
m	n	o
p	q	r

с точкой



без точки



с точкой

Не позже 1728 г. английским «черным кабинетом» взломан шифр императора Петра I, в котором использовалась (помимо номенклаторов) неоднородная замена $V \to W$ с причудливым шифроалфавитом (Петр I умер в 1725 г. — *Прим. ред.*).

Эдгар Аллан По, известный своими литературными трудами, в своей новелле «Золотой жук» (разд. 15.10.1) использовал довольно тривиальный алфавит из букв стандартных типографских шрифтов.

К этой же группе относится и шифр книготорговцев для шифрования цен и дат, представляющий взаимно однозначное отображение $Z_{10} \to Z_{26}$, порождаемое паролем («ключевой фразой»). Например,

1	2	3	4	5	6	7	8	9	0
M	I	L	C	H	P	R	O	B	E,

шаг шифрования с паролем *milchprobe* («молочная проба») использовался в течение многих лет в Германии для указания даты упаковки масла. Аналогично, в шифровальной машине ЭНИГМА Военно-морского флота Германии цифры иногда заменялись буквами:

1	2	3	4	5	6	7	8	9	0
q	w	e	r	t	z	u	i	o	p.

3.1.2. $V^{(1)} \to W$, неоднородное шифрование с омофонами и пустышками. Омофоны обнаруживаются уже в мусульманских источниках, например, у Калкашанди (1412 г.) и в шифре, применяемом для замены букв Симоном де Крема, секретарем герцога Мантуанского в 1401 г. Для гласных (обычно более частых символов в тексте) им были введены омофоны, первопричиной чего

было соображение о необходимости выравнять символьные частоты. Кроме того, в алфавит W были добавлены цифры. Введение омофонов фактически влечет необходимость введения пустышек; иначе омофоны легко могут распознаны по постоянному характеру букв, окружающих их в частых словах.

Методом с омофонами, используемым даже сегодня, является книжный шифр: в какой-нибудь неприметно выглядящей книге, которую отправитель и адресат имеют под рукой в идентичных изданиях, символы открытого текста выбираются один за другим; соответствующие координаты символа (страница x, строка y, позиция z) образуют для него шифрогруппу (x-y-z).

Выбрав для шифрования эту книгу, слово «молоко» можно было бы зашифровать в виде 7-2-6, 8-2-4, 8-6-1, 7-1-3, 9-8-5, 9-2-3.

3.2. Специальный случай $V \leftrightarrow V$ (перестановки)

В случае взаимно однозначного отображения $V \leftrightarrow W$ среди примеров в разд. 3.1.1 множество W называется перемешанным алфавитом (криптотекста) из N символов (фр. *alphabet désordonné*, *alphabet incohérent*, нем. *umgeordnetes Geheimtextalphabet*), который сопоставляется стандартному алфавиту (открытого текста) V из N символов (фр. *alphabet ordonné*, нем. *Standard-Klartextalphabet*).

Чтобы определить подстановку, достаточно перечислить некоторым способом пары соответствия символов открытого текста и криптотекста, например, для $V \triangleq W = Z_{26}$ (об использовании символов нижнего регистра и маленьких заглавных букв см. разд. 2.5.4):

```
u d c b m a v g k s t n w z e i h f q l j r o p x y
H E W A S R I G T O U D C L N M F Y V B P K J Q Z X
```

При шифровании, конечно, более удобно расположить символы открытого текста упорядоченными в стандартный алфавит (открытого текста); это задает перемешанный алфавит криптотекста:

```
a b c d e f g h i j k l m n o p q r s t u v w x y z
R A W E N Y G F M P T B S D J Q V K O U H I C Z X L
```

В математике эта «подстановочная нотация» является общепринятой. Для расшифровки, однако, лучше иметь символы криптотекста, упорядоченные в стандартный алфавит (криптотекста); это задает перемешанный алфавит открытого текста:

```
b i w n d h g u v o r z i e s j p a m k t q c y f x
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
```

Новая ситуация возникает в случае эндоморфизма $V \triangleq W$. В этом случае взаимно однозначное отображение $V \leftrightarrow V$ является перестановкой множества V; в электрическом шифраторе перестановка $V \leftrightarrow V$ может быть реализована перестановкой в пучке из N проводов (нем. *Umstecken*).

Часто для перестановок кроме подстановочной нотации математики используют «циклическую нотацию»:

(a r k t u h f y x z l b) (c w) (d e n) (g) (i m s o j p q v).

Здесь мы опускаем различие между прописными и маленькими заглавными буквами. При шифровании для каждого символа открытого текста отыскивается цикл, куда тот входит, и берется следующий за ним в цикле символ; при расшифровке переход производится к циклически предшествующему символу. Циклы длины 1 (1-циклы) часто опускаются; мы не будем следовать этому правилу.

3.2.1. Взаимообратные перестановки. Уже в самых древних источниках (кроме Египта; мы возвратимся к нему, рассматривая «коды») появляются взаимообратные («инволютивные») перестановки алфавита V: в Индии, в «*Камасутре*» писателя *Ватсяяна* упоминается секретное письмо как одно из шестидесяти четырех искусств; *Муладевьа* описывает процедуры шифрования и расшифрования, которые являются отражениями друг друга («инволюциями»):

$$V \overset{2}{\leftrightarrow} V: \quad \updownarrow \begin{matrix} \text{a} & \text{kh} & \text{gh} & \text{c} & \text{t} & \tilde{\text{n}} & \text{n} & \text{r} & \text{l} & \text{y} \\ \text{k} & \text{g} & \text{n} & \dot{\text{t}} & \text{p} & \dot{\text{n}} & \text{m} & \dot{\text{s}} & \text{s} & \acute{\text{s}} \end{matrix}$$

(остальные символы переводятся сами в себя, т. е. являются левоинвариантными для этой перестановки, так что перестановка не является, строго говоря, взаимообратной). При взаимообратной перестановке говорят, что алфавиты открытого текста и криптотекста обратны друг другу.

В древнееврейских священных книгах использовалась (хотя и не в криптографических целях) *бустрофодоническая* (boustrophedonic) подстановка, называемая *Атбаш* (Athbash), которая в латинском алфавите $V = Z_{20}$ выглядит так:

$$V \overset{2}{\leftrightarrow} V: \quad \updownarrow \begin{matrix} \text{a} & \text{b} & \text{c} & \text{d} & \text{e} & \text{f} & \text{g} & \text{h} & \text{i} & \text{l} \\ \text{z} & \text{v} & \text{t} & \text{s} & \text{r} & \text{q} & \text{p} & \text{o} & \text{n} & \text{m} \end{matrix}$$

Такая перестановка использует перевернутый («инверсный») алфавит. В случае отражения

$$V \overset{2}{\leftrightarrow} V: \quad \updownarrow \begin{matrix} \text{a} & \text{b} & \text{c} & \text{d} & \text{e} & \text{f} & \text{g} & \text{h} & \text{i} & \text{l} & \text{m} \\ \text{a} & \text{z} & \text{v} & \text{t} & \text{s} & \text{r} & \text{q} & \text{p} & \text{o} & \text{n} & \text{m} \end{matrix}$$

Чарльз Эйрауд говорит о дополнительном алфавите (фр. *alphabet complémentaire*); см. разд. 5.6. Эта перестановка, строго говоря, не является взаимообратной: буквы /a/ и /m/ являются левоинвариантными.

Очевидно, также является отражением со сдвинутым алфавитом (подобно древнееврейскому *Альбаму* (Albam)) шифр, использованный в 1589 г. Арджента с $V = Z_{20}$:

$$V \overset{2}{\leftrightarrow} V: \quad \updownarrow \begin{matrix} \text{a} & \text{b} & \text{c} & \text{d} & \text{e} & \text{f} & \text{g} & \text{h} & \text{i} & \text{l} \\ \text{m} & \text{n} & \text{o} & \text{p} & \text{q} & \text{r} & \text{s} & \text{t} & \text{v} & \text{z} \end{matrix}$$

или шифр, использованный Джованни Баттиста Порта в 1563 г. (см. рис. 53) с $V = Z_{22}$:

$$V \overset{2}{\leftrightarrow} V: \quad \updownarrow \begin{matrix} a & b & c & d & e & f & g & h & i & l & m \\ n & o & p & q & r & s & t & v & x & y & z \end{matrix}$$

Наиболее общий бустрофодонический случай, демонстрирующий использование пароля, представлен следующим примером ($V = Z_{26}$):

$$V \overset{2}{\leftrightarrow} V: \quad \updownarrow \begin{matrix} a & n & g & e & r & s & b & c & d & f & h & i & j \\ z & y & x & w & v & u & t & q & p & o & m & l & k \end{matrix}$$

Отражения, кроме преимущества компактности записи, имеют свойство, которому некоторые люди придают большое значение, а именно, идентичность процедур шифрования и расшифрования.

В циклической нотации перестановок, последние пять примеров выглядят (с цепочками, упорядоченными в алфавитном порядке) следующим образом:

(a,z)	(b,v)	(c,t)	(d,s)	(e,r)	(f,q)	(g,p)	(h,o)	(i,n)	(l,m)			
(a)	(b,z)	(c,v)	(d,t)	(e,s)	(f,r)	(g,q)	(h,p)	(i,o)	(l,n)	(m)		
(a,m)	(b,n)	(c,o)	(d,p)	(e,q)	(f,r)	(g,s)	(h,t)	(i,v)	(l,z)			
(a,n)	(b,o)	(c,p)	(d,q)	(e,r)	(f,s)	(g,t)	(h,v)	(i,x)	(l,y)	(m,z)		
(a,z)	(b,t)	(c,q)	(d,p)	(e,w)	(f,o)	(g,x)	(h,m)	(i,l)	(j,k)	(n,y)	(r,v)	(s.u)

В узком смысле взаимно обратная перестановка есть перестановка без 1-циклов, т. е. состоящая исключительно из 2-циклов («обменов»). Она же является желанной целью для криптоаналитических атак (разд. 14.1), которые перестают работать, если некоторые из циклов являются 1-циклами («женские» циклы).

Для двоичного алфавита $V = Z_2$ единственной нетривиальной перестановкой является отражение:

$$V \overset{2}{\leftrightarrow} V: \quad \updownarrow \begin{matrix} O \\ L \end{matrix}$$

3.2.2. Перекрестное соединение. При реализации в электрической схеме отражение выполняется перестановкой пары проводов, простым использованием двусторонних разъемов (рис. 25). Такие отражения использовались в коммутационной панели ЭНИГМЫ (нем. *Steckerbrett*).

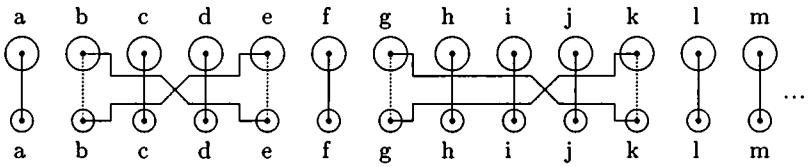


Рис. 25. Взаимообратная перестановка с помощью перекрестного соединения двусторонних разъемов

Число отражений $d(k, N)$ зависит от N и числа k используемых разъемов:

$$d(k, N) = \frac{N!}{2^k \cdot (N-2k)! \cdot k!} = \binom{N}{2k} \cdot \frac{(2k)!}{2^k k!} = \binom{N}{2k} \cdot (2k-1)!!,$$

где

$$(2k-1)!! = (2k-1)\cdot(2k-3)\cdot\ldots\cdot 5\cdot 3\cdot 1 = \frac{(2k)!}{2^k k!}.$$

Соответственно взаимно обратные перестановки («истинные» отражения) требуют, чтобы число $N = 2\nu$ было четным. Число $d\left(\frac{N}{2}, N\right)$ всех «истинных» отражений равно тогда (с относительной ошибкой $< 10^{-3}$ при $N > 6$)

$$d\left(\frac{N}{2}, N\right) = (N-1)!! = (N-1)\cdot(N-3)\cdot\ldots\cdot 5\cdot 3\cdot 1 = \frac{(2\nu)!}{\nu!2^\nu} \approx \frac{\sqrt{(2\nu)!}}{\sqrt[4]{\pi\cdot\left(\nu+\frac{1}{4}\right)}}.$$

Приближенное значение является довольно хорошей верхней границей для числа $(N-1)!!$. Отметим тем не менее, что при фиксированном N величина $d(k, N)$ достигает максимума для $k = [\nu - \sqrt{(\nu+1)/2}]$:

$$\begin{array}{lll} d(5,\,26)\approx 5.02\cdot 10^9, & d(6,\,26)\approx 1.00\cdot 10^{11}, & d(7,\,26)\approx 1.31\cdot 10^{12}, \\ d(8,\,26)\approx 1.08\cdot 10^{13}, & d(9,\,26)\approx 5.38\cdot 10^{13}, & d(10,\,26)\approx 1.51\cdot 10^{14}, \\ d(11,\,26)\approx 2.06\cdot 10^{14}, & d(12,\,26)\approx 1.03\cdot 10^{14}, & d(13,\,26)\approx 7.91\cdot 10^{12}, \end{array}$$

и $d(3,\,10) = 3150$, $d(4,\,10) = 4725$, $d(5,\,10) = 945$.

Шифровальная машина ЭНИГМА I, стоявшая на вооружении *Рейхсвера* в 1930 г., первоначально использовала шесть двусторонних двухканальных разъемов; начиная с 1 октября 1936 г. ЭНИГМА имела от пяти до восьми, с 1 января 1939 г. — от семи до десяти и с 19 августа 1939 г. — десять двусторонних двухканальных разъемов для перекрестного соединения.

3.2.3. Циклические перестановки. В компактной записи также легко описывается циклическая перестановка, порядок которой равен N: например, цикл стандартного алфавита Z_{20} с $N = 20$

$$V \overset{N}{\leftrightarrow} V\colon \quad (\mathrm{a\ b\ c\ d\ e\ f\ g\ h\ i\ l\ m\ n\ o\ p\ q\ r\ s\ t\ v\ x})$$

или его третья степень

$$V \overset{N}{\leftrightarrow} V\colon \quad (\mathrm{a\ d\ g\ l\ o\ r\ v\ b\ e\ h\ m\ p\ s\ x\ c\ f\ i\ n\ q\ t})$$

описываются в подстановочной записи как

$$\begin{array}{cccccccccccccccccccc} \mathrm{a} & \mathrm{b} & \mathrm{c} & \mathrm{d} & \mathrm{e} & \mathrm{f} & \mathrm{g} & \mathrm{h} & \mathrm{i} & \mathrm{l} & \mathrm{m} & \mathrm{n} & \mathrm{o} & \mathrm{p} & \mathrm{q} & \mathrm{r} & \mathrm{s} & \mathrm{t} & \mathrm{v} & \mathrm{x} \\ \mathrm{B} & \mathrm{C} & \mathrm{D} & \mathrm{E} & \mathrm{F} & \mathrm{G} & \mathrm{H} & \mathrm{I} & \mathrm{L} & \mathrm{M} & \mathrm{N} & \mathrm{O} & \mathrm{P} & \mathrm{Q} & \mathrm{R} & \mathrm{S} & \mathrm{T} & \mathrm{V} & \mathrm{X} & \mathrm{A}, \end{array}$$

$$\begin{array}{cccccccccccccccccccc} \mathrm{a} & \mathrm{b} & \mathrm{c} & \mathrm{d} & \mathrm{e} & \mathrm{f} & \mathrm{g} & \mathrm{h} & \mathrm{i} & \mathrm{l} & \mathrm{m} & \mathrm{n} & \mathrm{o} & \mathrm{p} & \mathrm{q} & \mathrm{r} & \mathrm{s} & \mathrm{t} & \mathrm{v} & \mathrm{x} \\ \mathrm{D} & \mathrm{E} & \mathrm{F} & \mathrm{G} & \mathrm{H} & \mathrm{I} & \mathrm{L} & \mathrm{M} & \mathrm{N} & \mathrm{O} & \mathrm{P} & \mathrm{Q} & \mathrm{R} & \mathrm{S} & \mathrm{T} & \mathrm{V} & \mathrm{X} & \mathrm{A} & \mathrm{B} & \mathrm{C}. \end{array}$$

Последний шаг шифрования, использовавшийся (согласно Светонию) Юлием Цезарем, состоит в замене каждого символа на третий после него символ в алфавите. Его преемник Август, уступавший во многих отношениях

Цезарю, использовал лишь первый шаг шифрования (возможно, он не мог уверенно считать до трех); Светоний рассказал, что он также заменял букву x на AA.

Всякая степень цикла стандартного алфавита порождает алфавит ЦЕЗАРЯ. Мы возвратимся к нему в главе 5 (сложение ЦЕЗАРЯ). Здесь лишь отметим следующее: в то время как оба шага шифрования, описанных выше, имеют порядок двадцать, вторая степень этих шагов имеет порядок десять, а десятая степень имеет порядок всего лишь два: она является отражением, которое рассматривалось выше. $(N-1)$-я степень моноциклической перестановки обратна ей и дает шаг расшифрования.

Одноалфавитная подстановка с шагом шифрования ЦЕЗАРЯ была введена в 1915 г. в русской армии после того, как на службу были призваны новые кадры, которые оказались не в состоянии использовать при шифровании что-либо более сложное. Для Людвига Дюбнера и Германа Покорны, глав криптоаналитических служб Германии и Австрии, дешифровка таких сообщений была приятным простым делом.

По самой своей природе дорожка диска, обод шайбы или склеенная в виде окружности полоска могут использоваться для представления полного цикла. Такие приспособления нашли широкое применение и использовались определенным образом (разд. 7.5.3) Томасом Джефферсоном (около 1795 г.) и Этьеном Базерье (1891 г.). q-я степень циклической перестановки получалась путем последовательного отсчета внутри цикла q символов.

3.2.4. Перемешанные алфавиты. Для невзаимообратной и нециклической перестановки $V \leftrightarrow V$ в наиболее общем случае перемешанного алфавита (фр. *alphabet désordonné*, нем. *permutiertes Alphabet*) обычно используется подстановочная запись:

$V \leftrightarrow V$:
a b c d e f g h i j k l m n o p q r s t u v w x y z
S E C U R I T Y A B D F G H J K L M N O P Q V W X Z

Краткая циклическая запись также полезна и здесь. Это показывает разложение

$V \leftrightarrow V$: (a s n h y x w v q l f i) (b e r m g t o j) (c) (d u p k) (z),

состоящее из одного цикла длины 12, одного цикла длины 8, одного цикла длины 4 и двух циклов длины 1 (циклическое разбиение типа $12+8+4+1+1$).

3.2.4.1. Более перемешанные алфавиты получаются циклическим сдвигом одной из двух строк в подстановочной записи (перемешанный алфавит со сдвигом, фр. *alphabet désordonné parallule*, нем. *verschobenes permutiertes Alphabet*):

$V \leftrightarrow V$:
a b c d e f g h i j k l m n o p q r s t u v w x y z
E C U R I T Y A B D F G H J K L M N O P Q V W X Z S

$V \leftrightarrow V$:
a b c d e f g h i j k l m n o p q r s t u v w x y z
C U R I T Y A B D F G H J K L M N O P Q V W X Z S E

или в циклической записи

(a e i b c u q m h) (d r n j) (f t p l g y z s o k) (v) (w) (x),
(a c r o l h b u v w x z e t q n k g) (f y s p m j) (d i).

3.2.4.2 Итерированная подстановка, называемая также «степенной», порождает степени перемешанного алфавита. Например, вторая степень рассмотренной выше подстановки SECURITY... в циклической записи имеет вид

(a n y w q f) (b r g o) (c) (dp) (e m t j) (h x v l i s) (k u) (z),

причем все циклы четной длины исходной подстановки расщепляются пополам. В подстановочной нотации она имеет вид:

$V \leftrightarrow V$: a b c d e f g h i j k l m n o p q r s t u v w x y z
N R C P M A O X S E U I T Y B D F G H J K L Q V W Z

Сдвиг подстановки, с одной стороны, и возведение ее в степень — с другой, в общем случае не дают один и тот же результат; они являются двумя сильно различающимися методами генерации семейства (состоящего из N; иногда меньше) сопутствующих алфавитов (гл. 7).

3.2.5. Алфавиты, порождаемые паролями. Примеры, рассмотренные нами выше, показывают способ построения (эндоморфной) простой подстановки $V \leftrightarrow V$ с помощью пароля (фр. *mot-clef*, нем. *Kennwort*, *Losung*), возможно, мнемонического ключа или ключевой фразы (лозунга). Классический метод использует слово (с неповторяющимися буквами) в алфавите V и дополняет его оставшимися буквами в алфавитном порядке. Метод восходит к Джованни Баттиста Ардженти, применившим его примерно в 1580 г. Даже в XX веке он все еще входит в шифровальный арсенал[1)].

Эта конструкция, однако, является уязвимой: можно просто угадать отсутствующую парольную часть подстановки (в конце концов, наиболее частые гласные /e/ и /a/ (в английском языке — *прим. перев.*) как правило заменяются символами из пароля, если тот имеет длину 5 и более символов). Слабым утешением служит то, что нет нужды в большой длине пароля.

Более продвинутые методы используют переупорядочивание алфавита с парольным началом, например, записывая его сначала в виде таблицы по строкам и затем читая его по столбцам (метод Чарльза Вейтстоуна 1854 г.:

[1)] Разрешение повторений плохо: это ведет к многозначности, например, ключевая фраза шифра

a b c d e f g h i j l m n o p q r s t u v x y z
L E G O U V E R N E M E N T P R O V I S O I R E

укорачивает алфавит криптотекста (в нашем случае до 13 символов, так как {b, g, j, m, z} $\mapsto$ $\mapsto$ E).

перестановка, которая будет систематически исследована в разд. 6.2):

SECURITY	aeilorux
ABDFGHJK	bfjmpsvy
LMNOPQVW	cgknqtwz
XZ	dh.

Это порождает перестановку алфавита

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
S	A	L	X	E	B	M	Z	C	D	N	U	F	O	R	G	P	I	H	Q	T	J	V	Y	K	W

или в циклической записи

(a s h z w v j d x y k n o r i c l u t q p g m f b) (e)

с 1-циклом (e).

Следующее развитие метода состоит в заполнении столбцов алфавита открытого текста (в рамках заполнения той же таблицы) в соответствии с алфавитным порядком символов пароля; применительно к нашему примеру получаем следующий порядок заполнения столбцов:

третий, второй, шестой, пятый, первый, седьмой, четвертый, восьмой,

что дает в результате

SECURITY	ndaukhrx
ABDFGHJK	oebvlisy
LMNOPQVW	pfcwmjtz
XZ	qg.

Это порождает перестановку алфавита

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
C	D	N	E	B	M	Z	I	H	Q	R	G	P	S	A	L	X	T	J	V	U	F	O	Y	K	W

или в циклической записи

(a c n s j q x y k r t v f m p l g z w o) (b d e) (h i) (u).

Метод может также использоваться для конструирования циклов. Например, предложение «*évitez les courants d'air*», «избегайте сквозняков» (Базерье, разд. 7.4.3) порождает цикл

$V \leftrightarrow V$: (e v i t z l s c o u r a n d b f g h j k m p q x y).

3.2.6. Подсчеты. В следующей таблице даются (для значений $N = 26$, $N = 10$ и $N = 2$) числа $Z(N)$ имеющихся алфавитов $V \leftrightarrow V$ в рассмотренных нами классах перестановок:

число перестановок	$Z(N)$	$Z(26)$	$Z(10)$	$Z(2)$
всего	$N!$	$4.03 \cdot 10^{26}$	3 628 800	2
моноциклических	$(N-1)!$	$1.55 \cdot 10^{25}$	362 880	1
отражений всего	$\approx N \cdot (N!)^{1/2}$	$5.33 \cdot 10^{14}$	9 496	2
истинные отражения,	$\approx (N!)^{1/2}$	$7.91 \cdot 10^{12}$	945	1
порожденные паролями (мнемоническими словами)		$10^4 \ldots 10^6$		

3.2.7. Шифровальные диски и линейки. Чтобы механизировать подстановку, можно зафиксировать соответствие символов открытого текста и криптотекста, разместив их в соответствие с подстановочной записью на цилиндре или линейке. Причем можно закрыть поверхности с символами, оставив лишь два окна (перемещающихся вдоль символов), чтобы в любой момент видеть только два символа открытого текста и криптотекста, соответствующие друг другу. Разместить окна можно таким образом, чтобы только хозяин видел символ открытого текста, в то время как секретарь (писарь) видел бы лишь окно криптотекста и не мог понять значение сообщения (см. рис. 54 в разд. 7.5.2: машина Грипенштерна).

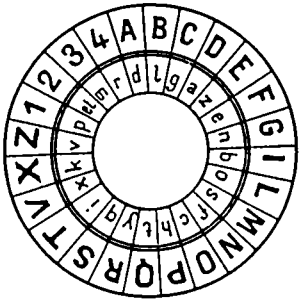


Рис. 26. Шифровальный диск Леона Баттисты Альберти (согласно Ланге-Сударту, 1935 г.)

Если одно из окон может перемещаться (вдоль своего алфавита), то мы можем выбрать N дополнительных сдвинутых алфавитов. Другой возможностью является сдвиг алфавита открытого текста относительно алфавита криптотекста. Это ведет к использованию двух дисков (рис. 26) или двух полос (рис. 27). В последнем случае необходимо повторение одного из алфавитов (дублирование).

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	…
S	E	C	U	R	I	T	Y	A	B	D	F	G	H	J	K	L	M	N	O	P	Q	V	W	X	Z															

Рис. 27. Шифровальная линейка с размноженным алфавитом открытого текста

Шифровальные диски (фр. *cadran*, нем. *Chiffrierscheibe*), механические устройства для общей подстановки со сдвинутыми перемешанными алфавитами были описаны еще в 1466 г. Леоном Баттиста Альберти[2)] (см. вклейку В). Шифровальные линейки (фр. *reglette*, нем. *Chiffrierschieber*) использовались

[2)] В иллюстрации Альберти заглавные буквы используются для открытого текста, малые символы для криптотекста, что отличается от современной традиции. Символ /et/, по-видимому, заменяет символ m. Начальная установка диска заключается в выставлении напротив друг друга ключевого символа, скажем D, с фиксированным символом, например /a/.

в елизаветинской Англии примерно в 1600 г. В XIX столетии они получили название линейки Сен-Сира по имени известной французской Военной академии. Шифровальные стержни (фр. *bâtons*, нем. *Chiffrierstäbchen*) имели то же самое назначение.

3.2.8. Циклы с окнами. Механическая реализация циклической перестановки может быть выполнена на основе циклической записи. Цикл символов снова располагается на цилиндре или на линейке (в этом случае первый символ должен быть продублирован). Два смежных окна в каждом положении позволяют видеть только два символа. Левый из них берется из открытого текста, а в другом окне показывается соответствующий символ криптотекста.

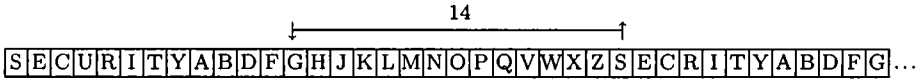


Рис. 28. Шифровальная линейка с окнами для реализации степеней алфавита

Выбор из возможных степеней (вплоть до N) перемешанного алфавита реализуется путем варьирования расстояния между окнами (в случае линейки необходимо дублировать весь цикл). При этом q-я степень циклической перестановки получается при установке расстояния между окнами в q символов (на рис. 28 имеем $q = 14$).

3.3. Случай $V^{(1)} \to W^m$. Простые подстановки мультиграммами

3.3.1. $m = 2$, простая подстановка биграммами $V^{(1)} \to W^2$. Подстановки биграммами (двухсимвольные подстановки) были известны еще в античности. Полибием была описана пятеричная ($|W| = 5$) двухсимвольная подстановка для греческих букв. В современной версии подстановки буквами латинского алфавита Z_{25} заполняется шахматная доска 5×5:

	1	2	3	4	5
1	a	b	c	d	e
2	f	g	h	i	k
3	l	m	n	o	p
4	q	r	s	t	u
5	v	w	x	y	z

или

	1	2	3	4	5
1	a	f	l	q	v
2	b	g	m	r	w
3	c	h	n	s	x
4	d	i	o	t	y
5	e	k	p	u	z

Расшифровав с помощью «квадрата Полибия» (правый квадрат) текст семаграммы

33515141234333514512432411343411343442 3 31144424333

из рис. 3 разд. 1.2, получаем открытый текст

n e e d m o n e y f o r a s s a s s i n a t i o n

(нужны деньги для убийства).

В свое время Полибий описал, как можно представить числа от 1 до 5 факелами; в более близкие к нам времена для этого использовалась «тюремная азбука» перестукивания. Помещенный выше специальный шифр $Z_{25} \to Z_5 \times Z_5$ распространен повсеместно, являясь поистине интернациональным шифром перестукивания, он использовался в тюрьмах от Алкатраца до Плотцензее как преступниками, так и политическими заключенными. Обычная скорость передачи для него составляет 8–15 слов в минуту.

В царской России такой шифр перестукивания (с русским алфавитом в квадрате 6×6) применялся достаточно широко и проник в Западную Европу с русскими анархистами как часть «шифра Нигилиста» (разд. 9.4.4), он также использовался стеганографически, см. разд. 1.2. Артур Кестлер в романе «*Слепящая тьма*» («*Sonnenfinsternis*») и Александр Солженицин в «*Архипелаге Гулаг*» описали его использование в Советском Союзе.

В общем случае используется пароль, который дописывается построчно, начиная с первой строки, а оставшееся место заполняется остальными буквами алфавита. Граф Оноре де Мирабо, французский революционер 18-го века, использовал этот метод в своей переписке с маркизой Софи де Моньер, дополняя его стеганографией и добавляя в качестве пустышечных символов цифры 6 7 8 9 0.

Система шифрования ADFGVX, изобретенная Фрицем Небелем (1891–1967 гг.) и внедренная в 1918 г. на немецком Западном фронте генерал-квартирмейстером Эрихом Людендорфом для радиосвязи (с криптоалфавитом Z_6, см. разд. 2.5.2), работала с $|W| = 6$ и шахматными досками, подобными следующей:

	A	D	F	G	V	X
A	c	o	8	x	f	4
D	m	k	3	a	z	9
F	n	w	l	0	j	d
G	5	s	i	y	h	u
V	p	1	v	b	6	r
X	e	q	7	t	2	g

Применяются также прямоугольные таблицы. Примерно в 1580 г. Джованни Баттиста Ардженти использовал следующую схему (с $W = Z_{10}$):

	0	1	2	3	4	5	6	7	8	9
1	p	i	e	t	r	o	a	b	c	d
2	f	g	h	l	m	n	q	s	u	z

в которой впервые был применен пароль.

Во общем случае двухсимвольная подстановка дает большой простор для использования омофонов:

	0	1	2	3	4	5	6	7	8	9
9, 6, 3	a	b	c	d	e	f	g	h	i	
8, 5, 2	j	k	l	m	n	o	p	q	r	
7, 4, 1	s	t	u	v	w	x	y	z		

В этом примере символ 0 может служить в качестве пустышки. Символ 0, первоначально *цифра нуль* (*nulla ziffra*), нигде до сих пор всерьез не используется.

Как правило, омофоны применяются для выравнивания символьных частот в криптотексте. Так как буквы e t a o n i r s h в английском языке имеют в сумме частоту примерно 70%, то мы достигнем хорошего равновесия в распределении символьных частот криптотекста, пользуясь подстановкой

	1	2	3	4	5	6	7	8	9	
4, 5, 6, 7, 8, 9, 0	e	t	a	o	n	i	r	s	h	71.09%
2, 3	b	c	d	f	g	j	k	l	m	19.46%
1	p	q	u	v	w	x	y	z		9.45%.

В следующем методе используется пароль из 4 символов, который задает, таким образом, начала циклов (**00** ... 24), (**25** ... 49), (**50** ... 74), (**75** ... 99) в определении (с $V = Z_{25}$ и $W = Z_{10}^2$) омофонного шифра, например, с паролем *KILO*:

	a	b	c	d	e	f	g	h	i	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
K	16	17	18	19	20	21	22	23	24	**00**	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
I	42	43	44	45	46	47	48	49	**25**	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
L	65	66	67	68	69	70	71	72	73	74	**50**	51	52	53	54	55	56	57	58	59	60	61	62	63	64
O	87	88	89	90	91	92	93	94	95	96	97	98	99	**75**	76	77	78	79	80	81	82	83	84	85	86

Десятеричный ($|W| = 10$) двухсимвольный шифр не должен иметь омофонов — подстановка не должна быть сюръективной и некоторые пары символов могут не использоваться. Такой шифр применял шведский баронет Фридрик Грипенштерн, основываясь, возможно, на предложении Кристофера Полхейма. Забавная форма двухсимвольного шифра с омофонами была предложе-

1	2	3	4	5	6	7	8	9	8	
I	P	I		O	U	O		P	N	1
W	E	U	T	E	K		L	O		2
E	U	G	N	B	T	N		S	T	3
T	A	Z	M	D		I	O	E		4
S	V	T	J		E		Y		H	5
N	A	O	L	N	S	U	G	O	E	6
	C	B	A	F	R	S		I	R	7
I	C	W	Y	R	U	A	M		N	8
M	V	T		H	P	D	I	X	Q	9
L	S	R	E	T	D	E	A	H	E	0

Рис. 29. Двухсимвольный шифр, использованный в Лос-Аламосе в 1944 г. для телефонных разговоров

на во время создания атомной бомбы в США бригадным генералом Лесли Р. Гровсом и подполковником Пиром да Силва в Лос-Аламосе (рис. 29) для сокрытия особо важных имен и должностей при телефонных разговорах. Дело в том, что требовалось время для поиска букв, и, таким образом, омофоны выбирались чаще наугад, нежели предопределялось шифратором.

3.3.2. $m=3$, простая подстановка триграммами $V^{(1)} \to W^3$. Подстановка триграммами (трехсимвольная подстановка) была предложена Тритемием в «Полиграфии» (Polygraphiae) с $|W|=3$ (заметим, что $3^3=27>26$) для стеганографического применения. Тем не менее, подобные троичные подстановки достаточно редки.

3.3.3. $m=5$, простая подстановка пентаграммами $V^{(1)} \to W^5$. Подстановка с помощью пятисимвольных групп криптотекста (пятисимвольная подстановка) с $|W|=2$ использовалась Фрэнсисом Бэконом в связи со стеганографическими средствами (заметим, что $2^5=32>26$). Пятисимвольное двоичное шифрование возродилось в шифровальной машине Вернама в 1918 г. (разд. 8.3.2) и во время Второй мировой войны в шифровальных телетайпах Siemens T52 (*Geheimschreiber*) и Lorenz SZ40/42 (*Schliisselzusatz*), см. разд. 9.1.3.

3.3.4. $m=8$, простая подстановка октетами $V^{(1)} \to W^8$. Вновь $|W|=2$ (8-разрядный код, двоичный EBCDIC код (расширенный двоично-десятичный код обмена информацией), код ASCII с проверочным битом), эта восьмисимвольная простая подстановка совпадает с односимвольной подстановкой байтами (Z_{256}) в современных компьютерах.

3.4. Общий случай $V^{(1)} \to W^m$. Словарь криптотекста «с разбросом»

Общий случай $V^{(1)} \to W^m$ естественным образом приводит к использованию пустышек и омофонов.

Симон де Крема из Мантуи (1401 г.) использовал только омофоны (с $m=1$). Начиная с $m=2$, помимо использования омофонов и пустышек вступает в игру новая важная идея: разброс длин слов криптоалфавита (нем. *Spreizen*), или, другими словами, отображение V в $W^1 \cup W^2$. Шифр, применявшийся при Священном престоле (папском дворе), изобретенный Маттео Ардженти после 1590 г., демонстрирует использование омофонов, пустышек и разброса длин слов. Для алфавита Z_{24}, обогащенного наборами /et/, /con/, /non/, /che/ и с цифрами 5 и 7 в качестве пустышек, шаг шифрования $Z_{24}^{(1)} \to Z_{24}^1 \cup Z_{24}^2$ имеет вид

a	b	c	d	e	f	g	h	i	l	m	n	o
1	86	02	20	62 82	22	06	60	3	24	26	84	9
p	q	r	s	t	v	z	et	con	non	che	ε	
66	68	28	42	80 40	04	88	08	64	00	44	5 7	

3.4.1. Предостережение. Шаги шифрования с разбросом длин слов алфавита криптотекста имеют одно ограничение, а именно: порожденное ими шифрование должно быть лево-однозначным, т. е. границы между односимвольными и двухсимвольными элементами шифра должны правильно определяться, обеспечивая, таким образом, корректное разложение криптотекста на эти элементы. Как было указано в разд. 2.4, Дж. Б. и М. Ардженти знали об этом. Их шифры удовлетворяли следующему условию: W делится на две группы символов: символы, используемые для односимвольных элементов шифра $W' = \{1, 3, 5, 7, 9\}$ и символы, используемые для двухсимвольных элементов шифра, начинающихся с символов из $W'' = \{0, 2, 4, 6, 8\}$. Ардженти сделали ошибку, ограничив выбор множеством вторых символов в криптоалфавите W''. Это выявляет разброс длин слов криптоалфавита. С другой стороны, они же дали несколько более практичных рекомендаций: подавлять букву u, следующую за буквой q, и подавлять повторяющиеся буквы.

Так называемые шпионские шифры, использованные советским НКВД и его последователями, являются шифрами с разбросом длин слов алфавита криптотекста. О них стало известно от осужденных разведчиков. По аналогии с квадратами Полибия эти шифры также описываются прямоугольными массивами, например,

	0	1	2	3	4	5	6	7	8	9
	s	i	o	e	r	a	t	n		
8	c	x	u	d	j	p	z	b	k	q
9	·	w	f	l	/	g	m	y	h	v

(*)

где первая строка содержит односимвольные элементы шифра.

При $W = Z_{10}$ доступны 28 элементов шифра, что достаточно для Z_{26} и двух специальных символов: «stop» и «/» для перехода буква/цифра. Так как этот шифр подвергался дальнейшему шифрованию («закрытию», разд. 9.2.1), им можно было зашифровать числовые данные (обрамляя числа признаком перехода буква/цифра), удваивая цифры в целях защиты от возможных ошибок при передаче.

При конструировании таких массивов использовались также и пароли. Доктор Пер Мейрлинг, их шведский сподвижник, делал это в 1937 г. следующим образом: он записывал пароль из 8 символов (М. Дельвайо был испанским коммунистом) и под ним остальные буквы алфавита; столбцы нумеровались в обратном порядке:

	0	9	8	7	6	5	4	3	2	1
	m	d	e	l	v	a	y	o		
1	b	c	f	g	h	i	j	k	n	p
2	q	r	s	t	u	w	x	z	·	/.

Эта процедура имела тот недостаток, что несколько наиболее частых символов не получали шифрогруппы из одной цифры. Этот недостаток был также характерен для метода шведского разведчика Бертила Эрикссона, использо-

ванного им в 1941 г.: он нумеровал столбцы согласно алфавитному порядку букв пароля (разд. 3.2.5):

	6	0	8	7	5	4	9	1	2	3
3	p	a	u	s	o	m	v	e	j	k
9	b	c	d	f	g	h	i	l	n	q
	r	t	w	x	y	z.				

Пароль брался из шведского перевода рассказа Ярослава Гашека *Paus som Svejk sjalv avbrot...* Так как шифрование наиболее частых символов шифрогруппами из одноразрядных чисел также уменьшает время передачи по телеграфу, НКВД пришел в 1940 г. к методу конструирования шифров, который принимал во внимание это соображение. Макс Клаузен, радист русского разведчика в Токио Рихарда Зорге должен был запомнить фразу «*a sin to err*» («грешно допускать ошибку» — очень хороший совет для шпиона), содержащую восемь наиболее частых букв английского языка (в сумме 65.2%). Начинаясь с пароля /subway/, прямоугольник Полибия далее заполнялся остающимися символами алфавита. Вслед за этим, двигаясь по столбцам слева направо, сначала буквам из набора {a s i n t o e r} назначались числа 0... 7 в порядке их появления; затем оставшимся символам назначались числа 80... 99:

s	u	b	w	a	y
0	82	87	91	5	97
c	d	e	f	g	h
80	83	3	92	95	98
i	j	k	l	m	n
1	84	88	93	96	7
o	p	q	r	t	v
2	85	89	4	6	99
x	z	·	/		
81	80	90	94		

Таким способом для прямоугольника Полибия, помеченного выше (*), получается более компактная запись.

Для кириллического алфавита подходящим является подразбиение на семь одноразрядных и тридцать двухразрядных шифрогрупп (всего 37 шифрогрупп); это позволяет добавить в алфавит 5 специальных символов. Шифр, выданный перебежчиком Реино Хейханеном (он был помощником русского суперразведчика Рудольфа Абеля), использовал русское слово, подобное слову СНЕГОПАД, первые семь букв которого имеют суммарную частоту 44.3%. Прямоугольник Полибия формировался обычным способом:

С	Н	Е	Г	О	П	А	·	·	·
Б	В	Г	Д	Ж	И	Й	К	Л	М
Р	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ
Ы	Ь	Э	Ю	Я	·	·	·	·	·

и затем перестраивался с помощью ключа, который менялся от сообщения к сообщению и размещался в заранее оговоренном месте шифрованного сообщения. Наконец, производилось перешифрование (разд. 9.2.1).

3.4.2. Русское соединение. В этом методе (он также стал известен из практики русских агентов, что послужило основанием назвать его «русским соединением») сообщение разбивалось на две части приблизительно одинаковой длины, и эти части менялись местами, пряча таким образом заметные стандартные фразы в начале и конце сообщения куда-нибудь в его середину. Уинстон Черчилль назвал Россию «решето, обернутое тайной внутри загадки». Это справедливо также и для русской криптологии.

Глава 4

Шаги шифрования: многосимвольная подстановка и кодирование

Простая (односимвольная) подстановка требует полного разложения открытого текста на отдельные символы. Многосимвольная подстановка допускает многосимвольные шаги шифрования, то есть шаги шифрования вида $V^{(n)} \to W^{(m)}$ с $n > 1$.

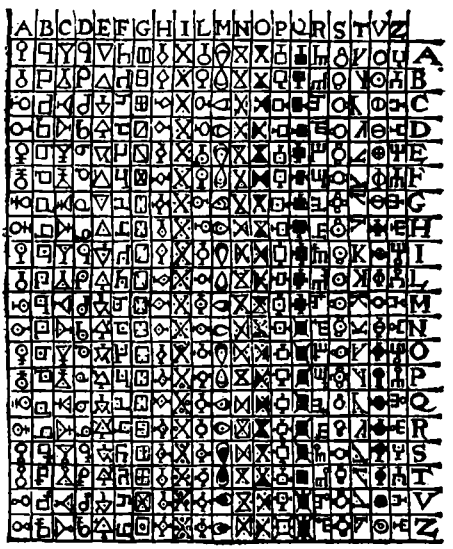


Рис. 30. Старинная двухсимвольная подстановка Джованни Баттисты Порта, 1563 г. (Giambattista Della Porta)

4.1. Случай $V^2 \to W^{(m)}$ (двухсимвольные подстановки)

4.1.1. Графемы. Самое раннее многосимвольное шифрование этого типа встречается в книге Порта 1563 г. «*Defurtivis literarum notis*» (рис. 30), а именно, отображение $V^2 \to W^1$. Порта проявил большую изобретательность в придумывании 400 причудливых символов для криптоалфавита.

4.1.2. Двухсимвольный шаг шифрования биграммами $\mathbf{V^2 \to V^2}$. Чаще всего, для его представления используется матрица. В случае взаимной однозначности $V^2 \leftrightarrow V^2$ он является перестановкой биграмм.

Ниже приводится пример $V^2 \overset{2}{\longleftrightarrow} V^2$ взаимообратной перестановки биграмм:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	...
a	XZ	KJ	YJ	HP	PL	EL	VB	CI	DW	XN	ZL	YP	VN	HH	CC	
b	LP	QT	HE	RS	UR	CR	ZH	GV	WC	HL	YN	KT	WT	MC	KH	
c	DX	MN	AO	NH	SF	GI	WL	MN	AH	GR	BZ	HS	ZU	YM	WU	
d	KM	YZ	RY	FP	TR	CR	XE	JK	NY	PO	GJ	JR	PE	MO	VB	
e	QU	HP	QG	JQ	YQ	OB	SA	NL	PX	OP	VS	AF	XK	XR	UQ	
⋮																

Здесь взаимообратные биграммы (ao $\mapsto$ CC, cc $\mapsto$ AO; ah $\mapsto$ CI, ci $\mapsto$ AH; af $\mapsto$ gL, cl $\mapsto$ AF) не бросаются сразу в глаза.

Следующие шаги шифрования $V^2 \leftrightarrow V^2$ снова могут быть получены с помощью паролей, например, таких как /amerika/ и /equality/:

	a	m	e	r	i	k	b	c	d	f	g	h	j	l	n	...
e	XZ	KJ	YJ	HP	PL	EL	VB	CI	DW	XN	ZL	YP	VN	HH	CC	
q	LP	QT	HE	RS	UR	CR	ZH	GV	WC	HL	YN	KT	WT	MC	KH	
u	DX	MN	AO	NH	SF	GI	WL	MN	AH	GR	BZ	HS	ZU	YM	WU	
a	KM	YZ	RY	FP	TR	CR	XE	JK	NY	PO	GJ	JR	PE	MO	VB	
l	QU	HP	QG	JQ	YQ	OB	SA	NL	PX	OP	VS	AF	XK	XR	UQ	
⋮																

(с тем эффектом, что исчезают взаимообратные биграммы и само шифрование становится более громоздким).

Конструирование матрицы требует основательной работы по выравниванию символьных частот (разд. 3.1.2). При построенной надлежащим способом перестановке подобная имитация распределения частот символов соответствующего языка возможна. Идеальным результатом является матрица, у которой в каждой строке и в каждом столбце любая буква встречается лишь один раз как первый и как второй символ биграммы, например,

AB	BC	CA		AC	BA	CB	DD		AA	BB	CC	DD	EE
CC	AA	BB	или	BD	AB	DA	CC	или	BC	CD	DE	EA	AB
BA	CB	AC		DB	CD	BC	AA		CE	DA	EB	AC	BD
				CA	DC	AD	BB		DB	EC	AD	BE	CA
									ED	AE	BA	CB	DC

Такие матрицы называют «греко-латинскими квадратами»[1]. Помимо случая $N = 6$ («задача о 36 офицерах» Эйлера, 1779 г.), для всех натуральных чисел $N > 2$ существуют греко-латинские квадраты и, как правило, их бывает несколько. В любом случае матрица, пример, приводимый Элен Фуше Гейнс,

[1] В русскоязычной литературе такие матрицы принято называть латинскими квадратами. — *Прим. ред.*

$$\begin{array}{cccc} AA & AB & AC & AD & \dots \\ BA & BB & BC & BD & \dots \\ CA & CB & CC & CD & \dots \\ DA & DB & DC & DD & \dots \\ \vdots & \vdots & \vdots & \vdots \end{array}$$

не является подходящей, так как она приводит к односимвольному двухалфавитному шифрованию (многоалфавитное шифрование, разд. 8.2), вытекающему из перестановок символов в парах символов.

Следующая таблица показывает для $N = 26$, $N = 10$ и $N = 2$ значения чисел $Z(N)$ имеющихся квадратов $V^2 \leftrightarrow V^2$ (сравните с разд. 3.2.6):

число квадратов	$Z(N)$	$Z(26)$	$Z(10)$	$Z(2)$
всего	$(N^2)!$	$1.88 \cdot 10^{1621}$	$9.33 \cdot 10^{157}$	24
истинные отражения	$\approx (N^2!)^{1/2}$	$7.60 \cdot 10^{809}$	$2.72 \cdot 10^{78}$	3

На рис. 31 а показан классический пример шага $V^2 \leftrightarrow V^2$, применявшегося для шифрования позывных радиосети R.S.H.A. в Норвегии. Десять таких таблиц были составлены в VI управлении (возглавляемом до 1942 г. Йостом, а за-

K 1 Norw.	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	
a	ca	fn	bl	ou	ih	oo	il	bv	bw	er	rm	qm	mn	ab	zm	ns	wl	yc	zy	tr	du	wo	oa	ho	ic	pu	a
b	sk	wm	dg	ia	cw	pf	if	vd	da	xz	fo	dh	px	rr	iv	gh	mu	ae	qr	tb	og	sr	vu	qg	zt	pm	b
c	hp	no	ij	xp	ji	yf	eo	xh	zu	pl	ft	yv	qw	am	qp	lz	bg	be	lc	nw	ap	vx	rs	yi	wy	gi	c
d	ov	gg	tk	ys	hm	tx	eq	qa	iu	zo	ud	gj	lh	bn	fm	ta	ej	hi	jc	sv	vp	rd	br	rh	kt	tw	d
e	di	wz	qo	pz	ag	wk	fl	uo	ll	oe	ph	jq	gl	vy	lf	af	vt	cj	vq	yz	rz	fc	ps	pq	ro	aq	e
f	cu	rf	nt	xr	ya	tg	xj	db	sc	hg	zr	hs	em	xv	vr	ul	wn	sh	ku	my	va	ad	fg	zp	ut	lb	f
g	sx	hd	vk	st	lk	xf	gn	lv	yr	yd	xg	kr	hc	xl	xw	pa	au	eb	gb	li	id	rj	tz	xq	wd	rn	g
h	bq	oy	sb	mw	qx	zd	ar	po	on	rx	sj	om	as	mb	vs	ke	yy	xy	uj	hb	rc	jg	co	fq	jr	pe	h
i	cb	sl	ri	cf	qt	ek	un	kl	nx	to	hk	ew	yo	wp	kj	kh	su	xi	jo	of	dt	ml	zi	bk	qq	gu	i
j	vv	tf	fi	mp	ky	hl	qc	iq	na	gd	up	tq	hq	xs	xb	wt	ez	mm	hj	vg	eh	dc	qe	ti	uk	cg	j
k	uv	bt	bf	ux	kz	zw	ex	nh	ac	av	tt	aw	ye	dw	dy	nv	wf	dn	sf	eg	lg	wc	kx	ur	pc	od	k
l	ir	ea	kn	le	jb	nu	at	hu	zl	fw	ce	ka	jv	bm	ev	ak	cp	gm	yn	cd	kd	ue	xm	ig	fy	ht	l
m	mv	el	yg	ny	bu	cq	fk	wq	pk	oo	ms	sz	rl	pr	qi	te	qn	kf	gs	uc	kv	kc	dl	kp	cl	lp	m
n	je	sq	gz	ts	dk	vo	xo	ge	mj	qv	mi	dp	vf	rb	yj	bj	mg	vl	qs	uw	rq	pb	mh	lt	oz	qk	n
o	vc	gk	al	vz	np	vm	by	cm	re	wv	uz	yt	ww	gp	js	en	tv	jn	bo	tm	sp	or	fj	ub	ck	td	o
p	hr	ah	ik	xn	mo	zk	ds	in	dz	ym	ci	qu	dv	df	nk	yk	pt	iz	ef	ws	es	ip	fz	ss	jk	ct	p
q	ec	xc	jj	vb	vh	ot	pg	ib	ty	ch	pd	qz	qf	fd	oh	sa	bc	zj	ba	fp	nq	wa	ie	vi	oq	lw	q
r	wi	uq	ln	ja	gq	lo	rp	sd	ko	iy	si	mc	uu	io	yh	ru	xx	qy	fr	hy	ob	ox	nl	uh	fh	ga	r
s	zg	nf	sy	jw	nn	kq	vn	ld	go	mt	pn	jf	he	um	ua	za	xt	bb	op	qh	gf	yl	md	os	ju	ei	s
t	yw	wg	mx	ol	sw	se	rv	yp	us	rk	dx	zs	bz	dj	cn	mf	hx	de	it	ai	ug	mk	ql	cs	ix	pi	t
u	gy	fa	ow	gr	vw	bh	ly	kw	ry	mz	pj	sg	jz	gt	dd	nd	et	az	tp	jh	cx	iw	la	zq	rw	lm	u
v	gv	bi	oi	ii	zb	lj	hz	zh	nb	ks	cy	yq	jx	dq	ma	hf	wr	lq	jp	ng	gw	jl	rg	tl	lr	wh	v
w	aj	gx	nr	qb	uf	ok	rt	xu	bp	wb	qd	jt	mr	aa	pv	yu	nj	xd	eu	mq	hw	nz	ze	km	uy	tn	w
x	kb	yx	ui	pw	we	xk	fe	vj	gc	pp	ep	hh	zn	ha	zf	ax	do	py	nm	xe	ff	so	tc	sm	fb	fx	x
y	fs	ay	ni	wj	wu	fu	ed	an	fv	xa	cv	cz	bs	ve	th	cc	bx	ra	cr	im	ne	hn	zv	oj	yb	tj	y
z	kg	bd	wx	zz	zx	lu	jy	sn	zc	tu	is	ao	dr	ki	ls	ey	qj	ee	lx	hv	nc	dm	jd	me	jm	kk	z
	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	Pnr. 0033

Рис. 31а. Двухсимвольное шифрование биграммами позывных радиосети R.S.H.A. в Норвегии.

тем Шелленбергом), в отделе зарубежной разведки гиммлеровского R.S.H.A. (Reichssicherheitshauptamt), по-видимому, Андреасом Фиглем (1873–1967 гг.). R.S.H.A. захватило в 1938 г. вместе с полезными документами австрийского полковника Андреаса Фигля, последнего главу австрийской Dechiffrierdienst («Chiffrengruppe»), когда Австрия была присоединена к Германии *(Аншлюс)*. Важность этой добычи была «обнаружена» молодым австрийским штурмбаннфюрером СС доктором Вильгельмом Хеттлем (1915–1999 гг.), который позже стал заместителем руководителя венского отделения VI E[2]). Фигль, живший до 1941 г. под «охраной» СС, работал в качестве «советника» в Ванзее под Берлином.

Хеттль также помог VI управлению R.S.H.A., когда в середине 1944 г. привлек к работе группу венгерских армейских криптологов в Будапеште, возглавляемую майором Бибо, который в 1944 г. успешно проник на линию связи Аллена Даллеса Берн—Вашингтон.

Фигль, очень умный криптоаналитик, был капитаном, когда в 1911 г. создавал криптоаналитическое бюро *k. u. k. Armee*. В 1915 г. майор Фигль уже вскрывал итальянские криптограммы, а в 1926 г. он имел чин полковника, написав хороший учебник «*Systeme des Chiffrierens*» (243 с., 45 приложений, Грац, 1926 г.). Печать подготовленного второго тома, «*Systeme des Dechiffrierens*», в 1926 г. была запрещена; копия (но не оригинал рукописи) теперь, кажется, стала доступна.

Взаимообратное двухсимвольное шифрование биграммами использовалось с 1 мая 1937 г. для перешифрования индикаторов *(Spruchschlüssel)*, которые помещались в начале радиограмм, зашифрованных с помощью ENIGMA ВМФ. Имелся выбор между десятью такими таблицами с именами наподобие FLUSS (рис. 31b), BACH, STROM, TEICH, UFER, бывшими в употреблении; они были известны англичанам, которые получили их с потопленных субмарин (U-110 в мае и Gedania в июне 1941 г.; VP 5904 в январе 1942 г.; U505 в июне 1944 г.) или, возможно, выкрали их или реконструировали позже. Индикатор, наподобие /psq/, должен был выбираться наугад; затем он удваивался до /psqpsq/ и зашифровывался на ENIGMA с заданной установкой роторов *(Grundstellung)*, например, /iaf/; шифр, скажем SWQRAF, разбивался на две части:

$$\begin{matrix} S & W & Q & * \\ * & R & A & F \end{matrix}$$ и заполнялся пустышками («символы заполнения»): $$\begin{matrix} S & W & Q & X \\ P & R & A & F. \end{matrix}$$

Двухсимвольное шифрование производилось вертикальными парами, например, такими:

$$\begin{matrix} S \\ P \end{matrix} \leftrightarrow \begin{matrix} Q \\ A, \end{matrix} \quad \begin{matrix} W \\ R \end{matrix} \leftrightarrow \begin{matrix} F \\ P, \end{matrix} \quad \begin{matrix} Q \\ A \end{matrix} \leftrightarrow \begin{matrix} C \\ D, \end{matrix} \quad \begin{matrix} X \\ F \end{matrix} \leftrightarrow \begin{matrix} S \\ Z. \end{matrix}$$

[2]) После Второй мировой войны Хеттль сыграл неудачную роль в австрийской правой партии *Wahlpartei der Unabhangigen*; он был тщеславен («я был главным шпионом Гитлера») и также написал книги («*Секретный фронт*», 1954 г. и под псевдонимом Вальтер Хаген «*Бумажное оружие*», 1955 г.).

Geheim! Prüfnr. 516

Kennwort: Fluß

Doppelbuchstabentauschtafel für Kenngruppen — Tafel B

AD	AE	AJ	AK	AL	AN	BC	BD	BE	CE	CF	CG	DG
AA = RN	BA = IK	CA = KJ	DA = PK	EA = TC	FA = XP	GA = NE	HA = JR	IA = NN	JA = WE	KA = EI	LA = EU	MA = RG
B = KW	B = RT	B = PO	B = EZ	B = JX	B = OI	B = JO	B = NO	B = VF	B = OY	B = GW	B = KH	B = IP
C = PM	C = EY	C = JV	C = AW	C = OM	C = IU	C = BK	C = GY	C = DN	C = NQ	C = IM	C = VO	C = WW
D = YE	D = AK	D = BM	D = JM	D = MJ	D = RB	D = FL	D = TB	D = FW	D = KK	D = SE	D = YA	D = TA
E = NR	E = OW	E = MZ	E = WB	E = HY	E = PA	E = ZT	E = ZI	E = RP	E = TN	E = AG	E = CV	E = BQ
F = UC	F = WQ	F = EK	F = XY	F = AS	F = DZ	F = SA	F = QY	F = EO	F = VS	F = JH	F = SC	F = KV
G = KE	G = QA	G = KT	G = ZA	G = PU	G = NV	G = LR	G = OA	G = WS	G = FR	G = PN	G = JU	G = NS
H = XU	H = ZZ	H = AZ	H = BS	H = WO	H = ZK	H = TP	H = CU	H = NU	H = KF	H = DT	H = ZQ	H = VK
I = PC	I = OG	I = ND	I = MT	I = HA	I = QR	I = MW	I = QS	I = TM	I = PM	I = LV	I = RX	I = XC
J = JP	J = HQ	J = TQ	J = OB	J = GZ	J = LN	J = AU	J = IS	J = XO	J = SV	J = CA	J = WZ	J = ED
K = BD	K = GC	K = GX	K = FP	K = CF	K = EL	K = QN	K = PG	K = BA	K = IT	K = JD	K = EM	K = ZF
L = QI	L = PR	L = RE	L = RI	L = FK	L = GD	L = WH	L = KR	L = MS	L = UP	L = TO	L = OK	L = DR
M = HT	M = CD	M = WA	M = VV	M = LK	M = AC	M = PB	M = SF	M = KC	M = DD	M = BW	M = TR	M = SU
N = MR	N = NL	N = OS	N = IC	N = TY	N = CP	N = OX	N = SZ	N = QZ	N = PX	N = UX	N = FJ	N = LO
O = BZ	O = US	O = DY	O = YJ	O = IP	O = VE	O = JT	O = FY	O = YV	O = GB	O = QC	O = MN	O = NX
P = XI	P = SX	P = FN	P = HF	P = NC	P = DK	P = RY	P = MX	P = MB	P = AJ	P = VJ	P = BT	P = FZ
Q = OZ	Q = ME	Q = QF	Q = GU	Q = WV	Q = PY	Q = IZ	Q = BJ	Q = OV	Q = XH	Q = RS	Q = IV	Q = OJ
R = UK	R = YN	R = XJ	R = ML	R = KS	R = JG	R = CY	R = OP	R = SH	R = HA	R = HL	R = GG	R = AN
S = EF	S = DH	S = ZB	S = QG	S = QW	S = UE	S = RF	S = RJ	S = HJ	S = YZ	S = ER	S = NW	S = IL
T = IY	T = LP	T = SW	T = KH	T = XD	T = SR	T = XV	T = AM	T = JK	T = GO	T = CG	T = UF	T = DI
U = GJ	U = XK	U = IH	U = WH	U = LA	U = WX	U = DQ	U = UQ	U = FC	U = LG	U = XZ	U = XW	U = DY
V = QU	V = TI	V = LE	V = HW	V = RL	V = TL	V = UM	V = LZ	V = LQ	V = CC	V = MF	V = KI	V = UT
W = DC	W = KM	W = VP	W = SO	W = SK	W = ID	W = KB	W = DV	W = PH	W = QL	W = AB	W = PW	W = GI
X = UV	X = VY	X = UG	X = HT	X = UZ	X = YS	X = CK	X = WJ	X = UD	X = EB	X = ZY	X = PP	X = HP
Y = SG	Y = MU	Y = GR	Y = CO	Y = DC	Y = HO	Y = HC	Y = VN	Y = AT	Y = TU	Y = NZ	Y = QD	Y = VD
Z = CH	Z = AO	Z = YJ	Z = FF	Z = DG	Z = MP	Z = EJ	Z = YD	Z = GQ	Z = UW	Z = WP	Z = HV	Z = CE

Fortsetzung s. Rückseite!

Kennwort: Fluß

Tafel B

DH	DJ	DR	DS	DT	EG	EH	EJ	EK	EQ	ER	ES	ET
NA = TZ	OA = HG	PA = FB	QA = BG	RA = QH	SA = GF	TA = MD	UA = QX	VA = ON	WA = CM	XA = TX	YA = LD	ZA = DG
D = QV	B = ZX	B = GM	B = ZD	B = FD	B = OT	B = HD	B = SD	B = MY	B = DE	B = UL	B = VG	B = CS
C = EP	C = TH	C = AI	C = KO	C = PL	C = LF	C = EA	C = AP	C = ZO	C = QJ	C = MI	C = WL	C = SI
D = CI	D = XS	D = NH	D = LY	D = OQ	D = UB	D = ZN	D = IX	D = SY	D = PV	D = ET	D = HZ	D = QB
E = GA	B = DJ	E = QT	E = TJ	E = CL	E = KD	E = YX	E = FS	E = FO	E = JA	E = WM	E = AD	E = TT
F = DP	F = QM	F = XX	F = CQ	F = GS	F = HM	F = RO	F = LT	F = IB	F = VT	F = ZL	F = OR	F = MK
G = XM	G = BI	G = HK	G = DS	G = MA	G = AY	G = WK	G = CX	G = YB	G = ZM	G = SS	G = VQ	G = RM
H = PD	H = NP	H = IW	H = RA	H = LB	H = IR	H = OC	H = ZJ	H = RK	H = GL	H = JQ	H = QQ	H = VL
I = TW	I = FB	I = ZR	I = AL	I = DL	I = ZC	I = BV	I = ST	I = XR	I = YR	I = AP	I = CZ	I = HE
J = VR	J = MQ	J = TS	J = WC	J = HS	J = PQ	J = QE	J = NM	J = KP	J = HX	J = CR	J = DO	J = UH
K = YU	K = LL	K = DA	K = SN	K = VH	K = EW	K = XN	K = AR	K = MH	K = TG	K = BU	K = WR	K = FH
L = BN	L = RZ	L = RC	L = JW	L = EV	L = VX	L = FV	L = XB	L = ZH	L = YC	L = ZP	L = SQ	L = XF
M = UJ	M = EC	M = JI	M = OF	M = ZG	M = ZV	M = II	M = GV	M = ZU	M = XE	M = NG	M = VW	M = WG
N = IA	N = VA	N = KG	N = GK	N = AA	N = QK	N = JE	N = YY	N = HY	N = DU	N = TK	N = BR	N = TD
O = IB	O = UU	O = CB	O = VZ	O = TF	O = DW	O = KL	O = TV	O = LC	O = EH	O = IJ	O = PZ	O = VC
P = OH	P = HR	P = LX	P = XT	P = IE	P = RV	P = GH	P = JL	P = CW	P = KZ	P = FA	P = ZS	P = XL
Q = JC	Q = RD	Q = SJ	Q = YH	Q = UR	Q = YL	Q = CJ	Q = HU	Q = YG	Q = BF	Q = YT	Q = RW	Q = LH
R = AE	R = YF	R = BL	R = FI	R = WU	R = FT	R = LM	R = RQ	R = NJ	R = YK	R = VI	R = WI	R = PI
S = MG	S = CN	S = UY	S = HI	S = KQ	S = XG	S = PJ	S = BO	S = JP	S = IG	S = OD	S = FX	S = YP
T = DX	T = SB	T = WY	T = PE	T = BB	T = UI	T = ZE	T = MV	T = WF	T = OU	T = QP	T = XQ	T = GE
U = II	U = WT	U = EG	U = AV	U = ZW	U = MM	U = JY	U = OO	U = YW	U = RR	U = AH	U = NK	U = VM
V = FG	V = IQ	V = WD	V = NB	V = SP	V = JJ	V = UO	V = AX	V = DM	V = EQ	V = GT	V = IO	V = SM
W = LS	W = DE	W = LW	W = ES	W = YQ	W = CT	W = NI	W = JZ	W = YM	W = MC	W = LU	W = VU	W = RU
X = MO	X = GN	X = JN	X = UA	X = LI	X = BP	X = XA	X = KN	X = SL	X = FU	X = PF	X = TE	X = OB
Y = EE	Y = JB	Y = FQ	Y = HF	Y = GP	Y = VD	Y = EN	Y = PS	Y = BX	Y = PT	Y = DF	Y = UN	Y = KX
Z = KY	Z = AQ	Z = YO	Z = IN	Z = OL	Z = HN	Z = NA	Z = EX	Z = QO	Z = LJ	Z = KU	Z = JS	Z = BH

Рис. 31b. Взаимообратная таблица биграмм FLUSS немецкого ВМФ (*Kriegsmarine*)

Это давало следующий результат:

Q F C S
A P D Z.

Зашифрованный индикатор QFCSAPDZ передавался. Получателем выполнялась обратная процедура: сначала (взаимообратная) двухсимвольная подстановка биграмм и удаление пустышек, затем (взаимообратное) шифрование на ENIGMA с заданной установкой роторов /iaf/; результат должен был иметь вид 123 123, первая часть которого указывала индивидуальную установку роторов для сообщения. Таким способом устанавливался ключ между двумя абонентами. Согласование ключей было и по сей день остается специфической слабостью криптологии.

Процедура выглядела достаточно сложной, что породило ощущение защищенности у ее авторов. Для англичан, однако, препятствия оказались преодолимыми. ENIGMA была успешно взломана.

Англичане должны были бы быть предупреждены. Однако британский торговый флот использовал двухсимвольную подстановку биграмм для перешифрования используемого им кода BAMS. Кодовая книга попала в руки немцев в 1940 г., когда немецкий рейдер «*Atlantis*» захватил судно «*City of Bagdad*» в Индийском океане. «*Служба наблюдения*» (*B-Dienst*) немецкого ВМФ до 1943 г. успешно снимала перешифрование радиосигналов союзнических судов.

Для перешифрования цифровых кодов годится перестановка $Z_{10}^2 \leftrightarrow Z_{10}^2$, называемая *Geheimklappe*. Это двухсимвольная подстановка биграмм была введена в марте 1918 г. немцами для использования в тактической связи на Западном фронте с одной таблицей для зашифрования и другой таблицей для расшифрования (рис. 32). К концу Первой мировой войны эта двухсимвольная подстановка биграммами менялась каждый день.

Verschlüsselungstafel.

	0	1	2	3	4	5	6	7	8	9
0	23	48	60	05	78	35	58	64	29	52
1	20	77	33	59	21	70	02	40	63	08
2	11	49	01	69	47	41	79	74	22	42
3	32	76	39	18	75	30	09	51	80	65
4	61	19	43	81	06	56	73	62	10	28
5	85	50	24	88	31	84	27	90	55	57
6	03	01	96	53	68	16	44	89	15	87
7	97	25	71	04	95	34	14	37	93	38
8	26	72	54	92	13	83	45	00	66	67
9	86	12	98	36	99	46	82	17	94	07

Entschlüsselungstafel.

	0	1	2	3	4	5	6	7	8	9
0	87	22	16	60	73	03	44	99	19	36
1	48	20	91	84	76	68	65	97	33	41
2	10	14	28	00	52	71	80	56	49	08
3	35	54	30	12	75	05	93	77	79	32
4	17	25	29	42	66	86	95	24	01	21
5	51	37	09	63	82	58	45	59	06	13
6	02	40	47	18	07	39	88	89	64	23
7	15	72	81	46	27	34	31	11	04	26
8	38	43	96	85	55	50	90	69	53	67
9	57	61	83	78	98	74	62	70	92	94

Рис. 32. Двухсимвольная подстановка биграммами (*Geheimklappe*) для перешифрования числовых кодов

4.1.3. Двухсимвольная подстановка триграммами $V^2 \to W^3$. Иногда используется десятеричная двухсимвольная подстановка триграммами ($V = Z_{26}$, $W = Z_{10}$):

	a	b	c	d	e	...
a	148	287	089	623	243	
b	243	127	500	321	601	
c	044	237	174	520	441	
d	143	537	188	257	347	
⋮	⋮	⋮	⋮	⋮	⋮	

Криптоаналитически, отображение $V^2 \to W^{(n)}$ при любом n попадает в один и тот же класс и может интерпретироваться как $|V|$-мерная омофонная простая подстановка нечетных символов плюс $|V|$-мерная омофонная простая подстановка четных символов. Поэтому при наличии достаточного объема доступного материала легко взломать шифр, если, как в примере отображения $V^2 \leftrightarrow V^2$ Элен Фуше Гейнс, используется стандартный табличный шифр. Эйрауд указывает, что сложность частного метода шифрования нарезкой сообщения на две половинки и записывании их в две строки, с использованием двухсимвольной подстановки для вертикальных пар, является *иллюзорной.*

4.2. Специальные случаи шифров Плейфейра и Деластеля: томографические методы

4.2.1. Шифр Плейфейра. В 1854 г. Чарльз Вейтстоун изобрел специальную двухсимвольную подстановку биграмм (рис. 33); его друг Лион Плейфейр рекомендовал ее высоким правительственным и военным чинам. Система, возможно, впервые применялась в Крымской войне и, как сообщалось, — в войне с бурами; имя Плейфейра стало ее именем. Военные оценили ее как полевой шифр, потому что она не требовала ни таблиц, ни аппарата. Британская армия приняла этот шифр на вооружение примерно на стыке столетий и продолжала сохранять его в тайне. Однако в Первой мировой войне, к середине 1915 г., немцы обычно успешно читали его.

Шаг шифрования PLAYFAIR производится следующим образом: с помощью пароля перемешанный алфавит Z_{25} (скажем, опустив букву J из алфавита Z_{26}) записывался в квадрат 5×5 (фр. *damier*)[3]:

P	A	L	M	E			T	O	N	R	S
R	S	T	O	N			D	F	G	B	C
B	C	D	F	G		или	K	Q	U	H	I
H	I	K	Q	U			X	Y	Z	V	W
V	W	X	Y	Z			L	M	E	P	A

который считался закольцованным подобно тору (так что эти два квадрата по сути означают одно и то же). Теперь, если два символа биграммы находятся в одной и той же строке (или столбце), каждый заменяется символом

[3] Вейтстоун фактически использовал алфавиты, которые были лучше перемешаны (разд. 3.2.5), и прямоугольные матрицы. Эти важные меры по обеспечению безопасности были вскоре понижены.

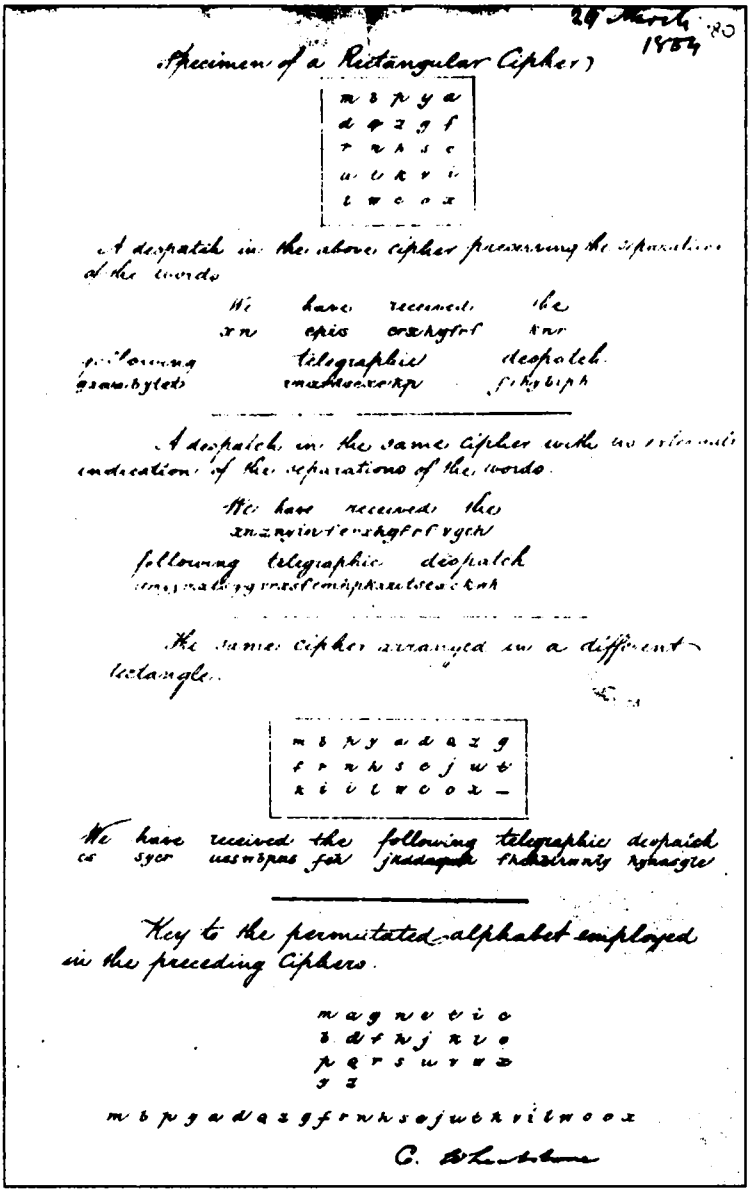


Рис. 33. Описание шифра PLAYFAIR, подписанного его изобретателем Чарльзом Вейтстоуном, 26 марта 1854 г.

справа (или ниже его соответственно); например, оба квадрата дают последовательно подстановки

$$\mathrm{am} \mapsto \mathrm{LE} \qquad \mathrm{dl} \mapsto \mathrm{KT}.$$

В противном случае первый символ заменяется символом в той же самой строке, но в столбце второго символа; аналогично второй символ заменен символом в той же самой строке, но в столбце первого символа («шаг перекрещивания», фр. *substitution orthogonale* и *diagonale*). Таким образом,

$$\mathrm{ag} \mapsto \mathrm{EC} \qquad \mathrm{ho} \mapsto \mathrm{QR}.$$

Шаг не определен, если биграмма является удвоенным символом[4)] или если последний символ не имеет пары. Эту ситуацию избегают, вставляя букву x:

ba ll oo n заменяется на ba lx lo on;
le ss se ve n на le sx sx se ve nx.

Это — опасная слабость. Тем не менее, шаг PLAYFAIR привлекает своей относительной простотой. Но из-за торической симметрии его комбинаторная сложность даже меньше, чем у простой подстановки.

В третьем случае из рассмотренных выше, шаг PLAYFAIR может интерпретироваться как произведение отображений: отображение биграмм открытого текста в пары координат строка/столбец, перестановка столбцовых координат и перевод в биграмму (наподобие перевертыша, разд. 6.1.2):

	1	2	3	4	5
1	P	A	L	M	E
2	R	S	T	O	N
3	B	C	D	F	G
4	H	I	K	Q	U
5	V	W	X	Y	Z

a g
12 35
×
15 32
E C

Такие композиции шифрований, которые равносильны разложению и рекомбинации, называют томографическими или методами фракций (fractionating) («*chiffres à damiers*», Огюст Л. А. Колон, 1899 г.); мы изучим их в разд. 9.4.4.

4.2.2. Модифицированный шифр PLAYFAIR. Такой шаг использовался немецкой армией и СД (государственная политическая полиция), начиная с середины 1941 г. под названием *Handschlüssel*, и регулярно взламывался до осени 1944 г. англичанами в Блэтчли Парк при полковнике Джоне Х. Тильтмане, который с 1924 г. был главой военного отдела. Он назывался двойным гробом (нем. *Doppdkas/set/tenverfahren*), а также двойным PLAYFAIR, двухтабличным PLAYFAIR, потому что в нем использовались

[4)] Совет Маттео Ардженти подавлять стоящие рядом одинаковые символы (разд. 3.4.1), вероятно, был забыт.

(при опущенной, скажем, букве J) два различных 5×5 квадрата, например,

A	Y	K	I	H		Y	X	U	H	A
L	B	M	N	P		T	R	K	B	I
Q	R	C	O	G		P	M	C	G	S
Z	X	V	D	S		F	D	L	Q	V
F	W	U	T	E		E	N	O	W	Z

Они не конструировались из паролей, а формировались «наугад» и затем распределялись. Как и в исходном шифре PLAYFAIR, здесь производился двухсимвольный биграммный шаг наподобие

$$\text{ah} \mapsto \text{AY}, \quad \text{nr} \mapsto \text{KP}, \quad \text{nb} \mapsto \text{IP},$$

если символы открытого текста расположены в той же самой строке (закольцованной подобно цилиндру), во всех других случаях используется «шаг перекрещивания»:

$$\text{xe} \mapsto \text{FW}, \quad \text{or} \mapsto \text{MC}, \quad \text{bx} \mapsto \text{RY}.$$

Кроме того, открытый текст нарезался на группы фиксированной длины. Так например, сообщение

anxobergruppenfuehrerxvonxdemxbachxkiewxbittexdreixtausendx
schussxpatronenxschickenstop

(с символом /x/ в качестве пробелов, которые не подавлялись (разд. 2.4.2)) нарезается на группы, скажем, по 17 символов, и каждая группа шифруется следующим образом:

a	n	x	o	b	e	r	g	r	u	p	p	e	n	f	u	e
h	r	e	r	x	v	o	n	x	d	e	m	x	b	a	c	h
A	K	F	M	N	Z	C	M	M	N	T	R	N	I	Z	O	W
Y	P	W	N	Y	S	W	E	Y	V	E	G	H	P	A	C	H

Эта процедура применяется еще раз: ay $\mapsto$ XY, kp $\mapsto$ YC, fw $\mapsto$ ZW, ... так, что окончательный шифротекст, разбитый на пятисимвольные группы, имеет вид

XYYCZ WRUPY VQGUT UTKID ...

Модифицированный PLAYFAIR подобно классическому, несколько громоздок и подвержен ошибкам; это приводит к частым повторным запросам и угрожает стойкости шифрования, идущего на компромисс (см. гл. 11). Сбои при шифровании со стороны немцев помогли англичанам в такой же степени, как и прусская склонность к «методичности и любезности» и любовь к титулам и другим формальностям.

4.2.3. Шифр Деластеля. Томографический метод в самой чистой форме («при поиске метода двухсимвольного шифрования, который бы не требовал громоздких кодовых таблиц размера 26 × 26», Кан) был опубликован в 1901 г. Феликсом Мария Деластелем (1840–1902 гг.), автором «*Traité Elementaire de Cryptographie*» (Gauthier-Villars, Париж 1902 г.). Это была взаимно однозначная простая (т. е. односимвольная) подстановка биграмм (очень похожая на квадрат Полибия) с последующим сдвигом на четыре позиции и, наконец, инверсия той же самой простой подстановки биграмм, например,

	1	2	3	4	5
1	B	O	R	D	E
2	A	U	X	C	F
3	G	H	I	J	K
4	L	M	N	P	Q
5	S	T	V	Y	Z

o	n
12	43
×	
14	23
D	X

или

o	n	
1	4	D
2	3	X

Шаг шифрования взаимообратен и приводит к двухсимвольной подстановке биграммами, наподобие рассмотренной в разд. 4.1.2. Для обращения подстановки может использоваться сопряженная простая подстановка биграмм; тогда взаимообратный символ исчезает.

Заметим, что простой сдвиг на одну позицию, *Kulisseaverfahren* (Rohrbach, 1948 г.) не дает желаемого эффекта:

...		a		b		s		a		l		o		m		...
3	2	1	1	1	5	1	2	1	4	1	1	2	4	2	3	
	H		B		E		O		D		B		C		X	

В данном случае криптоаналитику нет нужды восстанавливать квадрат: достаточно интерпретировать шаг шифрования как отображение $V \to V^2$ с омофонами:

$$\mathrm{a}\mapsto \begin{Bmatrix}\mathrm{O}\\ \mathrm{U}\\ \mathrm{H}\\ \mathrm{M}\\ \mathrm{T}\end{Bmatrix} \times \begin{Bmatrix}\mathrm{B}\\ \mathrm{O}\\ \mathrm{R},\\ \mathrm{D}\\ \mathrm{E}\end{Bmatrix} \quad \mathrm{b}\mapsto \begin{Bmatrix}\mathrm{B}\\ \mathrm{A}\\ \mathrm{G}\\ \mathrm{L}\\ \mathrm{S}\end{Bmatrix} \times \begin{Bmatrix}\mathrm{B}\\ \mathrm{O}\\ \mathrm{R},\\ \mathrm{D}\\ \mathrm{E}\end{Bmatrix} \quad \mathrm{s}\mapsto \begin{Bmatrix}\mathrm{E}\\ \mathrm{F}\\ \mathrm{K}\\ \mathrm{Q}\\ \mathrm{Z}\end{Bmatrix} \times \begin{Bmatrix}\mathrm{B}\\ \mathrm{O}\\ \mathrm{R},\\ \mathrm{D}\\ \mathrm{E}\end{Bmatrix} \times \begin{Bmatrix}\mathrm{B}\\ \mathrm{O}\\ \mathrm{R},\\ \mathrm{D}\\ \mathrm{E}\end{Bmatrix} \ldots$$

с граничным условием перекрытия:

$$\mathrm{a\,b\,s\,a\,l\,o\,m} \mapsto \mathrm{HB}\smile\mathrm{BE}\smile\mathrm{EO}\smile\mathrm{OD}\smile\mathrm{DB}\smile\mathrm{BC}\smile\mathrm{CX}.$$

А это открывает неожиданное направление атаки.

Более ранний пример цифрового томографического метода обнаруживается в публикации 1876 г. датского инженера Алексиса Коэла (разд. 9.4.6). Он связан с методом с Плиния Эрла Чейза, изобретенным в 1859 г. (разд. 9.5.4).

4.3. Случай $V^3 \to W^{(m)}$ (трехсимвольные подстановки)

4.3.1. Джиоппи. При триграммной замене в ее полном виде быстро возникают технические трудности. Бумага, к сожалению, не является трехмерной; так что распечатка триграмм является более громоздкой, и $26^3 = 17\,576$ триграмм является приличным числом, требующим буклет в 26 страниц для их перечисления. Триграммные замены сложно механизировать простыми средствами, хотя здесь может сильно помочь использование даже несложных калькуляторов. Специальные замены триграммами наподобие PLAYFAIR были не очень успешными; в 1897 г. в Милане такую систему опубликовал граф Луиджи Джиоппи ди Тюркхейм. Уильям Фридман также имел дело с заменами триграмм примерно в 1920 г., и это привлекает к ним некоторый интерес. В очень специальном случае линейных подстановок (гл. 5) имеются триграммные замены, применявшиеся Джеком Левином (1958, 1963 гг.).

4.3.2. Хенкельс. Шифровальная машина, которая механически выполняет четырехсимвольную подстановку квартетами, была запатентована в 1922 г. Хенкельсом.

4.4. Общий случай $V^{(n)} \to W^{(m)}$: коды

Вместо того, чтобы зашифровывать $26^3 = 17\,576$ триграмм, возможно было бы лучше зашифровать несколько сотен, тысячу или даже десятков тысяч часто встречающихся мультиграмм различной длины; это означает, что шаг шифрования оперирует с подмножеством C множества $V^{(n)}$ (с достаточно большим n) при условии, что каждый открытый текст $x \in V^{(n)}$ может быть разложен на элементы из множества C:

$$x = x_1 \star x_2 \star x_3 \star \ldots \star x_k \quad \text{(для некоторого } k \in \mathbb{N} \text{ и подходящих } x_j \in C \subseteq V^{(n)}\text{)}.$$

Это может быть гарантировано следующим «однобуквенным условием»: $C \supseteq V$.

Следуя Кану, такое шифрование называется кодом, если выбор множества C задается лингвистически: в кодовой книге перечисляются частые дифтонги, слоги, префиксы, окончания, слова, фразы вместе с их кодовыми группами.

Однобуквенное условие гарантирует, что могут быть зашифрованы даже диковинные, необычные, странные слова, включая биологические и химические термины или названия местностей, рек, гор и имена собственные. Конечно, это не означает, что каждое слово должно быть разложено на отдельные буквы, а каждое предложение должно быть разбито на слова — напротив, чем длиннее фраза, найденная в кодовой странице, тем лучше; наилучшим решением является то, которое нуждается в минимуме обращений к кодовой книге. В общем случае оптимум не может быть определен однозначно, но эта неопределенность не причинит вреда.

Поддерживать дисциплину кодирования трудно. В 1918 г. командующий *американскими экспедиционными силами* (A. E. F) во Франции имел причины предупредить штаб о том, что слово *boche* (бош — жаргонное имя немцев),

требующее пяти кодовых групп при записи по буквам, необходимо заменять словом *German*, которое записывается одной кодовой группой; и что восемнадцать кодовых групп, необходимых для записи *almost before the crack of dawn* (*почти перед рассветом*) лучше заменить на две кодовые группы для записи *day break* (*рассвет*). Использование кодов требует высокого уровня образования, так как хорошее кодирование является интеллектуальной задачей; плохое же кодирование помогает криптоаналитику взломать код. Таким образом, коды должны быть отвергнуты, если отсутствуют подходящие люди. Во время Первой мировой войны лейтенант Ягер из штаба 5-й немецкой армии оказал большую услугу противнику, когда из благих намерений поддержания безопасности связи регулярно подписывал свои распоряжения своим именем, которое, к сожалению, отсутствовало в кодовой книге и поэтому должно было записываться каждый раз по буквам. Как пишет Кан, он «был любимцем своих противников, потому что обеспечивал информацией о смене кода». В 1918 г. Ягер подставил под удар как перешифрование с помощью шифра *Geheimklappe*, так и новую кодовую книгу *Schlüsselheft*.

Дисциплина кодирования у американцев в Первой мировой войне была даже хуже, согласно оценке шефа G.2 A.6 майора Франка Мурмана, который возлагал за это ответственность на себя; это объясняется «хорошо известным американским игнорированием инструкций — особенно наиболее строгих из них» (Кан).

Во время Второй мировой войны ситуация улучшилась ненамного. В каждый штаб были назначены офицеры шифровального управления. Однако были еще и дипломаты. Отрицательным героем здесь является рузвельтовский дипломат Роберт Мёрфи (1894–1978 гг.), который по причинам престижа ставил при всяком использовании дипломатического кода стереотипные начала «For Murphy» или «From Murphy». Они помогли группе Рорбака из немецкой *Auswartiges Amt* взломать код. Фрейлейн Аста Фридрихс, принимавшая участие в этой работе, сказала после войны, что после своего задержания в Марбурге она видела его однажды и: «*Ich wollte ihn anhalten und ihm die Hand schüttein, — so viel hatte er für uns getan*». [Я хотела остановить его и пожать ему руку — так много сделал он для нас.]

4.4.1. Номенклаторы. Самыми старыми системами кодирования, известными на Западе, являются китайские идеограммы — хотя сами китайцы таковыми их не считают. В самом деле, недостаток криптологических достижений в древних культурах Китая объясняется тем фактом, что письменные сообщения были понятны только немногим. С другой стороны, египетские иероглифы основывались — 2000 лет до н. э. — на принципе, лежащем в основе построения ребуса и на акрофонии. Графический символ «ger» (свинья) сопоставлялся букве /r/, графический символ «wr», означая как «ласточку» так и «большой», сопоставлялся букве /wr/; специальные метки (детерминативы) указывают, в случае необходимости, на различия. В иероглифическом письме, если это необходимо, слово может быть даже расчленено на отдельные буквы для одиночных согласных. В этом письме нет элементов секретности. Но их также нет и в том случае, когда дипломаты, выучив постепенно используе-

мый код наизусть, способны выдать экспромтом речь, состоящую из кодовых слов и выражений, как это проделал с кодом GRAY американский консул в Шанхае в его речи на прощальном обеде в начале 1920-х гг. (разд. 4.4.7).

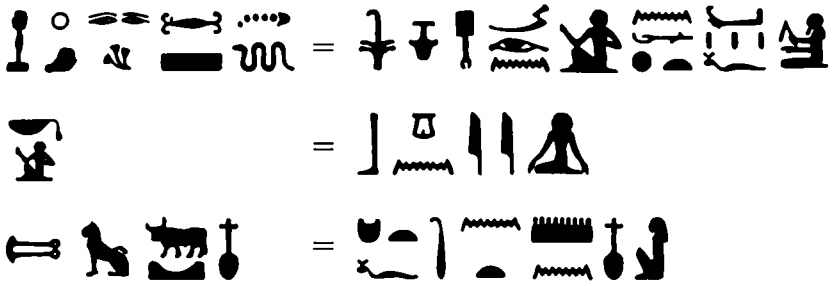


Рис. 34. Иероглифические надписи: необычные формы (слева) и обычные иероглифы (справа)

В египетских надписях (рис. 34) с необычными графемами, найденными рядом с открытым текстом, секретность не является самоцелью; напыщенный эпиграф на надгробной плите должен впечатлить и околдовать тайной магической мощью. На Западе первые смеси односимвольных подстановок и многосимвольных элементов обнаруживаются примерно в 1380 г., первоначально лишь для некоторых очень частых слов, среди которых *et* (см. рис. 26), *con*, *non*, *che* (разд. 3.4). Как только эти собрания стали несколько более пространными, их назвали номенклаторами. Ранний пример номенклатора показан на рис. 35.

Номенклаторы сохраняли свое большое значение на протяжении всей эпохи Возрождения. Король Англии Карл I использовал номенклатор с омофонами, восстановленный в 1860 г. Вейтстоуном: буква /a/ заменялась на 12 ... 17, /b/ — на 18 ... 19 и т. д., /france/ — на 9476. К 1600 г. некоторые номенклаторы содержали несколько сотен элементов, а цифровые группы в них имели коды из трех цифр.

Но и номенклаторы также взламывали. Испанский король Филипп II вручил своему посланнику Хуану де Морео номенклатор, содержащий приблизительно 400 кодовых групп. Виет трудился над его взломом с 28 октября 1589 г. до 15 марта 1590 г., после чего передал законченное решение своему королю Генриху IV. Филипп, узнав, что шифр, который он считал невскрываемым, скомпрометирован, пожаловался римскому папе на то, что Генрих использует черную магию. Римский папа был лучше информирован: его собственный криптолог Джованни Баттиста Ардженти также хорошо обслуживал папу, и Филипп заслужил лишь насмешку.

В политических интригах коды использовались, начиная от записки Марии Стюарт 1587 г. до осуждения французских анархистов, которые были обвинены в суде Сент-Этьена в 1892 г. на основе секретных сообщений, расшифрованных Базерье.

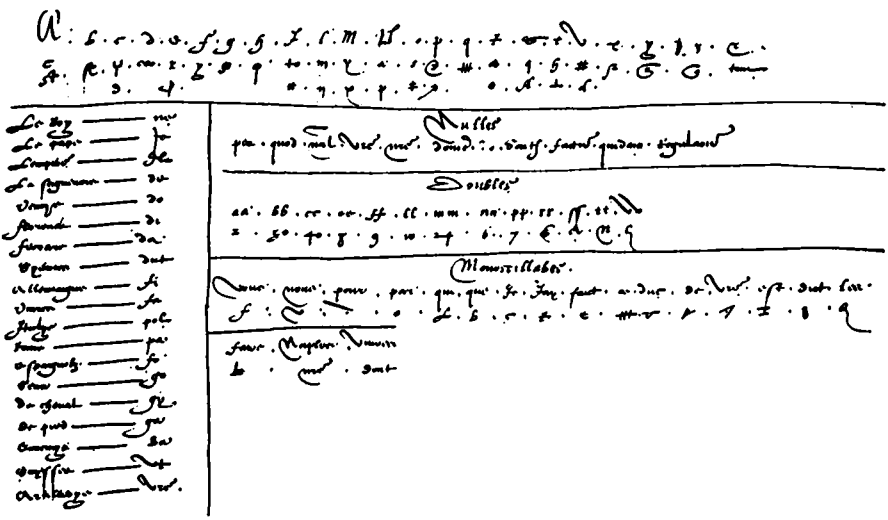


Рис. 35. Ранний номенклатор из Флоренции, 1554

В XVII веке не только итальянские княжества, но и все крупные европейские дворы имели свои «черные кабинеты» (фр. — *Cabinet Noir*, нем. — *Geheimkabinett*). Государственные деятели держали крупных криптологов в качестве помощников и доверенных лиц: у Людовика XIV служил Антуан Россиньоль, на русскую царицу работал Христиан Гольдбах, у Карла II служил Джон Валлис, а Мария Тереза имела барона Игнаца де Коха. Этим людям хорошо платили и оценивали важность их работы. Кан пишет:

«Хотя Валлис просил Ноттингема не предавать гласности его решения по дешифрованию из опасения, что Франция снова заменит свои шифры, что она делала ранее девять или десять раз (вероятно стараниями эксперта Россиньоля), слава о его мастерстве так или иначе распространилась. Король Пруссии подарил ему золотую цепь за решение криптограммы, а парламент Бранденбурга вручил медаль за прочтение 200 или 300 листов шифра. Парламент Ганновера, не желая зависеть от иностранного криптоаналитика, заполучил интеллектуального собрата Валлиса, барона Готтфрида фон Лейбница, который назойливо докучал Валлису соблазнительными предложениями обучить нескольких молодых людей своему искусству. Когда Валлис проигнорировал вопрос Лейбница про то, как он делает такие удивительные вещи, говоря, что не было никакого определенного метода, Лейбниц быстро согласился с этим и, намекая, что искусство Валлиса может умереть вместе с ним, настаивал на своей просьбе, чтобы тот обучил нескольких молодых людей этому искусству. Валлис, наконец, вынужден был сказать прямо, что он будет рад работать для парламента в случае необходимости, но он не может передать свои приемы работы за границу без разрешения короля».

Гольдбах (1690–1764 гг.), бывший с 1742 г. тайным советником в российском Министерстве иностранных дел, дешифровал письмо французского посла с неприятными замечаниями о правящей царице Елизавете Петровне, дочери Петра Великого. Иногда случается фиаско, подобное произошедшему с опытным де Кохом, когда вскрытое письмо герцогу Моденскому по ошибке было запечатано печатью герцога Пармы. Тем не менее, черный кабинет австрийских императоров *Kaiserliche Geheime Kabinetts-Kanzlei* просуществовал до XIX столетия и был способен читать среди прочего корреспонденцию Наполеона и Талейрана.

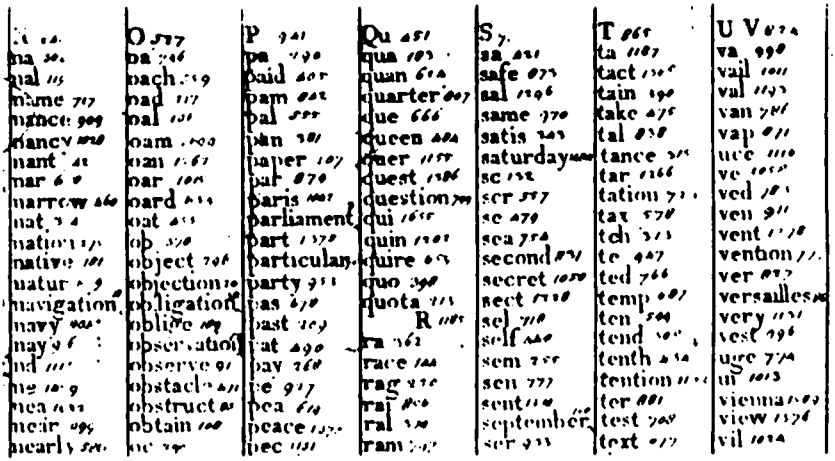


Рис. 36. Номенклатор 1785 г., сделанный Джефферсоном для использования Мэдисоном и Монро

В Новом Свете криптология также приобретала важное значение и известность. Джордж Вашингтон пользовался услугами двух агентов, Сэма Вудхалла и Роберта Таунсенда, скрывавшимися под псевдонимами Кульпер старший и Кульпер младший. В 1779 г. они использовали номенклатор приблизительно с 800 элементами, составленный майором Бенджамином Толлмаджем. Джефферсон также составил номенклатор для Джеймса Мэдисона в 1785 г. (рис. 36).

4.4.2. Книги с двумя частями. Приблизительно в 1630 г. Антуан Россиньоль (1600–1682 гг.) сделал большой шаг вперед в развитии номенклаторов. Кроме самых ранних примеров, до этого момента все номенклаторы имели вид отображений лексикографически упорядоченных элементов открытого текста на лексикографически упорядоченные кодовые группы из букв или математически упорядоченные цифровые кодовые группы. Это позволяло пользователю обходиться практически только одной кодовой книгой, которую можно было использовать для открытого текста так же, как и для шифротекста.

Эта система, однако, имела большой недостаток: если становился известным открытый текст, соответствующий одной кодовой группе, то сразу становилось понятным, что всем кодовым группам, стоящим в этой книге ниже, должны обязательно соответствовать открытые тексты, стоящие в кодовой книге также ниже. Поэтому по нескольким установленным парам открытый текст/кодовая группа может быть получена значительная информация обо всем коде.

Удивительно, но даже в течение Первой мировой войны эта устаревшая система одночастного номенклатора, одночастного кода (фр. *dictionnaire à table unique,* нем. *einteiliges Satzbuch)* была все еще в практическом употреблении в коде GREEN государственного департамента США, который имел до $20^3 \cdot 6^2 = 288\ 000$ кодовых групп типа $CVCVC$ (C заменяет согласную, V — гласную букву): FYTIG, MIHAK, PEDEK.

Код *Signalbuch* немецкого ВМФ времен Первой мировой войны, который был захвачен на крейсере *Магдебург* в августе 1914 г., был одночастным кодом (фрагмент кода показан ниже):

63940	OAT	Ohnmacht, -ig
63941	OAU	Ohr, Ohren-
63942	OAÜ	Okkupation, Okkupations, -ieren
63943	OAV	Okonomie, -isch
63944	OAW	Oktant.

Россиньоль ввел полностью перемешанные коды. Это означает, что для расшифровки необходима вторая часть, лексикографически или математически упорядоченный список кодовых групп, номенклатор с двумя частями, код с двумя частями (фр. *dictionnaire à table double, dictionnaire à deux tables,* нем. *zweiteiliges Satzbuch).* Примером является номенклатор Джефферсона. Ниже приведен современный пример с омофонами из военного жанра

⋮	⋮	⋮	⋮
flap	XYMAS	RATPA	ship
	TIBAL	RATPE	quite
flapjack	UPTON	RATPI	enough
flapper	UPABS	RATPO	happy
flare	OHPAP	RATPU	loxodromic
⋮	⋮	⋮	⋮

где для частого слова открытого текста наподобие /army/ имеется пять омофонов:

TORMA, RAFEM, LABAR, ROMUF, IBEXO.

В некоторых кодовых книгах пользуются цифровые группы (цифровые коды) наравне с кодовыми группами из букв. Например, фрагмент немецкого кода *Satzbuch* 1944 г. мог бы выглядеть так:

a	0809	XCL	b	1479	MLA
Abend	8435	PUV	Bad	1918	TID
aber	7463	NAS	bald	1492	LGD
acht	6397	DXL			
Achtung	1735	APS	z	2467	VBH
an	7958	EVG			
auf	6734	UNO	zyklotron	5116	JLD

Код *Kurzsignalheft* (краткая книга сигналов) немецкого ВМФ (с лета 1941 г.) был заголовочным кодом, содержащим кодовые группы для стереотипных команд:

AAAA *Beabsichtige gemeldete Feindstreitkrafte anzugreifen*
AAEE *Beabsichtige Durchfuhrung Unternehmung wie vorgeseheu*
AAFF *Beabsichtige Durchfuhrung Unternehmung mit vollem Einsatz*
AAGG *Beabsichtige Durchfuhrung Unternehmung unter Vermeidung vollen Einsatzes*

Код *Wetterkurzschlüssel* (краткий метеошифр) немецкого ВМФ кодировал воздушные температуры многозначным однобуквенным кодом (буква X отсутствовала!):

$$\begin{array}{cccccccc} A\hat{=}+28^\circ & B\hat{=}+27^\circ & C\hat{=}+26^\circ & D\hat{=}+25^\circ & \ldots & W\hat{=}+6^\circ & Y\hat{=}+5^\circ & Z\hat{=}+4^\circ \\ A\hat{=}+3^\circ & B\hat{=}+2^\circ & C\hat{=}+1^\circ & D\hat{=}0^\circ & E\hat{=}-1^\circ & F\hat{=}-2^\circ & \ldots & Z\hat{=}-21^\circ. \end{array}$$

Подобным же способом кодировались в определенном порядке температуры воды, атмосферное давлением, влажность, направление и скорость ветра, видимость, степень облачности, географические широта и долгота; сообщение о погоде состояло из единственного короткого слова. Это, казалось, было очень экономно и также делало трудным перехват сообщения, но криптологически это было совершенно глупо: перешифрованные метеосообщения, которые предписывалось регулярно передавать подводным лодкам, для криптоанализа противника были почти столь же хороши, как и открытый текст.

4.4.3. Современные коды. К 1700 г. номенклаторы имели 2000–3000 элементов и продолжали увеличиваться. Современные многозначные коды с омофонами показаны на рис. 37, 38.

В середине XIX века «черные кабинеты» в Европе были распущены (в Англии в 1844 г., в Вене и Париже в 1848 г.); закончилось the surreptitious — тайное вскрытие дипломатической и другой почты. Век Просвещения одержал свою победу. Индустриальная революция породила телеграф и, как следствие, коммерческие кодовые книги с их основной задачей сжатия телеграмм для снижения времени передачи.

В 1845 г. Фрэнсис О. Дж. Смит опубликовал код — *The Secret Corresponding Vocabulary. Adapted for Use to Morse's Electro-Mechanic Telegraph* (*Словарь секретной переписки, приспособленный к использованию с электромеханическим телеграфом Морзе*) — это произошло даже раньше, чем была введена

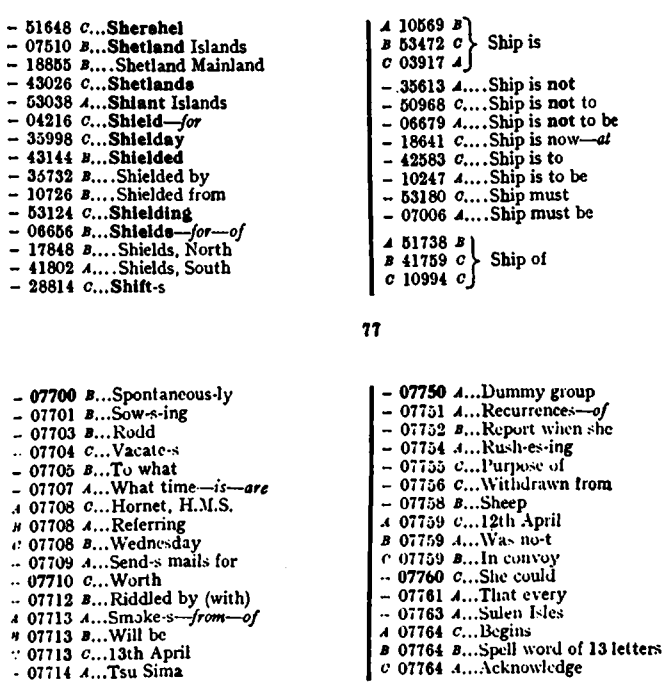


Рис. 37. Шифр SA британского Адмиралтейства (1918 г.). Одна страница части омофонного кодирования и одна страница части многозначого декодирования

азбука Морзе. Код Смита имел 50 000 кодовых групп, и лишь 67 предложений. Его кодовые группы строились из цифр (цифровой код) и предназначались для перешифрования (разд. 9.2). Позже стали использовать коды с кодовыми группами, построенными из букв (литеральный код), главным образом с группами из пяти букв; число предложений перешло на сотни, число кодовых групп выросло до 100 000. Из-за большого объема вновь стали использоваться одночастные коды. Это происходило даже в дипломатических службах и военных штабах, хотя там тайна была жизненно важна.

Мало-помалу появились сотни коммерческих кодов; среди самых первых были коды Генри Роджерса и Джона Виллса (оба в 1847 г.). Согласно Фридману, в 1860 г. «человек по имени Белл опубликовал в Буффало свой *Коммерческий шифр для сжатия телеграмм*». В 1874 г., спустя восемь лет после завершения прокладки трансатлантического кабеля, стал широко использоваться код ABC Уильяма Клозен–Туэ; родился пятисимвольный код. Другие пятисимвольные коды были созданы Болтоном (*Dictionnaire pour la Correspondance anglais*) Кроном в Берлине (1873 г.) и Уолтером в Винтертюре (1877 г.). Четырехсимвольный код (*Chiffrier-Wörterbuch*) был опубликован Катшером в Лейпциге (1889 г.), а трехсимвольный код (*Dictionnaire télégraphique, écono-*

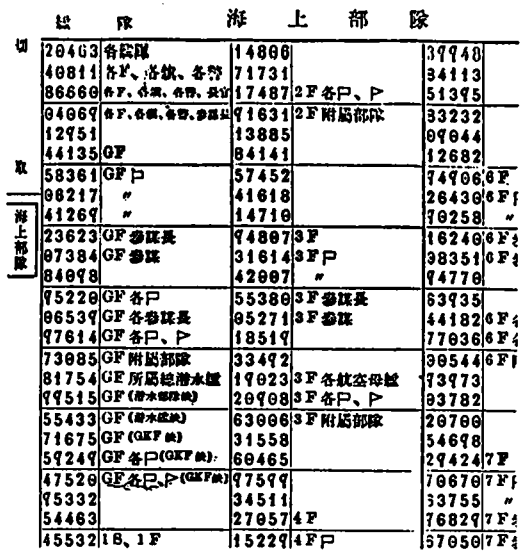


Рис. 38. Фрагмент кодирующей части японского Морского кода (1943 г.)

mique et secret) Маме-Гальяном в Париже (1874 г.). В США известные коды названы в честь Джона Чарльза Хартфилда (1877 г., развивался с 1890 г. его сыном Джоном Уильямом Хартфилдом) и Генри Харви (1878 г.). Кодовая книга Бенджамина Франклина Либера с 75 800 кодовыми группами была также переведена на французский и немецкий языки. Даже семисимвольные коды нашли применение, такие как *Ingemeur-Code* (в Германии) Галланда. Во всех кодах преследовалась, главным образом, цель уменьшить стоимость телеграфной передачи (*Коммерческий шифр для сжатия телеграмм* Белла, 1860 г.). Это было особенно важно для трансатлантического трафика.

В Европе предпочитали цифровые коды, которые допускали простое дополнительное перешифрование. Важным прототипом четырехразрядного кода стал *Dictionnaire abréviatif chiffré* Ф. Ж. Ситтлера в Париже (1868 г.), далее следует код *Dictionnaire pour la Correspondance télégraphique secrète* Брунсвика в Париже (1868 г.) и код *Dictionnaire chiffré* Нила; код Базерье (1893 г.), так же порождающий коды де Виари; другими четырехразрядными кодами были *Dizionario per corrispondenze in cifra* Баравелли в Турине (1896 г.), *Chiffrier-Wörterbuch* Фридмана в Берлине, и *Chiffrierbuch* Штейнера и Штерна в Вене (1892 г.).

4.4.4. Телеграфные коды. Тарифная политика Международного телеграфного союза (разд. 2.5.4) привела в 1890 г. к широко распространенному использованию пятиразрядных кодов. Брэчет (Brachet) опубликовал в Париже такой код в 1850 г. (*Dictionnaire chiffré*), другими кодами были *Diccionario para la correspondencia secreta* Вас Субтиля в Лиссабоне (1871 г.), *Wörterbuch* Низе в Берлине (1877 г.) и *Dictionnaire pour la Correspondance secrète* Н. К. Луи

в Париже (1881 г.). Позже появились коды *Dictionnaire chiffré Diplomatique et Commercial* Айренти и *Telescand Code* во Франции, *Diccionario Cryptographico* в Лиссабоне (1892 г.), *Nuovo Cifrario* Менгарини в Риме (1898 г.), *Cifrario per la corrispendenza segreta* Цицеро (Cicero) в Риме (1899 г.), *Slater's Code* Слейтера в Лондоне (1906 г.), и *Clave telegrafica* Дархана в Мадриде (1912 г.).

4.4.5. Коммерческие коды. В 20-м столетии во многих кодовых книгах используются как цифровые, так и буквенные кодовые группы. До недавнего времени часто использовались коды Bentley's *Code* (с 1922 г.), *ABC Code* (*6-е изд.*) (с 1925 г.), *Peterson's Code* (*3-е изд.*) Эрнста Ф. Петерсона, *Acme Code* Уильяма Дж. Митчела, *Rudolf Mosse Code* (с 1922 г.), *Lombard Code* и *AZ Code*. Самая большая кодовая книга, когда-либо бывшая в общем пользовании, создана Кирусом Тиббальсом для *Western Union Code;* она содержала 379 300 элементов, в то время как *ABC Code* имел только 103 000 элементов.

Во Второй мировой войне Союзники использовали код BAMS («Broadcasting for Allied Merchant Ships» — «Радиовещание для торговых судов союзников»), который был широко скомпрометирован как основа для перешифрования. Открытый текст был бы не худшим злом.

В течение долгих лет код *Internationaler Hotel-Telegrapheiischliissel für Zimmerbestellung* воспроизводился на немецких еженедельниках со следующими кодовыми группами:

ALBA для «1 *Zimmer mit* 1 *Bett*»,	ARAB для «1 *Zimmer mit* 2 *Betten*»,
ABEC для «1 *Zimmer mit* 3 *Betten*»,	BELAB для «2 *Zimmer mit je* 1 *Bett*»,
BIRAC для «2 *Zimmer mit* 3 *Betten*»,	BANAD для «2 *Zimmer mit* 4 *Betten*»,
CIROC для «3 *Zimmer mit* 3 *Betten*»,	CALDE для «3 *Zimmer mit* 5 *Betten*»,
CARID для «3 *Zimmer mit* 4 *Betten*»	и т. д.

Некоторые коды были переведены на иностранные языки. Код Маркони, переведенный Джеймсом К. Ч. Макбетом, является действительно многоязычным (девять языков в четырех томах), воплощая в реальности мечту Атаназиуса Кирхера (рис. 39).

4.4.6. Коды с обнаружением ошибок и коды, исправляющие ошибки. В 1880 г. Дж. Ч. Хартфилд в целях проверки ввел «двухсимвольный дифференциал» кодовых групп (в алфавите Z_{27} из 27 символов с $27^5 = 14\,348\,907$ кодовыми группами для пятисимвольного кода имеется $27^4 = 531\,441$ дифференциалов). Около 1925 г. У. Дж. Митчел ввел также проверку на наличие перестановок смежных символов (ведущих к изменениям наподобие в LABED и ALBED), которая уменьшила число пригодных для использования кодовых групп в примере до 440 051. Идея Митчела относительно «ограничения смежных символов» быстро распространилась. Оба кода японского ВМФ, взломанные американской службой OP-20-G, одночастный код JN-25a (1.06.1939; взломан в сент. 1940 г.) и двухчастный код JN-25b использовали кодовые группы из пяти цифр, делящиеся на 3. Эти коды были предшественниками кодов с обнаружением ошибок и кодов с исправлением ошибок, введенных Ричардом У. Хеммингом в 1950 г.; сегодня этот принцип провер-

	M.	N.	O.	P.	R.	S.	T.	U.	Y.	Z.
	0	1	2	3	4	5	6	7	8	9

19140	UVVIM	slackness.	relâchement.	flojedad, descuido.
19141	UVVON	Slag(s).	Scorie(s).	Escoria(s).
19142	UVWEO	Slander(s).	Diffame(r), diffamation.	Calumnia(r), calumnia(s).
19143	UVWUP	slandered.	diffamé.	calumniado.
19144	UVWYR	slandering.	diffamant.	calumniando.
19145	UVYDS	slanderous.	diffamatoire.	calumnioso.
19146	UVYCT	Slate(s).	Ardoise(s).	Pizarra(s).
19147	UVYDU	Sleeper(s).	Traverse(s) (chemins de fer).	Traviesa(s), durmiente(s) (f.c.).
19148	UVYFY	Sleeve-valve.	Soupape à manchon.	Válvula de manguito.
19149	UVYMZ	Slide(s).	Glisse(r), glissière(s). tiroir de distribution.	Resbala(r), corredera(s). válvula de distribución, de corredera.
19150	UVYUM	slide-valve.	glissant, à coulisse.	resbalando, resbaladizo, deslizante.
19151	UVYVN	sliding.		
19152	UVYWO	sliding scale.	échelle mobile.	escala móvil.
19153	UVYZP	Slight.	Léger, peu important.	Ligero, leve.
19154	UVZUR	slightest.	le (la) moindre.	lo más ligero, leve.
19155	UVZYS	not the slightest.	pas le (la) moindre.	no lo más ligero, mínimo.

	M.	N.	O.	P.	R.	S.	T.	U.	Y.	Z.
	0	1	2	3	4	5	6	7	8	9

19140	UVVIM	slackness.	Schlaffheit, Geschäftsstille.	slapheid, stilte.
19141	UVVON	Slag(s).	Schlacke(n).	Schuim, slak(ken), metaalschuim(-slakken).
19142	UVWEO	Slander(s).	Verleumden(-e,-t), Verleumdung(en).	Belasteren, belaster(s), laster.
19143	UVWUP	slandered.	verleumdet.	belasterd.
19144	UVWYR	slandering.	verleumdend.	belasterend.
19145	UVYDS	slanderous.	verleumderisch.	lasterlijk.
19146	UVYCT	Slate(s).	Schiefer, Schiefertafel(n).	Lei(en).
19147	UVYDU	Sleeper(s).	(Bahn-)Schwelle(n).	Dwarsligger(s).
19148	UVYFY	Sleeve-valve.	Muffenventil.	Mofklep.
19149	UVYMZ	Slide(s).	Gleiten(-e, -et), Gleitbahn(en), Gleitführung(en).	Glijden(-t); schuif (schuiven), leibaan (leibanen), windklep(pen).
19150	UVYUM	slide-valve.	Schieberventil, Ventilschieber.	schuif, stoomschuif, schuifklep.
19151	UVYVN	sliding.	gleitend.	verschuifbaar, glijdend.
19152	UVYWO	sliding scale.	Gleitskala.	kalibermaat, proportionele schaal.
19153	UVYZP	Slight.	Gering, von geringer Wichtigkeit.	Gering, onbeduidend.
19154	UVZUR	slightest.	geringst.	geringste.
19155	UVZYS	not the slightest.	nicht das (der, die) geringste.	niet de (het) geringste.

Рис. 39. Код Маркони: соответствующие страницы из англо-французско-испанского и англо-немецко-голландского изданий

ки присутствует в штрих-кодах европейского Кода изделий (European Article Number) или в системе ISBN.

4.4.7. Срок жизни кодов. В отличие от коммерческих кодовых книг, которые являются (при не очень низкой цене) общедоступными и поэтому должны иметь максимально возможный срок жизни, в дипломатических и военных кодах должно быть предусмотрено «запланированное устаревание» (разд. 2.1.1) и, соответственно, они должны заменяться как можно чаще. Поэтому представляется безнадежным перечислить все коды, использованные в 20-м столетии в этих областях, хотя часто скупость и лень служили помехами регулярным заменам кодов. Среди кодов, использовавшихся слишком долго, можно назвать такие дипломатические коды США, как пятизначные коды RED, BLUE (до 1914 г.) и GREEN (примерно 1914–1919 гг.). Около 1920 г. появился код GRAY, упомянутый в разд. 4.4.1; он использовался еще и при Франклине Делано Рузвельте в 1941 г. Рузвельт 6 декабря 1941 г. послал записку

Корделлу Халту: «*Дорогой Корделл — Отправьте это Грэю* [посол США в Токио] — *я думаю, подойдет серый код — сэкономит время, я не возражаю, если сообщение перехватят — ФДР*». При помешанном на секретности Рузвельте, в середине 1930-х годов был введен двухчастный код BROWN; после 1939 г., однако, подозрительный государственный деятель предпочел системы шифрования морского министерства «для вящей предельной секретности», как он выразился. Последующие дипломатические коды A-1, B-1, C-1, D-1 не изменяли жесткого мнения Рузвельта о слабой защищенности кодов Государственного департамента. Около 1940 г. появился код BLACK.

Иногда создатели кода сталкиваются с неожиданностями: когда в 1917 г. в Первую мировую войну *Американские экспедиционные силы* (A. E. F.) приняли участие в боевых действиях во Франции, оказалось, что *телеграфный код военного министерства* (*War Department Telegraph Code*), введенный в 1915 г., не обеспечивал безопасность и не подходил для тактического применения. В большой спешке *подотдел создания кодов* (*Code Compilation Subsection*) МИ-8 в июле 1917 г. начал работать над приемлемым кодом: 1 июля 1918 г., через год код был готов. *Военный секретный код № 5* (*Military Intelligence Code No. 5*) был кодом одночастным, хотя и с двусимвольной разностью кодовых групп типа *VCVCV*, *VCCVC* или *CVCCV*. Несмотря на то, что вскоре стал доступен лучший двухчастный код (*Military Intelligence Code* № 9), № 5 оставался в работе с уровнем «СЕКРЕТНО» до 1 сентября 1934 г., затем под кратким названием SIGCOT он был понижен до уровня «КОНФИДЕНЦИАЛЬНО». Аналогично, код № 9, который был выведен из употребления около 1923 г., был реанимирован 1 апреля 1933 г. с понижением до уровня «КОНФИДЕНЦИАЛЬНО» под кратким названием SIGSYG для шифрования и SIGPIK в части расшифрования. Причина этого лежала в недостатке денег, но даже перешифрование, которое не нуждается ни в каких инвестициях, не было предусмотрено.

4.4.8. Полевые коды. На нижнем армейском уровне коды имели иногда лучшую, иногда худшую репутацию. Немецкая армия перешла в 1917 г. от поворотных трафаретов (разд. 6.1.4) к кодам. Для радиопереговоров в 3-километровой боевой зоне в марте 1917 г. была введена простая двухсимвольная подстановка *Befehlstafel*. Уже в 1916 г. французы выпустили трехсимвольный код, *carnet réduit*, с названиями типа *olive* и *urbain*. В нем кодовые группы упорядочивались по заголовкам вроде пехоты, артиллерии, числам, часам времени, общим словам, названиям мест, названиям укрытий и т. д., так что это был заголовочный код.

В марте 1918 г. немцы стали использовать перешифрование — шаг, предсказанный Союзниками, но их код все еще оставался одночастным трехзначным кодом. Перешифрование распространялось только на первые две цифры и было реализовано в полном объеме лишь в коде *Geheimklappe* (разд. 4.1.2), который был часто изменяемым. Фиксированная третья цифра допускала появление шаблонов и, таким образом, помогала криптоанализу.

Для удовлетворения более высоких требований криптоаналитической защиты, вне 3-километровой боевой зоны, немцы ввели в июне 1917 г. трехсим-

вольный двухчастный код (*Satzbuch*). Перешифрование в нем не предполагалось; криптоаналитическая защита этой криптосистемы была изначально основана на запланированном устаревании (около 14 дней). Код содержал большое число омофонов (например, KXL, ROQ, UDZ для *Anschlußfehlt)* и пустышек. У Союзников он назывался кодом KRU, потому что все кодовые группы начинались с одной из букв K, R, U. Назывался он также кодом *Fritz.* Позже в нем добавились кодовые группы, начинающиеся с буквы S и, кроме того, кодовые группы, начинающиеся с буквы A (код KRUSA). Окончательно 26 символов алфавита были дополнены диакритическими гласными Ä, Ö, Ü (этот код также назывался — не очень систематически — кодом KRUSA).

Перемирие в ноябре 1918 г. закончило эту немирную эпоху «траншейных кодов». Но о полевых кодах не забывали. В руководстве американского Военного министерства (1944 г.) можно найти такой пассаж: «шифровальные машины не могут, как правило, выноситься из большего штаба, дивизии. Следовательно, кодовые методы могут преобладать в нижних эшелонах и войсковых подразделениях».

Это можно интерпретировать как понимание того, что в военном радиообмене кодирование без перешифрования допустимо только на самом низком уровне секретности. По этой причине Фридман вводил многоалфавитное шифрование, используя удобное устройство M-94 в качестве полевого шифратора армии США. Во время Второй мировой войны в США даже это уже не считали достаточным. Борис Хагелин, который в мае 1940 г. отправился в последнюю минуту из Швеции через Германию в Геную и затем приплыл на *Conte di Savoia* в Соединенные Штаты, везя в своем багаже два экземпляра

Рис. 40. Шифратор Хагелина M-209 в боевой обстановке

своей машины C-36 (разд. 8.5.1), произвел впечатление на Фридмана и американские войска связи ее улучшенной версией C-38. Хагелин должен был ждать в течение целого года, пока его машина не была полностью протестирована. В июне 1941 г. было принято решение относительно механической машины для подразделений нижнего уровня. На рис. 40 изображен солдат с винтовкой за спиной, работающий на шифровальной машине Хагелина M-209 в центре сообщений командного поста 3-й дивизии американской пехоты в Хипчонге (Корея) 1 октября 1951 г. Действительно, Борис Хагелин весьма рано предусмотрел возможность применения механических шифровальных машин на линии фронта. Для машины Хагелина C-35 была изготовлена опорная плита, позволяющая в полевых условиях размещать машину даже на колене оператора. Как пишет Хагелин, в случае необходимости оператор мог передвигаться с машиной, прикрепленной к его колену — если ему это нравится, в открытом положении. Для французских полицейских сил Хагелин в 1950-х годов сконструировал карманную шифровальную машинку примерно той же мощности (CD 55, CD 57).

Тем временем, борьба между кодовыми книгами и шифровальными машинами перестала быть актуальной — микроэлектроника устранила различия между ними и открыла полностью новые направления в шифровании. В основе одного из них лежит специальный случай линейных подстановок, которые мы обсудим в следующей главе.

Глава 5

Шаги шифрования: линейная подстановка

> Хотя сама по себе система шифрования Хилла на практике почти не использовалась, она оказала огромное влияние на криптологию.
>
> *Дэвид Кан*, 1967 *г.*

Линейная («аффинная») подстановка является специфической многосимвольной подстановкой. Инъективный шаг шифрования многосимвольного блочного шифрования всегда можно рассматривать как отображение

$$\chi\colon V^n \to W^m$$

при достаточно больших n и m. Конечные алфавиты V и W здесь мы будем считать, *как правило*, линейно упорядоченными от первого символа $\alpha(V)$ до последнего символа $\omega(V)$. Упорядоченный алфавит называется стандартным алфавитом в соответствующем смысле.

В этом упорядочивании каждый символ x, кроме последнего, имеет однозначно определенный *следующий* символ $\operatorname{succ} x$; для последнего символа следующим символом будем считать первый символ: $\operatorname{succ} \omega(V) = \alpha(V)$. Таким образом, отображение succ однозначно определяет обратное отображение pred; циклически замкнутый стандартный алфавит обладает конечным неветвящимся (т. е. линейным) циклическим квазипорядком.

При $V = Z_{|V|}$ и $W = Z_{|W|}$, кроме того, можно рекурсивно определить сложение: для $a, b \in V$ или $a, b \in W$ определим

$$a + b = \operatorname{succ} a + \operatorname{pred} b,$$
$$a + \alpha = a.$$

Это означает, что множества $Z_{|V|}$ и $Z_{|W|}$ отображаются однозначно с сохранением порядка на $\mathbb{Z}_{|V|}$ и $\mathbb{Z}_{|W|}$, где $\mathbb{Z}_N$ обозначает группу классов вычетов по

модулю натурального числа N, элементы которой можно представить циклом натуральных чисел $\{0, 1, \ldots, N-1\}$. Сложение в V и W соответствует сложению классов вычетов (в соответствующих группах). Обычно алфавит $\{\alpha, \ldots, \omega\}$ отождествляется с числами цикла («циклотомическими числами») $\{0, 1, \ldots, N-1\}$, где $N = |V|$ или $N = |W|$[1)].

Сложение в V или в W мы теперь можем покомпонентно перенести на V^n и W^m. Кроме того, мы отождествляем $V = W = \mathbb{Z}_N$, $V^n = \mathbb{Z}_N^n$, $W^m = \mathbb{Z}_N^m$.

В соответствие с этими определениями, отображение $\varphi: \mathbb{Z}_N^n \to \mathbb{Z}_N^m$ называется аддитивным тогда и только тогда, когда

$$\forall x, y \in \mathbb{Z}_N^n: \varphi(x+y) = \varphi(x) + \varphi(y),$$

иными словами: «образ суммы равен сумме образов». Следовательно,

$$\forall x \in \mathbb{Z}_N^n: \varphi(x + x + \ldots + x) = \varphi(x) + \varphi(x) + \ldots + \varphi(x),$$

т. е. «образ кратного элемента есть кратное образа этого элемента». Действительно, $\mathbb{Z}_N$ является кольцом, а $\mathbb{Z}_N^n$ — векторным пространством с началом координат $\mathbf{o} = (00 \ldots 0)$; φ является линейным отображением векторного пространства $\mathbb{Z}_N^n$ в векторное пространство $\mathbb{Z}_N^m$. Когда $N = p$ является простым числом (и только в этом случае!), $\mathbb{Z}_N$ является полем Галуа $\mathbb{F}(p)$. Однако в дальнейшем мы не будем требовать простоты числа N.

Для представления отображения мы будем использовать квадратную матрицу T с элементами из $\mathbb{Z}_N$; для φ: $\varphi(x) = xT$, а для обратного отображения $\varphi^{-1}(y) = yT^{-1}$.

Линейная подстановка χ: $\mathbb{Z}_N^n \to \mathbb{Z}_N^m$ определяется как сумма однородной части, линейного отображения φ, представляемого матрицей $T \in \mathbb{Z}_N^{n,m}$, и сдвига начала координат, представляемого вектором $t \in \mathbb{Z}_N^m$:

$$\chi(x) = xT + t.$$

Если T — единичная матрица, то имеет место специальный случай, когда подстановка является чистым сдвигом $\chi(x) = x + t$ (многосимвольное сложение ЦЕЗАРЯ).

Если линейное отображение φ инъективно, то оно регулярно, т. е. имеет единственное обратное отображение φ^{-1}, определенное на образе отображения φ. Если мы имеем дело с эндоморфным случаем ($m = n$), то регулярное линейное отображение является взаимно однозначным.

Пример. Определенные на Z_{26} квадратная 3×3 матрица T и трехкомпонентный вектор t

$$T = \begin{pmatrix} 15 & 2 & 7 \\ 8 & 10 & 23 \\ 0 & 2 & 8 \end{pmatrix}, \qquad t = \begin{pmatrix} 17 & 4 & 20 \end{pmatrix}$$

[1)] Для $Z_{26} \leftrightarrow \mathbb{Z}_{26}$, отождествление производится следующим образом («алгебраический алфавит»):

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25.

определяют трехсимвольную замену триграмм. Образ триграммы /mai/ $\hat{=}$ $\hat{=}$ (12 0 8) получается из следующих вычислений по *модулю* 26:

$$(12 \quad 0 \quad 8)\begin{pmatrix}15 & 2 & 7\\ 8 & 10 & 23\\ 0 & 2 & 8\end{pmatrix} + (17 \quad 4 \quad 20) \overset{26}{\simeq}$$

$$\overset{26}{\simeq} (24 \quad 14 \quad 18) + (17 \quad 4 \quad 20) \overset{26}{\simeq} (15 \quad 18 \quad 12) \hat{=} /\mathrm{psm}/.$$

Но триграмма /ecg/ $\hat{=}$ (4 2 6) имеет тот же самый образ /psm/, так как

$$(4 \quad 2 \quad 6)\begin{pmatrix}15 & 2 & 7\\ 8 & 10 & 23\\ 0 & 2 & 8\end{pmatrix} \overset{26}{\simeq} (24 \quad 14 \quad 18).$$

Поэтому шифрование при помощи данной матрицы T не является инъективным: на самом деле матрица T не регулярна и не имеет обратной. Матрица T отображает вектор (8 24 2) в вектор (0 0 0).

5.1. Взаимообратные линейные подстановки

Возникает естественный вопрос: когда эндоморфная линейная подстановка является взаимообратной? Очевидно, в том случае, когда для $\forall x$ $\chi(x) = xT + t = \chi^{-1}(x)$. Таким образом

$$x = \chi(\chi(x)) = (xT + t)T + t = xT^2 + tT + t,$$

откуда следует, что

$$T^2 = I \quad \text{и} \quad tT + t = \mathbf{o}.$$

Это означает, что матрица T однородной части φ является взаимообратной, если она имеет в качестве собственных значений только 1 или $N-1$, а вектор сдвига t или равен нулю, или является собственным вектором T для собственного значения $N-1$. В частности, χ может задаваться равенством

$$\chi(x) = x + (1 - \gamma(x))v \quad \text{с} \quad v \neq \mathbf{o},$$

в котором линейный функционал γ удовлетворяет условию $\gamma(v) = 2$ (в случае $N = 2$ условию $\gamma(v) = \mathbf{o}$). Тогда $\chi(v) = \mathbf{o}$ и $\chi(\mathbf{o}) = v$. Простое вычисление подтверждает тождество $\chi^2(x) = x$:

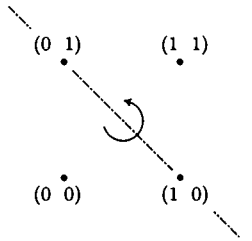
$$\begin{aligned}\chi(\chi(x)) &= \chi(x) + [1 - \gamma(\chi(x))]v = \\ &= x + (1 - \gamma(x))v + [1 - \gamma(x) - (1 - \gamma(x))\gamma(v)]v = \\ &= x + (1 - \gamma(x))v + [(1 - \gamma(x)) - 2(1 - \gamma(x))]v = x.\end{aligned}$$

Пример. $\mathbb{Z}_2^2 \to \mathbb{Z}_2^2$ $(N = 2,\ n = 2)$

$$\chi\big((x_1 \quad x_2)\big) = (x_1 \quad x_2) + (1 - x_1 - x_2)\ (1 \quad 1) = (1 - x_2 \quad 1 - x_1) =$$

$$= (x_1 \quad x_2)\begin{pmatrix}0 & -1\\ -1 & 0\end{pmatrix} + (1 \quad 1).$$

$$\chi((0\quad 0)) = (1\quad 1),$$
$$\chi((0\quad 1)) = (0\quad 1),$$
$$\chi((1\quad 0)) = (1\quad 0),$$
$$\chi((1\quad 1)) = (0\quad 0).$$



Обычный текст

0010011101101100 01
1110010001100011 01

порождает криптотекст и наоборот.

5.2. Однородные линейные подстановки

5.2.1. Хилл. Специальный случай однородной линейной замены ($t = 0$) в качестве инструмента шифрования изучался Хиллом (шаг шифрования ХИЛЛА). Как получить одновременно матрицу T и ей обратную T^{-1}? Рассмотрим пример (квадратной) матрицы над кольцом $\mathbb{Z}$ с определителем, равным $+1$ при $n = 4$:

$$T = \begin{pmatrix} 8 & 6 & 9 & 5 \\ 6 & 9 & 5 & 10 \\ 5 & 8 & 4 & 9 \\ 10 & 6 & 11 & 4 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} -3 & 20 & -21 & 1 \\ 2 & -41 & 44 & 1 \\ 2 & -6 & 6 & -1 \\ -1 & 28 & -30 & -1 \end{pmatrix}.$$

При вычислениях удобно использовать в качестве представителей классов вычетов маленькие отрицательные числа; поэтому в Z_{26} обратную матрицу

$$T^{-1} \overset{26}{\simeq} \begin{pmatrix} 23 & 20 & 5 & 1 \\ 2 & 11 & 18 & 1 \\ 2 & 20 & 6 & 25 \\ 25 & 2 & 22 & 25 \end{pmatrix}$$

можно представить в виде

$$T^{-1} \overset{26}{\simeq} \begin{pmatrix} -3 & -6 & 5 & 1 \\ 2 & 11 & -8 & 1 \\ 2 & -6 & 6 & -1 \\ -1 & 2 & -4 & -1 \end{pmatrix}.$$

Пример. Образом тетраграммы /ende/=(4 13 3 4), преобразованной матрицей T, является тетраграмма /jhbl/=(9 7 1 11):

$$(4 \quad 13 \quad 3 \quad 4) \begin{pmatrix} 8 & 6 & 9 & 5 \\ 6 & 9 & 5 & 10 \\ 5 & 8 & 4 & 9 \\ 10 & 6 & 11 & 4 \end{pmatrix} \overset{26}{\simeq} (9 \quad 7 \quad 1 \quad 11),$$

$$\begin{pmatrix}9 & 7 & 1 & 11\end{pmatrix}\begin{pmatrix}23 & 20 & 5 & 1\\ 2 & 11 & 18 & 1\\ 2 & 20 & 6 & 25\\ 25 & 2 & 22 & 25\end{pmatrix}\overset{26}{\simeq}\begin{pmatrix}4 & 13 & 3 & 4\end{pmatrix}.$$

5.2.2. Неоднородный случай. При $t=(3\ 8\ 5\ 20)$ и матрице T, приведенной выше, получается неоднородная линейная подстановка χ:

$$\chi\big((x_1\ \ x_2\ \ x_3\ \ x_4)\big)=(x_1\ \ x_2\ \ x_3\ \ x_4)\begin{pmatrix}8 & 6 & 9 & 5\\ 6 & 9 & 5 & 10\\ 5 & 8 & 4 & 9\\ 10 & 6 & 11 & 4\end{pmatrix}+(3\ \ 8\ \ 5\ \ 20)$$

с обратной подстановкой

$$\chi^{-1}\big((y_1\ \ y_2\ \ y_3\ \ y_4)\big)=(y_1\ \ y_2\ \ y_3\ \ y_4)\begin{pmatrix}23 & 20 & 5 & 1\\ 2 & 11 & 18 & 1\\ 2 & 20 & 6 & 25\\ 25 & 2 & 22 & 25\end{pmatrix}+(3\ \ 24\ \ 21\ \ 14).$$

5.2.3. Перечисление. Количество регулярных $n\times n$ матриц над $\mathbb{Z}_N$ зависит от простоты числа N. Известный результат (см. Л. Е. Диксон, Линейные группы. Лейпциг, 1901) звучит так:

Теорема. *Пусть $N=p$ — простое число. Количество $g(p,n)$ регулярных матриц из $\mathbb{Z}_p^{n,n}$ равно числу базисов векторного пространства $\mathbb{Z}_p^{n,n}$, а именно:*

$$g(p,n)=(p^n-1)(p^n-p)(p^n-p^2)\ \dots\ (p^n-p^{n-1}).$$

Количество всех различных $n\times n$ матриц равно p^{n^2}. Таким образом

$$g(p,n)=p^{n^2}\cdot\rho(p,n),\quad\text{где}$$
$$\rho(p,n)=\prod_{k=1}^{n}\Big(1-\Big(\frac{1}{p}\Big)^k\Big).$$

Для двоичного случая $N=2$ имеем: $g(2,1)=1$, $g(2,2)=2^1\cdot 3$, $g(2,3)=2^3\cdot 3\cdot 7$, $g(2,4)=2^6\cdot 3\cdot 7\cdot 15$, $g(2,5)=2^{10}\cdot 3\cdot 7\cdot 15\cdot 31$, $g(2,6)=2^{15}\cdot 3\cdot 7\cdot 15\cdot 31\cdot 63$.

Можно вычислить предел $\rho(p,n)$ при n, стремящемся к бесконечности (Эйлер, 1760 г.):

$$\lim_{n\to\infty}\rho(p,n)=h\Big(\frac{1}{p}\Big),\quad\text{где}$$

$$h(x)=1+\sum_{k=1}^{\infty}(-1)^k\big[x^{(3k^2-k)/2}+x^{(3k^2+k)/2}\big]=$$
$$=1-x-x^2+x^5+x^7-x^{12}-x^{15}+x^{22}+x^{26}+\dots$$

Функция $h(x)$ представляет собой «лакунарный» ряд, связанный с тетарядами и эллиптическими функциями. За деталями можно обратиться к книге. Р. Ремерта «Теория функций», I, (изд. Springer, Берлин, 1984, стр. 263). Функция $h(1/p)$ при больших n дает довольно хорошую оценку для $\rho(p, n)$; ниже в таблице приводятся значения функций для некоторых простых p:

$N=p$	$h\left(\frac{1}{p}\right)$	$\ln h\left(\frac{1}{p}\right)$
2	0.28879	−1.24206
3	0.56013	−0.57959
5	0.76033	−0.27400
7	0.83680	−0.17817
11	0.90083	−0.10444
13	0.91716	−0.08647
17	0.93772	−0.06430
19	0.94460	−0.05699

Для $p > 10$ первые три члена ряда $1 - 1/p - 1/p^2$ дают уже пять правильных цифр для $h(1/p)$; $1/(3/2 - p)$ аппроксимирует $\ln h(1/p)$ с относительной ошибкой меньшей, чем $1/p^2$.

Если размер алфавита N является степенью простого числа, то ситуация становится более сложной.

Теорема (Манфред Брой, 1981). *Пусть $N = p^s$ и $A \in \mathbb{Z}_N^{n,n}$. Тогда существуют матрицы $A_i \in \mathbb{Z}_N^{n,n}$, $0 \leqslant i < s$ такие, что A можно однозначно представить в виде суммы $A = \sum_{i=0}^{s-1} A_i p^i$. Матрица A является регулярной тогда и только тогда, когда регулярна матрица A_0.*

Из этой теоремы получаем

$$g(p^s, n) = g(p, n) \cdot (p^{s-1})^{n^2} = (p^s)^{n^2} \cdot \rho(p, n) = N^{n^2} \cdot \rho(p, n).$$

Наконец, для общего случая $N = p_1^{s_1} \cdot p_2^{s_2} \ldots p_k^{s_k}$ количество регулярных $n \times n$ матриц равно

$$g(N, n) = N^{n^2} \cdot \rho(p_1, n) \cdot \rho(p_2, n) \ldots \rho(p_k, n).$$

Число N^{n^2} меньше числа $(N^n)!$ всех подстановок n-символьных n-кортежей: для $N = 25$, $n = 4$ мы имеем $(N^n)! \approx 10^{2\,184\,284}$, тогда как $N^{n^2} = 2.33 \cdot 10^{22}$ и $g(N, n) = 1.77 \cdot 10^{22}$. Оно становится близким к числу $6.20 \cdot 10^{23}$ простых циклических перестановок для $N = 25$.

Для $N = 25$ имеем:

$$g(25, 1) = 20, \quad g(25, 2) = 300\,000, \quad g(25, 3) = 2\,906\,250\,000\,000;$$

Для $N = 26$ имеем:

$$g(26, 1) = 12, \quad g(26, 2) = 157\,248, \quad g(26, 3) = 1\,634\,038\,189\,056.$$

5.2.4. Построение пар матриц. Строить регулярную квадратную матрицу проще всего в виде произведения нижне- и верхнетреугольной регулярных матриц. Это означает, что диагональные элементы матрицы должны быть обратимыми в соответствующем кольце (разд. 5.5, табл. 1); проще всего выбрать их равными 1. Кроме того, верхнетреугольная матрица может быть выбрана как результат транспонирования нижнетреугольной матрицы, что порождает симметричную матрицу. В итоге обращение треугольных матриц методом исключения приведет нас к обратной матрице.

Выбрав единицы на диагонали, что также возможно, хотя не всегда предпочтительно, проделаем все вычисления сначала в $\mathbb{Z}$ и затем перейдем к классам вычетов.

Пример.

$$\begin{pmatrix}1&&\\3&1&\\5&2&1\end{pmatrix}\begin{pmatrix}1&3&5\\&1&2\\&&1\end{pmatrix}=\begin{pmatrix}1&3&5\\3&10&17\\5&17&30\end{pmatrix};$$

$$\begin{pmatrix}1&&\\3&1&\\5&2&1\end{pmatrix}^{-1}=\begin{pmatrix}1&&\\-3&1&\\1&-2&1\end{pmatrix},\quad\begin{pmatrix}1&3&5\\&1&2\\&&1\end{pmatrix}^{-1}=\begin{pmatrix}1&-3&1\\&1&-2\\&&1\end{pmatrix},$$

$$\begin{pmatrix}1&-3&1\\&1&-2\\&&1\end{pmatrix}\begin{pmatrix}1&&\\-3&1&\\1&-2&1\end{pmatrix}=\begin{pmatrix}11&-5&1\\-5&5&-2\\1&-2&1\end{pmatrix}.$$

Переходя к $\mathbb{Z}_{26}$ и $\mathbb{Z}_{25}$, имеем, соответственно, пары (симметричных) взаимно обратных матриц

$$\begin{pmatrix}1&3&5\\3&10&17\\5&17&4\end{pmatrix},\ \begin{pmatrix}11&21&1\\21&5&24\\1&24&1\end{pmatrix}\quad\text{и}\quad\begin{pmatrix}1&3&5\\3&10&17\\5&17&5\end{pmatrix},\ \begin{pmatrix}11&20&1\\20&5&23\\1&23&1\end{pmatrix},$$

при переходе к $\mathbb{Z}_{10}$, $\mathbb{Z}_2$ эти же пары выглядят так:

$$\begin{pmatrix}1&3&5\\3&0&7\\5&7&0\end{pmatrix},\ \begin{pmatrix}1&5&1\\5&5&8\\1&8&1\end{pmatrix}\quad\text{и}\quad\begin{pmatrix}1&1&1\\1&0&1\\1&1&0\end{pmatrix},\ \begin{pmatrix}1&1&1\\1&1&0\\1&0&1\end{pmatrix}.$$

Если для данных n и N выбирать произвольные нижнетреугольные матрицы L, верхнетреугольные матрицы U (с единицами на диагонали) и диагональные матрицы с обратимыми элементами D, то пары матриц LDU и $U^{-1}D^{-1}L^{-1}$ дадут нам все пары взаимообратных матриц с точностью до порядка строк и столбцов.

5.2.5. Построение взаимообратной матрицы едва ли более трудно. Если для данных n и N $(X,\ X^{-1})$ — пара взаимно обратных матриц и J является взаимообратной диагональной матрицей с элементами $+1$ или -1, то XJX^{-1} также является взаимообратной матрицей.

Пример.

$$\begin{pmatrix}1&3&5\\3&10&17\\5&17&30\end{pmatrix}\begin{pmatrix}1&&\\&-1&\\&&1\end{pmatrix}\begin{pmatrix}11&-5&1\\-5&5&-2\\1&-2&1\end{pmatrix}=\begin{pmatrix}31&-30&12\\100&-99&40\\170&-170&69\end{pmatrix}.$$

При переходе к $\mathbb{Z}_{26}$, $\mathbb{Z}_{25}$, $\mathbb{Z}_{10}$, $\mathbb{Z}_2$ получаются следующие взаимообратные матрицы:

$$\begin{pmatrix}5&22&12\\22&5&14\\14&12&17\end{pmatrix},\quad\begin{pmatrix}6&20&12\\0&1&15\\20&5&19\end{pmatrix},\quad\begin{pmatrix}1&0&2\\0&1&0\\0&0&9\end{pmatrix},\quad\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

Над $\mathbb{Z}_2$ единичная матрица является единственной взаимообратной диагональной матрицей. Не известно никаких простых выражений для числа взаимообратных $n\times n$ матриц над $\mathbb{Z}_N$.

5.3. Двоичные линейные подстановки

Для $\mathbb{Z}_2$, т. е. для двоичных слов (длины n_0) открытого текста и криптотекста, техническая реализация линейных подстановок особенно проста. Арифметика по *модулю* 2 может быть реализована параллельно двоичной электрической схемой с шиной ширины n_0 (для не очень больших значений n_0, скажем до 64). Сравнивая пространство $\mathbb{Z}_{2^s}^{n_0\times n_0}$ $(n=n_0,\ N=2^s)$ с пространством $\mathbb{Z}_2^{s\cdot n_0\times s\cdot n_0}$ $(n=s\cdot n_0,\ N=2)$ с помощью разложения 2^s символов $\mathbb{Z}_{2^s}^{n_0\times n_0}$ в двоичные слова длины s, получаем следующее заключение: число всех регулярных линейных подстановок равно $2^{s\cdot n_0^2}s\cdot\rho(2,\,n_0)=K\cdot\rho(2,\,n_0)$ в первом случае и $2^{s^2\cdot n_0^2}s\cdot\rho(2,\,s\cdot n_0)=K^S\cdot\rho(2,\,s\cdot n_0)$ — во втором случае.

5.4. Общие линейные подстановки

Включая N^n сдвигов, всего имеется N^{n^2+n} линейных подстановок. Взаимообратные однородные линейные подстановки (с $N=26$) были предложены Лестером С. Хиллом[2)] в 1929 г. (его предшественником был Ф. Дж. Бакк в 1772 г.; в 1867 г. Л. Ж. д'Ариоль использовал двухсимвольный шифр биграмм $V^2\to V^2$, который, возможно, был специальной линейной подстановкой). Идеи Хилла были возрождены в 1941 г. А. А. Альбертом на волне патриотико-математического энтузиазма. К этому времени идеи Хилла уже оказали свое воздействие на У. Ф. Фридмана в США и В. Кюнце из немецкой *Auswärtiges*

[2)] Лестер С. Хилл был доцентом математики в колледже Хантера в Нью-Йорке. Степень доктора философии он получил в 1926 г. в Йеле в возрасте 35 лет, будучи некоторое время учителем в школе. Статья была опубликована в *The American Mathematical Monthly* под заголовком «Криптография в алгебраическом алфавите» («*Cryptography in an Algebraic Alphabet*», vol. 37, p. 135–154, March 1931). Хилл получил патент США 1845947 на свой аппарат 16 февраля 1932 г. До 1960 г. он был профессором колледжа Хантера и умер 9 января 1961 г.

Amt[3]. Важность изобретения Хилла состоит в том, что в настоящее время ценность математических методов в криптологии стала бесспорной. Как следствие в 1930-х годах математики пошли в шифровальные бюро: Соломон Кульбак (1907–1994 гг.), Абрахам Синков (1907–1998 гг.), Вернер Кюнце, голландский статистик Мауриц де Вриз и многие другие, чьи имена остаются нераскрытыми.

Лестер С. Хилл разработал машину для линейных подстановок ($n = 6$, патент США 1 845 947). Это простое механическое устройство со сцепленными колесами было весьма медленным, поэтому машины Хилла во время Второй мировой войны использовались лишь для перешифрования трехсимвольных кодовых групп радиопозывных, что по сравнению с ручным вычислением давало значительную экономию времени.

5.5. Декомпозиция линейных подстановок

Имеется дополнительный специальный случай, когда линейная подстановка содержит некоторое многоалфавитное шифрование. Это происходит, если T разлагается в прямую сумму: $T = T_1 \oplus T_2 \oplus \ldots \oplus T_r$, т. е. ее матрица имеет блочно-диагональный вид

$$\left(\begin{array}{c|c|c|c} T_1 & 0 & \ldots & 0 \\ \hline 0 & T_2 & \ldots & 0 \\ \hline \vdots & \vdots & \ldots & 0 \\ \hline 0 & 0 & 0 & T_r \end{array}\right),$$

где матрица T_i имеет размерность $n_i \times n_i$. В этом случае, каждое T_i вместе с соответствующей частью t_i, $t = t_1 \oplus t_2 \oplus \ldots \oplus t_r$, по-прежнему остается многосимвольной подстановкой, шифрующей n_i-граммы. Если эти r подстановок попарно различны, шаг шифрования становится r-кратным многоалфавитным линейным многосимвольным шагом. Другими словами, весь период периодического многоалфавитного шифрования содержится в одной матрице. Подробнее об этом мы поговорим в разд. 7.4.1.

Чрезвычайно важным является случай $n_i = 1$, $r = n$. В этом случае T является диагональной матрицей и каждая строка в ней соответствует простой линейной подстановке, односимвольной подстановке 1-грамм T: $V^1 \to V^1$.

Изучим эту подстановку, являющуюся перестановкой, более подробно. Она задается равенством $\chi(x) = h \cdot x + t$ и, конечно, регулярна при $h = 1$, приобре-

[3]Д-р Вернер Кюнце (р. примерно в 1890 г.) изучал математику, физику и философию в Гейдельберге, служил в кавалерии в Первую мировую войну и в январе 1918 г. начал заниматься криптологией в *Auswärtiges Amt*. В 1923 г. он взломал перешифрованный французский дипломатический код, в 1936 г. — ORANGE и позже RED, две японских роторные шифровальные машины. Кюнце был, возможно, первым профессиональным математиком, обслуживающим современное криптоаналитическое бюро. Кюнце был, подобно Майборну, неплохим скрипачем, а Оливер Стречи был известен как хороший музыкант, в то время как Пейнвин был превосходным виолончелистом. Ламброуз Д. Каллимахес из АНБ был известным флейтистом.

тая в этом случае вид: $\chi(x) = x + t$. Таким образом, простая линейная подстановка с $h = 1$ является односимвольным сложением ЦЕЗАРЯ $\chi(x) = x + t$ с обратной подстановкой $\chi^{-1}(x) = x - t$ (при подходящем $t \neq 0$).

Для $\chi(x) = x + i$ с этого момента мы будем писать также $\chi(x) = \rho^i(x)$, где $\rho(x) = x + 1$. Имеет место более общее утверждение:

Простая линейная подстановка $\chi(x) = h \cdot x + t$ регулярна и $\chi^{-1}(x) = h^{-1} \times$ $\times (x - t)$ тогда и только тогда, когда h взаимно просто с N.

В таблицах 1 а и 1 б для некоторых значений N приведены взаимообратные пары h и h^{-1}, включая некоторые взаимообратные h, приводящие к инволютивным перестановкам.

5.6. Алфавиты с выбыванием

Однородный случай $t = 0$ включает тривиальные случаи:

$$h = h^{-1} = 1 \qquad \text{(неизменяемый алфавит) и}$$

$$h = h^{-1} = N - 1 \qquad \text{(дополнительный алфавит),}$$

иначе говоря, алфавиты с выбыванием (фр. *alphabets chevauchants*, нем. *dezimierte Alphabete*), изучавшиеся еще Эйраудом: алфавиты, которые заполняются h-кратными целых чисел по *модулю* N (предполагается взаимная простота h и N). Таким образом, алфавиты получаются прохождением с шагом h («символьное умножение», «выбивание каждого h-го») исходного алфавита.

Примеры для $N = 8$:

$$h = 1: \quad \begin{matrix} a & b & c & d & e & f & g & h \\ a & b & c & d & e & f & g & h \end{matrix} = (a)\,(b)\,(c)\,(d)\,(e)\,(f)\,(g)\,(h),$$

$$h = 7: \quad \begin{matrix} a & b & c & d & e & f & g & h \\ a & h & g & f & e & f & c & b \end{matrix} = (a)\,(bh)\,(cg)\,(df)\,(e),$$

$$h = 3: \quad \begin{matrix} a & b & c & d & e & f & g & h \\ a & d & g & b & e & h & c & f \end{matrix} = (a)\,(bd)\,(cg)\,(e)\,(fh),$$

$$h = 5: \quad \begin{matrix} a & b & c & d & e & f & g & h \\ a & f & c & h & e & b & g & d \end{matrix} = (a)\,(bf)\,(c)\,(dh)\,(e)\,(g).$$

Есть различие между дополнительным алфавитом с

$$\chi(x) = (N - 1) \cdot x \overset{N}{\simeq} N - x \overset{N}{\simeq} -x$$

и обращенным алфавитом, полученным из неоднородного случая с

$$\chi(x) = (N - 1) \cdot (x + 1) \overset{N}{\simeq} (N - 1) - x \overset{N}{\simeq} -x - 1.$$

Число $g(N, 1)$ *регулярных* однородных простых линейных подстановок совпадает с функцией Эйлера $\varphi(N)$, равной количеству чисел из набора $1, \ldots, N - 1$, взаимно простых с N.

$N=2$	1									$\mathcal{M}_2=\mathcal{C}_1$
$N=3$	1	**2**								$\mathcal{M}_3=\mathcal{C}_2$
$N=4$	1	**3**								$\mathcal{M}_4=\mathcal{C}_2$
$N=5$	1	**2** 3	4							$\mathcal{M}_5=\mathcal{C}_4$
$N=6$	1	**5**								$\mathcal{M}_5=\mathcal{C}_2$
$N=7$	1	2 4	**3** 5	6						$\mathcal{M}_7=\mathcal{C}_6$
$N=8$	1	**3**	5	**7**						$\mathcal{M}_8=\mathcal{C}_2\times\mathcal{C}_2$
$N=9$	1	**2** 5	4 7	8						$\mathcal{M}_9=\mathcal{C}_6$
$N=10$	1	**3** 7	9							$\mathcal{M}_{10}=\mathcal{C}_4$
$N=11$	1	**2** 6	3 4	5 9	7 8	10				$\mathcal{M}_{11}=\mathcal{C}_{10}$
$N=12$	1	**5**	7	**11**						$\mathcal{M}_{12}=\mathcal{C}_2\times\mathcal{C}_2$
$N=13$	1	**2** 7	3 9	4 10	5 8	6 11	12			$\mathcal{M}_{13}=\mathcal{C}_{12}$
$N=14$	1	**3** 5	9 11	13						$\mathcal{M}_{14}=\mathcal{C}_6$
$N=15$	1	**2** 8	4	7 13	11	**14**				$\mathcal{M}_{15}=\mathcal{C}_4\times\mathcal{C}_2$
$N=16$	1	**3** 11	5 13	7	9	**15**				$\mathcal{M}_{16}=\mathcal{C}_4\times\mathcal{C}_2$
$N=17$	1	2 9	**3** 6	4 13	5 7	8 15	10 12	11 14	12	$\mathcal{M}_{17}=\mathcal{C}_{16}$

Таблица 1 а. Взаимно обратные пары в $\mathbb{Z}_N$ для N от 2 до 17 (жирным выделены порождающие элементы мультипликативной группы $\mathcal{M}_N$). Символ C_k обозначает циклическую группу порядка k

$N=18$	1	**5** 11	7 13	17												$\mathcal{M}_{18}=C_6$
$N=19$	1	**2** 10	3 13	4 5	6 16	7 11	8 12	9 17	14 15	18						$\mathcal{M}_{19}=C_{18}$
$N=20$	1	**3** 7	9	11	13 17	**19**										$\mathcal{M}_{20}=C_4\times C_2$
$N=21$	1	**2** 11	4 16	5 17	8	10 19	13	**20**								$\mathcal{M}_{21}=C_6\times C_2$
$N=22$	1	3 15	5 9	**7** 19	13 17	21										$\mathcal{M}_{22}=C_{10}$
$N=23$	1	2 12	3 8	4 6	**5** 14	7 10	9 18	11 21	13 16	15 20	17 19	22				$\mathcal{M}_{23}=C_{22}$
$N=24$	1	**5**	7	11	**13**	17	19	**23**								$\mathcal{M}_{24}=C_2\times C_2\times C_2$
$N=25$	1	**2** 13	3 17	4 19	6 21	7 18	8 22	9 14	11 16	12 23	24					$\mathcal{M}_{25}=C_{20}$
$N=26$	1	3 9	5 21	**7** 15	11 19	17 23	25									$\mathcal{M}_{26}=C_{12}$
$N=27$	1	**2** 14	4 7	5 11	8 17	10 19	13 25	16 22	20 23	26						$\mathcal{M}_{27}=C_{18}$
$N=28$	1	**3** 19	5 17	9 25	11 23	13	15	**27**								$\mathcal{M}_{28}=C_6\times C_2$
$N=29$	1	**2** 15	3 10	4 22	5 6	7 25	8 11	9 13	12 17	14 27	16 20	18 21	19 26	23 24	28	$\mathcal{M}_{29}=C_{28}$
$N=30$	1	**7** 13	11	17 23	19	**29**										$\mathcal{M}_{30}=C_4\times C_2$
$N=31$	1	2 16	**3** 21	4 8	5 25	6 26	7 9	10 28	11 17	12 13	14 20	15 29	18 19	22 24	23 27	30 $\mathcal{M}_{31}=C_{30}$
$N=32$	1	**3** 11	5 13	7 23	9 25	15	17	19 27	21 29	**31**						$\mathcal{M}_{32}=C_8\times C_2$
$N=33$	1	**2** 17	4 25	5 20	7 19	8 29	10	13 28	14 26	16 31	23	**32**				$\mathcal{M}_{33}=C_{10}\times C_2$

Таблица 1 b. Взаимно обратные пары в $\mathbb{Z}_N$ для N от 18 до 33 (жирным выделены порождающие элементы мультипликативной группы $\mathcal{M}_N$). Символ C_k обозначает циклическую группу порядка k

Для $N = p_1^{s_1} \cdot p_2^{s_2} \cdot \ldots \cdot p_k^{s_k}$ имеют место равенства

$$\varphi(N) = (p_1 - 1) \cdot p^{s_1 - 1} \cdot (p_2 - 1) \cdot p^{s_2 - 1} \ \ldots \ (p_k - 1) \cdot p^{s_k - 1} = \\ = N \cdot \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \ldots \left(1 - \frac{1}{p_s}\right).$$

5.7. Линейные подстановки с десятичными и двоичными числами

Обратим внимание на следующее различие: $\mathbb{Z}_N^n$ принадлежит V^n, в то время как $\mathbb{Z}_{N^n}$ принадлежит V; $\mathbb{Z}_{N^n}$ применяется, если V^n лексикографически упорядочено.

5.7.1. Случай $N = 10$ ($\mathbb{Z}_{10^n}$). Алфавиты с выбыванием (разд. 5.6) особенно интересны для любителей, зашифровывающих n-значные десятичные числа с помощью калькулятора, тем что они позволяют использовать помимо сложения еще и умножение.

Вычисления в $\mathbb{Z}_{10^n}$ легко производятся также с помощью механического сумматора. (Для перехода к вычислению в $\mathbb{Z}_{10}^n$ необходимо только отключить устройство переноса, см. разд. 8.3.3.)

Пример $n = 2$ ($\mathbb{Z}_{100}$): достаточно знать обратные по *модулю* 100 для простых чисел до 97 (исключая 2 и 5):

h =	3	7	11	13	17	19	23	29	31	37	41	43	47	53	59	61	67	71	73	79	83	89	97
h^{-1} =	67	43	91	77	53	79	87	69	71	73	61	7	83	17	39	41	3	31	27	19	47	9	33

Заметим, что последние цифры обратных величин сами являются обратными по *модулю* 10. Это наблюдение подсказывает пошаговую процедуру для больших значений n: на каждом шаге выбирается соответствующим образом только одна новая цифра.

Пример $n = 5$ ($\mathbb{Z}_{10^5}$): обратное по *модулю* 10^5 числа $h = 32413$ есть $h^{-1} = 3477$, что подтверждается следующими вычислениями:

$$\begin{array}{rlll}
3: & & & 3 \cdot \quad 7 = 2 \cdot 10 + 1 \\
13: & 2 + 1 \cdot 7 + 3 \cdot x \overset{10}{\simeq} 0 & x = 7 & 13 \cdot \quad 77 = 10 \cdot 10^2 + 1 \\
413: & 10 + 4 \cdot 7 + 3 \cdot x \overset{10}{\simeq} 0 & x = 4 & 413 \cdot \quad 477 = 197 \cdot 10^3 + 1 \\
2413: & 197 + 2 \cdot 7 + 3 \cdot x \overset{10}{\simeq} 0 & x = 3 & 2413 \cdot \quad 3477 = 839 \cdot 10^4 + 1 \\
32413: & 839 + 3 \cdot 7 + 3 \cdot x \overset{10}{\simeq} 0 & x = 0 & 32413 \cdot 03477 = 1127 \cdot 10^5 + 1
\end{array}$$

Число операций, достаточное для определения обратных n-значных чисел, пропорционально n^2.

5.7.2. Случай $N = 2$ ($\mathbb{Z}_2^n$). Для профессиональной работы предпочтительна двоичная система счисления. Случаи $n = 8$, 16, 32 и даже 64 соответствуют внутренней арифметической архитектуре микропроцессоров. Алгоритм для определения обратного по *модулю* 2^n полностью аналогичен рассмотренному

выше десятичному алгоритму, более того, он, подобно классическому алгоритму деления двоичных чисел проще, чем алгоритм для десятичных чисел.

Например, число $LOOO\ OOOO\ OOLL\ OLLL = 32\,823$ является обратным по *модулю* 2^{16} числу $OOLL\ OLOL\ LOOO\ OLLL = 13703$. Это следует из равенства $32823 \cdot 13703 = 449773569 = 1 + 6863 \cdot 2^{16}$.

5.7.3. Тьюринг. Уже осенью 1937 г., за два года до того, как он серьезно увлекся криптологией, Алан Тьюринг размышлял о шифровании умножением в двоичной системе счисления. Это, возможно, приходило в голову также и другим математикам. Тьюринг, однако, разработал схему из реле для умножения и при участии физика Малкольма МакФэйла (Принстон) построил несколько экземпляров. Обстоятельства побудили Тьюринга отложить этот проект после его возвращения в июле 1938 г. из Принстона. Несмотря на это он был хорошо подготовлен к работам по механическому криптоанализу 4 сентября 1939 г., спустя всего один день после начала войны, когда он прибыл в Блетчли Парк. Это место представляло собой викторианский особняк в Букингемшире, на полпути между Оксфордом и Кембриджем. Туда в августе 1939 г. была эвакуирована Правительственная школа кодов и шифров. Летом 1938 г., после своего возвращения, Тьюринг был приглашен в эту школу прочитать курс криптологии[4]. Так же он достаточно регулярно встречался с главным криптологом Дильвином Кноксом, который изо всех сил и с большими трудностями пытался вскрывать итальянские и, позже, испанские сообщения, зашифрованные коммерческой ЭНИГМОЙ без коммутационной панели[5] (и который умер в 1943 г.). Гордон Уэлчман также был завербован в разведывательную работу до того, как вспыхнула война.

Первым математиком, завербованным Правительственной школой кодов и шифров был Питер Твинн, оксфордский выпускник, который поступил на службу в феврале 1939 г. Ему говорили позже, что имелись некоторые сомнения относительно правильности привлечения математиков, «поскольку они расценивались как странные коллеги, печально известные непрактичностью» (Эндрю). Фактически, некоторые другие ранние обитатели Блетчли Парка подобно Алану Тьюрингу, Гордону Уэлчману и Деннису Бэббиджу имели по крайней мере некоторые навыки в шахматах, не говоря уже о шахматных мастерах Стюарте Милнере-Барри, Гарри Голомбке и Хью Александере, привлеченных с помощью Уэлчмана.

Великобритания была хорошо готова к назревавшей войне; лучшие люди из Оксфорда и Кембриджа, если они не хотели стать летчиками-истребителя-

[4] Другой курс Тьюринг прочитал в рождественские дни.

[5] Возможно, не стоит доверять истории, начатой Фредериком В. Винтербосем и рассказанной в книге Кейва Брауна, что Кнокс и Тьюринг ездили в середине 1938 г. в Варшаву на встречу, устроенную польской секретной службой. Поляк с псевдонимом Рихард Левинский, который предположительно работал в берлинской фирме «Heimsoeth & Rincke» в качестве математика и инженера, предложил доставить копию ЭНИГМы. Марианн Режевски в 1982 г. назвал это «басней». Однако Гарри Хинслей сообщает, что уже в 1938 г. польская секретная служба вошла в контакт с G. C. & C. S. и Кноксом по поводу ЭНИГМы. Этот первый контакт, однако, не был успешным; Кнокс назвал поляка «глупым и неосведомленным».

ми, работали в Блетчли Парк. ПШКШ — подразделение Министерства иностранных дел, с середины 1938 г. была мобилизована. Ни Соединенные Штаты, ни Германия не привлекали ученых. Таким образом, в Великобритании потребовалось больше времени по сравнению с Германией или с США, чтобы признать важность математики для криптоанализа, но Тьюринг и Уэлчман, оказавшиеся под рукой, были полностью готовы к предстоящей работе. Использование талантов нетрадиционных и эксцентричных лиц позволила Министерству иностранных дел создать самую способную команду криптоаналитиков в британской истории.

Глава 6

Шаги шифрования: перестановки

En un mot, les methodes de transposition
sont une salade des lettres du texte clair.

[Короче говоря, методы перестановок дают хорошее перемешивание букв открытого текста.]

Базерье.

В частном случае, не рассмотренном в гл. 5, требуется, чтобы матрица однородной линейной подстановки состояла только из нулей и единиц. При $N > 2$ это является серьезным ограничением. Дополнительное требование, чтобы в каждой строке и в каждом столбце единица встречалась только один раз и, таким образом, остальные элементы были бы нулевыми, приводит к матрице перестановок, производящей простую перестановку элементов базиса векторного пространства.

Перестановка[1)] (нем. *Würfelverfahren* или *Versatzverfahren*) представляет собой многосимвольную подстановку $V^n \to V^n$, *кодирование* специального вида

$$(x_1, x_2, \ldots, x_n) \mapsto (x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}),$$

где $\pi \in \gamma_k$ является перестановкой индексов $\{1 \ldots n\}$; здесь γ_k обозначает группу всех $n!$ перестановок.

Перестановка является не перестановкой символов алфавита, а перестановкой мест, которые они занимают в слове. Она изначально присутствует в анаграммах (*апельсин-спаниель*), используемых, в частности, при конструировании псевдонимов (*Améry-Mayer*).

6.1. Простейшие методы

Простые методы используют один или несколько шагов шифрования, выполняемых друг за другом над последовательностями не очень большой длины n («перестановка составных элементов»).

[1)] В оригинале — transposition.

6.1.1. Оборачивание. Наиболее простой перестановкой является обращение порядка или оборачивание (нем. *Krebs*): в словах сообщения буквы переставляются в обратном порядке или же в обратном порядке переписывается все сообщение: МОКНУСИР С УГИНК ИРТОМС (см. разд. 1.5).

Эта «перевернутая запись» включает анонимы, вроде РЕМАРК для КРАМЕР и АВЕ для ЕВА. Оборачивание известно также в музыке, например в обратном каноне.

Палиндромы — это слова или предложения, которые остаются неизменными при оборачивании:

Madam été	Reittier summus
Able was I ere I saw Elba	Ein Neger mit Gazelle zagt im Regen nie
Esope reste ici et se repose	in girum irnus nocte et consumimur igni
А роза упала на лапу Азора	Меня истина манит сияньем.

В каждом языке имеются свои палиндромы. Вот несколько примеров на русском языке (*добавлено переводчиком*):

Я иду с мечем, судия. (*Г. Державин*)

Мокнет Оксана с котенком. (*А. Земель*)

Ах, у лешего на ноге шелуха.

Да, небо дано, но надобен ад. (*Д. Авалиани*)

У девы в уме роется: «А я-с теорему выведу...» (*Б. Левин*)

Еще несколько примеров по-английски:

Red rum & murder.

A man, a plan, a canal: Panama.

Ma is as selfless as I am.

Was it a cat I saw? (*Генри Э. Дадни*)

Madam, I'm Adam. (*Сэм Ллойд*)

Lewd did I live, & evil I did dwel. (*Джон Тейлор*)

Draw pupil's lip upward.

Doc note, I dissent; a fast never prevents a fatness; I diet on cod. (*Питер Хилтон*).

6.1.2. Перестановка слогов. Безобидное некриптографическое использование перестановки слогов (нем. *Schüttelreim*) (с $n = 4$) обнаруживается в следующих примерах:

прав я, юродивый — рою я правдивый

ах, реклама — река хлама

а бог чей? — богачей

нарву ревень, навру — не верь
they hung Bags — they flung hags
dear old Queen — queer old Dean
wasted the term — tasted the worm
missed the history — hissed the mystery

Перестановка π: $\pi(1, 2, 3, 4) = (3, 2, 1, 4)$, обнаруживаемая в перестановках слогов, использовалась криптографически в медицинском греческом, увлечение которым в среде лондонских студентов-медиков, согласно Кану, приобрело масштабы умеренной эпидемии: например, POKE A SMIPE заменяло *smoke a pipe*.

6.1.3. Маршрутная перестановка. Следующим методом является маршрутная перестановка, или маршрутизация («блуждание», нем. *Wiirfel*): открытый текст записывается в l строк фиксированной длины k и считывается по ходу движения по некоторому предписанному маршруту. Таким образом, прямоугольник размера $k \times l$ используется для шага шифрования с $n = k \times l$. Часто используется квадрат, тогда $n = k^2$. Для квадрата 2×2 перестановка слогов включается как «самопересекающийся» маршрут.

Например, шифротекст может быть считан по столбцам (маршрутизация «строка–столбец»):

Я	Д	О	К	Т
О	Р	Э	Й	З
Е	Н	Х	А	Р
Т	Б	А	Р	Т

ЯОЕТДРНБОЭХАКЙАРТЗРТ

Этот метод уже нам встречался при формировании алфавита по мнемоническому паролю. Возможны варианты считывания по диагоналям:

ТЕБОНАЯРХРДЭАТОЙРКЗТ,

или бистрофедонически[2)], двигаясь по столбцам попеременно вниз и вверх (каждый второй столбец проходится в обратном порядке):

ЯОЕТБНРДОЭХАРТАЙКТЗРТ,

или даже по спирали:

ТРЗТКОДЯОЕТБАРАЙЭРНХ.

Более сложный маршрут считывания задается движением коня (нем. *Rösselsprungwürfel*) (рис. 41); если стартовая точка известна или, хотя бы, может быть предположена, расшифрование текста не очень трудно для человека, знакомого с задачей об обходе конем (шахматной) доски.

[2)] Греч. *bustropheidon*, нем. *fürchenwendig*, маршрут передвижения по полю волов при вспашке.

1	4	53	18	55	6	43	20
52	17	2	5	38	19	56	7
3	64	15	54	31	42	21	44
16	51	28	39	34	37	8	57
63	14	35	32	41	30	45	22
50	27	40	29	36	33	58	9
13	62	25	48	11	60	23	46
26	49	12	61	24	47	10	59

Рис. 41. Маршрут для перемещения коня ($n = 64$)

Время от времени вместо прямоугольников использовались другие геометрические фигуры, чаще всего треугольники, а также пересечения различных форм и другие массивы (рис. 42, 43). Здесь нет никаких пределов для фантазии. Но эти простые методы перестановок легки для криптоанализа.

a e i n r v z
b d f h k m o q s u w y
c g l p t x
A E I N R V Z B D F H K M O Q S U W Y C G L P T X

Рис. 42. Перестановка «Штакетник» (Смит), $n = 25$

b f k o s w
a c e g i l n p r t v x
d h m q u z
B F K O S W A C E G I L N P R T V X D H M Q U Z

Рис. 43. Перестановка «Греческий орнамент» (Мюллер), «Роза ветров» (Николс), $n = 24$

6.1.4. Решетки. Более удобными средствами для выполнения перестановок и более безопасными (если шаблон достаточно нерегулярен) по сравнению с методами маршрутизации являются решетки, называемые также сеточными шифрами (фр. *grille*, нем. *Raster*). Обычно сразу заготавливается целый набор готовых решеток. Важное практическое усовершенствование заключается в поворачивающейся решетке, которая обеспечивает различные положения окон одной и той же решетки при ее повороте. Такая решетка была описана в 1885 г. Жюль Верном (1828–1905 гг.) в повести «*Матиас Шандор*» (*Mathias Sandorff*). Решетки использовались в XVIII в., например в 1745 г. в правление голландского *Стадтхоудера* Вильяма IV.

Математик К. Ф. Гинденбург изучал поворачивающиеся решетки в 1796 г., его изыскания продолжили Мориц фон Прассе в 1799 г. и Иоган Людвиг Клю-

бер в 1809 г. Используются решетки как с двумя положениями (рис. 44) так и с четырьмя положениями (рис. 45). Последние часто называются решетками Флейсснера[3)].

Рис. 44. Поворачивающаяся решетка с двумя положениями

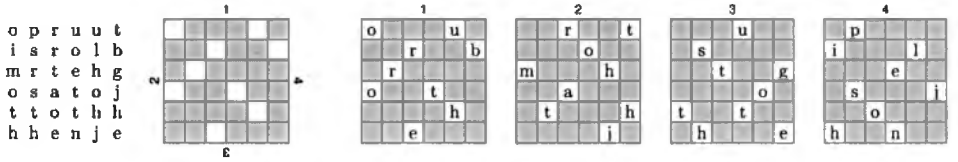


Рис. 45. Поворачивающаяся решетка с четырьмя положениями

Конструирование поворачивающихся решеток достаточно просто: клетки квадранта квадратной шахматной доски (с *четным числом* 2ν строк и столбцов) нумеруются числами от $1 \ldots \nu^2$. Эти номера путем последовательных поворотов решетки проставляются в остальных квадрантах. Затем для каждого числа выбирается позиция вращаемой решетки и вырезается соответствующее окно. Поворачивающаяся решетка на рис. 45 получена по этому алгоритму:

1	2	3	7	**4**	1
4	5	**6**	8	5	**2**
7	**8**	9	9	6	3
3	6	9	**9**	8	7
2	5	8	6	**6**	4.
1	4	**7**	3	2	1

Эта методика позволяет изготавливать поворачивающиеся решетки согласно комбинации ключей, которая перечисляет позиции решетки в последовательных шагах ее построения, например, следующим способом:

[3)] Эдуард Флейсснер фон Востровиц, австрийский полковник. Имеется в виду его книга «Neue Patronen-Geheimschrift» (Справочник по военной криптографии, Вена, 1881 г.). Термин «Patrone» означает «родительская форма», эталонная форма; он использовался в текстильной промышленности для рисунка ткацкого *образца* на разграфленной в клетку кальке. В романе Ярослава Гашека *Бравый солдат Швейк* упоминается «Справочник по военной криптографии» обер-лейтенанта фон Флейсснера; имеются и другие детали, указывающие на определенное знакомство Гашека с криптографией.

Нетрудно подсчитать мощность класса поворачивающихся решеток (для данного $n = 4\nu^2$): для решетки с двумя позициями имеется $2^{n/2}$ вариантов, для решетки с четырьмя позициями — $4^{n/4} = 2^{n/2}$ вариантов. Для $n = 36$ имеется приблизительно $2.62 \cdot 10^5$ решеток Флейсснера, число же всех перестановок равно $36! \approx 3.72 \cdot 10^{41}$.

Даже в конце XIX в. в армиях некоторых стран применяли шифры с использованием маршрутизации и поворачивающихся решеток в качестве полевых шифров. В Первую мировую войну в немецком армейском шифре перед 1917 г. периодически вводились новые поворачивающиеся решетки с именами типа *АННА* (5×5), *БЕРТА* (6×6), *КЛАРА* (7×7), *ДОРА* (8×8), *ЭМИЛЬ* (9×9) и *ФРАНЦ* (10×10). После четырех месяцев эта череда решеток оборвалась — к сожалению французов, которые легко взламывали этот шифр.

Маршрутизация и поворачивающиеся решетки приводят к перестановке букв внутри сообщения и ничему более. Их предшественник, решетка Кардано, не порождает *никаких* перестановок, вместо этого она помимо букв, видимых через ее окошки, добавляет большое количество пустышек (разд. 1.6). Решетки любого вида вообще не должны использоваться серьезными криптографами, но уровень их безопасности вполне достаточен для любителей. С другой стороны, к перестановкам в сочетании с подстановками следует отнестись весьма серьезно.

6.2. Перестановки столбцов

Серьезное использование перестановок возможно только при весьма больших значениях n, приближающихся к длине всего сообщения, что приводит к использованию паролей для выбора перестановок из очень большого множества шагов шифрования. По немецкой традиции такой пароль иногда называется *лозунгом* (*Losung*).

6.2.1. Пароли. Они уже используются в простой столбцовой перестановке (фр. *transposition simple â clef*): открытый текст записывается в строках выбранной длины k, получившиеся столбцы переупорядочиваются в соответствии с перестановкой $\pi \in \gamma_k$ (*лозунгом*) и криптотекст считывается *столбец за столбцом*:

п и с ь м о н е п о л у ч е н о $\quad \pi$: *2 1 4 3*

2 1 4 3	*1 2 3 4*	
п и с ь	И П Ь С	
м о н е	О М Е Н	И О О Е П М П Ч Ь Е У О С Н Л Н.
п о л у	О П У Л	
ч е н о	Е Ч О Н	

Эквивалентной с точки зрения криптоанализа является, очевидно, блочная перестановка или «перестановка составного модуля» (Гейнс), по-французски *variante de Richelieu* (Эйрауд), по-немецки *Gruppen-Transposition*, *Umstellung*, которая производится так же, как и столбцовая, за исключением того, что

криптотекст считывается *строка за строкой*:

п и с ь м о н е п о л у ч е н о π: *2 1 4 3*

2 1 4 3	*1 2 3 4*
п и с ь	И П Ь С
м о н е	О М Е Н
п о л у	О П У Л
ч е н о	Е Ч О Н

И П Ь С О М Е Н О П У Л Е Ч О Н.

Такое шифрование может также интерпретироваться следующим образом: открытый текст делится на блоки по k элементов и каждый блок переставляется в соответствии с перестановкой π, что означает многократный одноалфавитный многосимвольный шаг шифрования *блоков* ширины k: блочная перестановка — это перестановка составного модуля. По сравнению с блочной перестановкой простая столбцовая перестановка выполняется с карандашом на бумаге легче и с меньшим риском ошибиться. Хотя она распространяется на весь открытый текст, обеспечиваемая ею защита не больше, чем у перестановки составного модуля — перед нами пример *иллюзорной сложности*.

6.2.2. Прямоугольные схемы. Для маршрутизации и простой столбцовой перестановки расшифрование облегчается, если открытый текст полностью заполняет прямоугольник или квадрат из l строк. Чтобы достигнуть этого, часто используются пустышки. Однако выбирать их надо с большой осторожностью, иначе несанкционированное дешифрование может оказаться простым, например, если открытый текст дополняется пустышками q q ... q q. Ни в коем случае нельзя идти на подобное заполнение; длина последней строки определяется как остаток деления длины текста на l.

Простая столбцовая и блочная перестановки (даже с неполными прямоугольниками) могут быть легко взломаны. Поэтому следует обратить внимание на более сложные (составные) методы перестановок (см. разд. 9.1.1).

6.2.3. Двухшаговые методы. Дополнительная перестановка вводится в строчно-перемешанную столбцовую перестановку (фр. *transposition double*, Живарж, Эйрауд): открытый текст записывается в строках выбранной длины k в *соответствии с некоторой перестановкой* π_1, полученные столбцы переупорядочиваются в *соответствии с перестановкой* π_2, и криптотекст считывается столбец за столбцом:

п и с ь м о н е п о л у ч е н о π_1: *2 4 1 3* π_2: *2 1 4 3*

2 1 4 3		*1 2 3 4*	
1	п и с ь	*2*	м о н е	О М Е Н
2	м о н е	*4*	ч е н о	Е Ч О Н
3	п о л у	*1*	п и с ь	И П Ь С
4	ч е н о	*3*	п о л у	О П У Л

О Е И О М Ч П П Е О Ь У Н Н С Л.

Аналогично выполняется строчно-перемешанная блочная перестановка, в которой криптотекст считывается *строка за строкой*:

$$\underbrace{\text{письмонеполучено}} \qquad \pi_1\colon 2\,4\,1\,3 \quad \pi_2\colon 2\,1\,4\,3$$

2143		1234		
1	пись	2	моне	ОМЕН
2	моне	4	чено	ЕЧОН
3	полу	1	пись	ИПЬС
4	чено	3	полу	ОПУЛ

$\Big\}$ ОМЕНЕЧОНИПЬСОПУЛ.

Тот же самый эффект можно получить, если выполнять перестановку π_1 в конце процедуры:

$$\underbrace{\text{письмонеполучено}} \qquad \pi_1\colon 2\,4\,1\,3 \quad \pi_2\colon 2\,1\,4\,3$$

2143	1234			
пись	ИПЬС	1	ОМЕН	2
моне	ОМЕН	2	ЕЧОН	4
полу	ОПУЛ	3	ИПЬС	1
чено	ЕЧОН	4	ОПУЛ	3

$\Big\}$ ОМЕНЕЧОНИПЬСОПУЛ.

Если число строк равно числу столбцов, то в строчно-перемешанной столбцовой или блочной перестановке можно использовать одну и ту же перестановку: $\pi_1 = \pi_2$. Беря $\pi_2 = \pi_1^{-1}$, получим метод, который Керкхоффс в 1883 г. два раза приписывал русским *нигилистам* («перестановка Нигилиста»).

Рассмотрим эти перестановки, полагая, что открытый текст и криптотекст заполняют прямоугольную (или квадратную) матрицу X с l строками и с k столбцами.

В этом случае перестановка строк матрицы эквивалентна умножению этой матрицы слева на $l \times l$ матрицу перестановки π_1:

$$X \mapsto \pi_1 X,$$

а перестановка столбцов эквивалентна умножению справа на $k \times k$ матрицу перестановки π_2:

$$X \mapsto X\pi_2.$$

Строчно-столбцовая транскрипция означает *матричное* транспонирование (отражение относительно диагонали):

$$X \mapsto X^T.$$

Таким образом, блочная перестановка является преобразованием $X \mapsto$ $\mapsto X\pi_2$, тогда как столбцовая перестановка описывается перестановкой столбцов, и согласно строчно-столбцовой транскрипции есть

$$X \mapsto (X\pi_2)^T.$$

Следовательно (так как матрица перестановки ортогональна, то $\pi_2^T = \pi_2^{-1}$), перестановка строк описывается перестановкой π_2^{-1} строк транспонированной матрицы:

$$X \mapsto \pi_2^{-1} X^T.$$

Поэтому строчно-перемешанная блочная перестановка является преобразованием

$$X \mapsto (\pi_1 X)\pi_2 = \pi_1(X\pi_2),$$

тогда как строчно-перемешанная столбцовая перестановка — это

$$X \mapsto ((\pi_1 X)\pi_2)^T,$$

или в другой форме:

$$X \mapsto \pi_2^{-1} X^T \pi_1^{-1}.$$

6.2.4. Ubchi. Двойная столбцовая перестановка (фр. *double transposition*[4)], нем. *doppelte Spaltentransposition*, *Doppelwürfelverfahren*) использует простую столбцовую перестановку дважды. Это, в общем случае, предполагает использование двух различных паролей, что для квадратных матриц делается не всегда.

Двойную столбцовую перестановку можно интерпретировать как отображение

$$X \mapsto ((X\pi)^T\pi')^T = (\pi')^{-1} X\pi,$$

которое является инвариантным относительно строчно-перемешанной блочной перестановки. Эти отображения выполняются с помощью произведения перестановок $\pi_i \times \pi_k$. Для перестановки Нигилиста рассматриваемое отображение является преобразованием подобия, то же самое верно для шифра двойной перестановки армии США в случае $l = k$ (см. ниже). Все эти методы взламываются по существу теми же самыми криптоаналитическими средствами, что и в случае простой столбцовой перестановки.

Двойная столбцовая перестановка с одним паролем долгое время использовалась в армии США («шифр двойной перестановки армии США»). Она также использовалась в немецком армейском шифре, который французские криптоаналитики под командой майора, позже полковника и генерала Франсуа Картье называли *ubchi*, — этот шифр использовался на довоенных немецких маневрах, на которых сообщения помечались меткой *übchi*, сокращением слова *übungschiffrierung* (учебная стрельба). Французы научились взламывать этот шифр и в начале войны читали уже боевые сообщения вплоть до 18 ноября 1914 г.

Удивительно, но немецкий *Вермахт* не извлек уроков из этой истории и повторил свои ошибки: с начала Второй мировой войны до 1 июля 1941 г. и снова с 1 июня 1942 г. двойная столбцовая перестановка с паролем, который менялся каждый день, применялась в армейском аварийном шифре

[4)] Отметим различие французских терминов: *transposition double*= дополнительная перестановка (разд. 6.2.3) и *double transposition* = двойная перестановка (разд. 6.2.4).

(*Handschlüsselverfahren*), используемом на уровне полка и ниже, и в *Kriegsmarine* (шифр *Reserve-Handverfahren*, *Notschlüssel*). В этот раз его читали англичане. Заметим, что в данном случае не помогло бы и использование двух различных паролей и даже тройной столбцовой перестановки — подобные шифры вскрываются методом «кратных анаграмм». Подробнее криптоанализ столбцовой перестановки мы рассмотрим в гл. 21.

Двойная столбцовая перестановка (с одним паролем) была основной криптографической системой голландского Сопротивления и французских маки. Она также использовалась британскими шпионами и диверсионным *Управлением специальных операций* (*Special Operations Executive* — S. O. E), созданным *Черчиллем* в 1942 г., что было подтверждено в 1998 г. *Лео Марксом*, прежним главой отдела кодирования S. O. E.

Дешифрование простых перестановок усложняется после введения в открытый текст случайно размещенных пробелов (японская система 1941 г. PA-K2; немецкий армейский шифр *Rasterschlüssel 44*, введенный в марте 1944 г.). Несмотря на это, США обычно вскрывали шифр PA-K2, хотя часто и со значительной временной задержкой. Шифр *Rasterschlüssel 44* не стал бы в этом плане исключением, не появись он так поздно.

6.2.5. Конструирование перестановок. Можно предложить много схем для порождения перестановок, которые используются в столбцовых перестановках из мнемонического пароля. Один часто упоминаемый в литературе прием выполняется следующим образом.

Каждой букве пароля сопоставляется его номер в упорядоченном алфавите:

M	*A*	*C*	*B*	*E*	*T*	*H*
6	1	3	2	4	7	5

Это достаточно просто, если пароль не имеет повторяющихся символов. Если же таковые имеются, немного модифицированная схема упорядочивает повторные символы последовательно:

A	*M*	*B*	*A*	*S*	*S*	*A*	*D*	*E*	*D*	*A*	*L*	*L*	*E*	*M*	*A*	*G*	*N*	*E*
1	15	6	2	18	19	3	7	9	8	4	13	14	10	16	5	12	17	11.

6.3. Анаграммы

Перестановка оставляет неизменным состав[5)] букв открытого текста. Анаграмма порождает проблему восстановления открытого текста по его составу. Если бы анаграммы можно было решать систематическим образом, то рухнуло бы все шифрование перестановками.

6.3.1. Истоки. Использование анаграмм имеет богатую историю. Астроном Гюйгенс оставил следующую анаграмму:

[5)] То есть все повторяющиеся символы считаются как один символ. Это утверждение также тривиально справедливо для множеств.

$$a^7\ c^5\ d^1\ e^5\ g^1\ h^1\ i^7\ l^4\ m^2\ n^9\ o^4\ p^2\ q^1\ r^2\ s^1\ t^5\ u^5,$$

которую можно прочитать как *«annulo cingitur tenui piano, nusquam co-haerente, ad echpticam inclmato»* («[Сатурн] окружен тонким плоским кольцом, нигде не касающимся и наклоненным к эклиптике»)[6].

Ньютон послал Лейбницу такую анаграмму:

$$a^7\ c^2\ d^2\ e^{14}\ f^2\ i^7\ l^3\ m^1\ n^8\ o^4\ q^3\ r^2\ s^4\ t^8\ v^{12}\ x^1,$$

которая, вероятно, означала *«data aequatione quodcumque fluentes quantitates involvcnte, Quxiones fluxiones et vice versa»* («из данного уравнения с произвольным числом *флюент* (переменных), чтобы найти *флюксии* (производные) и наоборот»).

В XVII в. анаграммы были обычным развлечением ученых и они до сих пор привлекают внимание любителей.

Галилей отправил Кеплеру замаскированную анаграмму:

HAEC IMMATURA A ME IAM FRUSTRA LEGUNTUR O. Y.

(Эти незрелые вещи читаются ныне мною напрасно); что должно было означать *«cynthiae figuras aemulatur mater amorum»* (мать любви [= Венера] подражает фазам Синтии [= Луна]).

Современный пример — ASTRONOMES (астрономы), что можно прочитать как *moon starers* (наблюдающие за луной), но также и *no more stars* (нет больше звезд).

Баварский король Людвиг II (безумный строитель Neuschwanstein) записал (не очень подумав) замаскированную анаграмму MEICOST ETTAL, в которой должно читаться *l' état c'est moi* (государство — это я).

В фармацевтической промышленности также используются анаграммы: торговая марка KLINOMYCIN® означает препарат *Minocyclin* (миноциклин). Это только один пример использования игры слов в продвижении продукта.

В экспериментальной поэзии можно найти целые анаграммные поэмы, наподобие следующего фрагмента из поэмы Франческо Гальярди:

> Glück und Sommer weinen Waden, Röhricht neu,
> Rad und Röcke suchen Note: Glühweinwirnmern.
> Randenhügel, Wut und Nock: wie Öre schimrnern.
> Wandertürme, Gnom in Köchern wund, eil scheu.

с составом

$$a^1\ c^2\ d^2\ e^5\ g^1\ h^2\ i^2\ k^1\ l^1\ m^2\ n^4\ o^1\ r^3\ s^1\ t^1\ u^2\ w^2\ \ddot{o}^1\ \ddot{u}^1.$$

А вот примеры на русском языке (*добавлено переводчиком*):

[6] В анаграмме перечисляются буквы, входящие в состав сообщения вместе с их кратностью. — *Прим. перев.*

Что нам весна или за ней дано?
Одна мечта: знай сан и лей вино! (В. Брюсов)

Анаграммы до сих пор популярны в среде британских интеллектуалов (рис. 46). Они также встречаются в качестве загадок в немецких еженедельниках:

IRI BRÄTER, GENF	Briefträgerin
FRANK PEKL, REGEN	Krankenpfleger
PEER ASTIL, MELK	Kapellmeister
INGO DILMUR, PEINE	Diplomingenieur
EMIL REST, GERA	Lagermeister
KARL SORDORT, PEINE	Personaldirektor
GUDRUN SCHRILL, HERNE	Grundschullehrerin

admonition	domination	alarmingly	marginally
algorithms	logarithms	alienators	senatorial
ancestries	resistance	antagonist	stagnation
auctioning	cautioning	australian	saturnalia
broadsides	sideboards	catalogued	coagulated
catalogues	coagulates	certifying	rectifying
collapsing	scalloping	compressed	decompress
configures	refocusing	conserving	conversing
contenting	contingent	coordinate	decoration
countering	recounting	creativity	reactivity
dealership	leadership	decimating	medicating
decimation	medication	deductions	discounted
denominate	emendation	denotation	detonation
denouncers	uncensored	deposition	positioned
descriptor	predictors	directions	discretion
discoverer	rediscover	earthiness	heartiness
egocentric	geocentric	enduringly	underlying
enervating	venerating	enervation	veneration
excitation	intoxicate	filtration	flirtation
harmonicas	maraschino	impregnate	permeating
impression	permission	impressive	permissive
indiscreet	iridescent	introduces	reductions
mouldering	remoulding	nectarines	transience
ownerships	shipowners	percussion	supersonic
persistent	prettiness	persisting	springiest
pertaining	repainting	petitioner	repetition
platitudes	stipulated	positional	spoliation
procedures	reproduces	profounder	underproof

Рис. 46. Анаграммы из десятибуквенных слов (Хью Кейсмент)

6.3.2. Однозначность. Возникает вопрос, можно ли из кучи букв собрать более одного осмысленного сообщения. Джонатан Свифт уже ответил на этот вопрос, показав в своей сатире *Путешествия Гулливера*, как злонамеренный политический противник может превратить невинную фразу

OUR BROTHER TOM HATH JUST GOT THE PILES

(наш брат Том Хас нажил геморрой)

с помощью перестановки (*«анаграмматический метод»*) в конспиративное сообщение

Resist, — a Plot is brought home — The Tour

(Сопротивляться — заговор пришел в дом — Путешествие).

Действительно, как показывает опыт и как подтверждается теорией Шеннона, не существует длины, при которой анаграмма будет иметь единственную расшифровку.

В историческом плане нужно добавить, что самая ранняя форма перестановок встречается в древнегреческом *скитале* ($\sigma\kappa\upsilon\tau\lambda\varepsilon$), приспособлении для шифрования, которое известно с V в. до н. э. Он представлял собой деревянный цилиндр, на который наматывалась полоска папируса. Секретное сообщение записывалось на папирусе вдоль цилиндра.

После упадка классической культуры и краха Римской империи, первое шифрование перестановками обнаруживается, согласно Бернхарду Бишоффу, в средневековых монастырских рукописях: здесь встречается и оборачивание, и вертикальное письмо, и игра слов, часто довольно небрежная.

Перестановки утратили свое значение после распространения механических шифровальных машин в начале XX в., так как в механическом устройстве было трудно реализовать хранение большого числа символов. Однако с тех пор положение вещей изменилось. Нынешняя полупроводниковая технология обеспечивает достаточную память для эффективного шифрования перестановками, и крохотные чипы с очень малым временем доступа имеют емкость в миллионы битов и стоят не дороже автобусного билета. XXI век увидит, как перестановка восстанавливает свое значение.

Глава 7

Многоалфавитное шифрование: семейства алфавитов

Одноалфавитное шифрование многократно использует фиксированный шаг шифрования (возможно, многосимвольный). Все шаги шифрования, рассмотренные в гл. 3–6, могут использоваться с одним алфавитом — в примерах это предполагалось по умолчанию. Подлинное многоалфавитное шифрование требует, чтобы множество M доступных шагов шифрования имело по крайней мере два элемента, т. е. чтобы система шифрования M содержала, как минимум, $\theta = 2$ элемента. В частом случае $\theta = N$, где $N = |V|$, во французской литературе применяется термин *chiffre carré*.

Отдельные шаги шифрования могут иметь различную природу, например система шифрования M может состоять из одной или нескольких простых подстановок и одной или нескольких перестановок. Это вполне может свести с ума профессионального взломщика шифров, так как обычно все шаги шифрования одной системы принадлежат одному и тому же ограниченному классу — например, это подстановки, или линейные подстановки одного размера, или перестановки. Часто даже требуется, чтобы все шаги имели равную ширину шифрования, и более того, по техническим причинам может потребоваться блочное шифрование.

Главная проблема состоит в том, как простым способом охарактеризовать большое число различных шагов шифрования или, другими словами, как генерировать много различных алфавитов. Удивительно, но воображение изобретателей пока оставляет незадействованными многие возможности.

7.1. Итерирование подстановок

Естественная идея состоит в том, чтобы строить систему шифрования, систематически порождая все ее шаги из одного шага шифрования (первичный

алфавит, нем. *Referenzalphabet*). Это мы видели в случае простой подстановки в разд. 3.2.4, где семейства порожденных алфавитов получались применением всех сдвигов и возведений в степень.

Мы увидим, что оба эти семейства конструируются с использованием понятия итерированной подстановки S^i, определяемой выражением $pS^{j+1} = (pS)S^j$ для необходимо эндоморфной ($V = W$) подстановки S. Фактически, итерированная подстановка является основной при генерировании используемых алфавитов.

7.1.1. Сконцентрируем наше внимание на эндоморфном случае $V \doteqdot Q$, $W \doteqdot Q$. Пусть $S\colon Q^n \leftrightarrow Q^n$. Тогда $S^i\colon Q^n \leftrightarrow Q^n$ и мы имеем

$$(a^\circ) \quad \{S^i\colon i \in \mathbb{N}\},\ \textit{группу} \text{ степеней смешанного алфавита } S.$$

С подстановками $P_1\colon V \leftrightarrow Q^n$ и $P_2\colon Q^n \leftrightarrow W$ связано множество

$$(a) \quad \{P_1S^iP_2\colon i \in \mathbb{N}\}, \text{ где } P_1S^iP_2\colon V \leftrightarrow W,$$

с частными случаями $(')$ $V = Q^n$, $P_1 = \mathrm{id}$ и $('')$ $W = Q^n$, $P_2 = \mathrm{id}$. Кроме того, с дополнительной подстановкой $R\colon W^n \leftrightarrow Q^n$ мы связываем

$$(b^\circ) \quad \{S^iRS^i : i \in \mathbb{N}\},\ \textit{группу} \text{ подстановок, } S\text{-подобных подстановке } R.$$

Снова, с подстановками P_1, P_2, как и выше, связано множество

$$(b) \quad \{P_1S^iRS^iP_2\colon i \in \mathbb{N}\}, \text{ где } P_1S^iRS^iP_2\colon V \leftrightarrow W.$$

Семейства во всех случаях конечны, так как $Q^n \leftrightarrow Q^n$ при конечной мощности $|Q| = N$ содержит не более $(N^n)!$ различных перестановок.

Для подстановки S с порядком $h \leqslant (N^n)!$, т. е. $S^h = \mathrm{id}$ и $S^i \neq \mathrm{id}$ при $i < h$, степени S порождают h различных алфавитов. Заметим, что неравенство $h > N^n$ возможно: при $N = 5$, $q = 1$, подстановка (в циклической нотации) $(ab)(cde)$ имеет порядок 6. Может случиться, что h довольно мало: для взаимообратной подстановки S имеем $h = 2$.

7.1.2. Иногда может быть выгодным выбрать в качестве S циклическую перестановку σ. Порядок такой перестановки равен N^n. Имеется $(N^n - 1)!$ различных циклических перестановок. В этом случае порождение всех N^n степеней σ можно механизировать, как уже указывалось в разд. 3.2.8 (рис. 28).

7.2. Сдвигаемые и вращаемые алфавиты

Как только выделен стандартный алфавит в Q, сразу появляется стандартный алфавит в Q^n, фиксируемый лексикографическим упорядочением. Цикл, принадлежащий этому упорядочению (разд. 3.2.3), и соответствующую подстановку стандартного алфавита будем далее обозначать ρ. Подстановки P_1 и P_2, появившиеся выше, действуют тогда как первичные алфавиты.

7.2.1. Из ρ^i для S^i разд. 7.1.1 мы получаем степенные циклы

$$(\mathrm{a}^\circ) \quad \{\rho^i : i \in \mathbb{N}\} = \{\rho^i \rho : i \in \mathbb{N}\} = \{\rho\rho^i : i \in \mathbb{N}\},$$

группу сдвигаемых стандартных алфавитов (фр. *alphabets normalement Parallèle*, нем. *verschobene Standardalphabete*). Частные случаи (a):

(a′) $\{\rho^i P : i \in \mathbb{N}\}$ — множество горизонтально сдвигаемых (смешанных) P-алфавитов (фр. *alphabets désordonné et Parallèle*),

(a″) $\{P\rho^i : i \in \mathbb{N}\}$ — множество вертикально продолженных (смешанных) P-алфавитов (фр. *alphabets désordonné etendu verticalement*).

В общем случае (Эйрауд: *alphabets non-normalement Parallèle*)

(a*) $\{P_1\rho^i P_2 : i \in \mathbb{N}\}$ см. разд. 8.2.3 (и разд. 19.5.3).

Введенные обозначения станут понятными после просмотра приведенных ниже таблиц для семейств подстановок: с $V = Q = W = Z_{26}$ и $N = 26$. Для первичного алфавита P, порожденного мнемоническим паролем NEWYORK-CITY,

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
N	E	W	Y	O	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X	Z

множество $\{\rho^i P : i \in \mathbb{N}\}$ имеет следующее табличное представление (в форме *tabula recta*, т. е. с одинаковыми буквами по обратным диагоналям):

i	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	N	E	W	Y	O	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X	Z
1	E	W	Y	O	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X	Z	N
2	W	Y	O	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X	Z	N	E
3	Y	O	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X	Z	N	E	W
4	O	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X	Z	N	E	W	Y
5	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X	Z	N	E	W	Y	O
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
25	Z	N	E	W	Y	O	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X

а множество $\{P\rho^i : i \in \mathbb{N}\}$ имеет табличное представление такого вида:

i	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	N	E	W	Y	O	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X	Z
1	O	F	X	Z	P	S	L	D	J	U	B	C	E	G	H	I	K	M	N	Q	R	T	U	W	Y	A
2	P	G	Y	A	Q	T	M	E	K	V	C	D	F	H	I	J	L	N	O	R	S	U	V	X	Z	B
3	Q	H	Z	B	R	U	N	F	L	W	D	E	G	I	J	K	M	O	P	S	T	V	W	Y	A	C
4	R	I	A	C	S	V	O	G	M	X	E	F	H	J	K	L	N	P	Q	T	U	W	X	Z	B	D
5	S	J	B	D	T	W	P	H	N	Y	F	G	I	K	L	M	O	Q	R	U	V	X	Y	A	C	E
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
24	L	C	U	W	M	P	I	A	G	R	Y	Z	B	D	E	F	H	J	K	N	O	Q	S	T	V	X
25	M	D	V	X	N	Q	J	B	H	S	Z	A	C	E	F	G	I	K	L	O	P	R	T	U	W	Y

В *горизонтально сдвигаемом P-алфавите* первичный алфавит P в каждой строке сдвигается относительно предыдущей строки на одну позицию влево; в вертикально сдвигаемом P-алфавите первичный алфавит P фигурирует только в первой строке, а в остальных столбцах он продолжается в стандартном порядке, начиная с буквы из первой строки.

7.2.2. Дополнительно к подстановкам, S-подобным подстановке R из разд. 7.1.1, мы имеем

$$(\mathrm{b}^\circ) \quad \{\rho^{-i}R\rho^i : i \in \mathbb{N}\},$$

группу R-вращений (обозначение будет объяснено в разд. 7.3) стандартных алфавитов.

Для частного случая $P_1 = P$, $P_2 = P^{-1}$ из (b°) получаем

$$(\mathrm{b}^*) \quad \{P\rho^{-i}R\rho^i P^{-1} : i \in \mathbb{N}\},$$

множество R-вращений (смешанных) P-алфавитов.

Беря теперь, как и выше, в качестве R тот же самый первичный алфавит NEWYORKCITY, получим табличное представление множества $\{\rho^{-i}R\rho^i : i \in \mathbb{N}\}$:

i	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	N	E	W	Y	O	R	K	C	I	T	A	B	D	F	G	H	J	L	M	P	Q	S	U	V	X	Z
1	A	O	F	X	Z	P	S	L	D	J	U	B	C	E	G	H	I	K	M	N	Q	R	T	V	W	Y
2	Z	B	P	G	Y	A	Q	T	M	E	K	V	C	D	F	H	I	J	L	N	O	R	S	U	W	X
3	Y	A	C	Q	H	Z	B	R	U	N	F	L	W	D	E	G	I	J	K	M	O	P	S	T	V	X
4	Y	Z	B	D	R	I	A	C	S	V	O	G	M	X	E	F	H	J	K	L	N	P	Q	T	U	W
5	X	Z	A	C	E	S	J	B	D	T	W	P	H	N	Y	F	G	I	K	L	M	O	Q	R	U	V
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
21	M	F	X	D	O	V	W	Y	A	B	C	E	G	H	K	L	N	P	Q	S	U	I	Z	R	T	J
22	K	N	G	Y	E	P	W	X	Z	B	C	D	F	H	I	L	M	O	Q	R	T	V	J	A	S	U
23	V	L	O	H	Z	F	Q	X	Y	A	C	D	E	G	I	J	M	N	P	R	S	U	W	K	B	T
24	U	W	M	P	I	A	G	R	Y	Z	B	D	E	F	H	J	K	N	O	Q	S	T	V	X	L	C
25	D	V	X	N	Q	J	B	H	S	Z	A	C	E	F	G	I	K	L	O	P	R	T	U	W	Y	M

Первичный алфавит R фигурирует только в первой строке таблицы *R-вращаемых стандартных алфавитов* и *продолжается по диагоналям* в *стандартном* порядке, начиная с буквы из первой строки (предполагается, что столбцы закольцованы как цилиндр: за последним столбцом идет первый).

7.2.3. Семейство сдвигаемых первичных алфавитов может быть механизировано, как упоминалось в разд. 3.2.7, с помощью шифровального диска Альберти или шифровальной линейки. Мы будем говорить о шагах шифрования ALBERTI в случае горизонтально сдвигаемых P-алфавитов и шагах шифрования роторных машин в случае R-вращаемых стандартных алфавитов.

7.2.4. Для нас интересно циклическое разложение сопутствующих алфавитов.

Пример. Для $Q = \begin{pmatrix}\text{abcde}\\ \text{badec}\end{pmatrix}$ и $\rho = (\text{abcde}) = \begin{pmatrix}\text{abcde}\\ \text{bcdea}\end{pmatrix}$ получаем

$$\rho Q = \begin{pmatrix}\text{abcde}\\ \text{bcdea}\end{pmatrix}\begin{pmatrix}\text{bcdea}\\ \text{adecb}\end{pmatrix} = \begin{pmatrix}\text{abcde}\\ \text{cbead}\end{pmatrix}$$

$$Q\rho = \begin{pmatrix}\text{abcde}\\ \text{badec}\end{pmatrix}\begin{pmatrix}\text{badec}\\ \text{cbead}\end{pmatrix} = \begin{pmatrix}\text{abcde}\\ \text{cbead}\end{pmatrix}$$

$$\rho^{-1}Q\rho = \begin{pmatrix}\text{abcde}\\ \text{eabcd}\end{pmatrix}\begin{pmatrix}\text{eabcd}\\ \text{dcbea}\end{pmatrix} = \begin{pmatrix}\text{abcde}\\ \text{dcbea}\end{pmatrix}.$$

В подстановочной нотации алфавиты этого примера выглядят так:

i	a	b	c	d	e
0	B	A	D	E	C
1	A	D	E	C	B
2	D	E	C	B	A
3	E	C	B	A	D
4	G	B	A	D	E

i	a	b	c	d	e
0	B	A	D	E	C
1	C	B	E	A	D
2	D	C	A	B	E
3	E	D	B	C	A
4	A	E	C	D	B

i	a	b	c	d	e
0	B	A	D	E	C
1	D	C	B	E	A
2	B	E	D	C	A
3	B	C	A	E	D
4	E	C	D	B	A

В циклической нотации получаем следующее представление: для множества горизонтально сдвигаемых P-алфавитов

{(a b)(c d e), (a)(b d c e), (a d b e)(c), (a e d)(b c), (ac)(b)(d)(e)};

для множества вертикально продолженных P-алфавитов

{(a b)(c d e), (a c e d)(b), (a d b c)(e), (a e)(b d c), (a)(be)(c)(d)};

для множества R-вращаемых стандартных алфавитов

{(a b)(c d e), (b c)(d e a), (c d)(e a b), (d e)(a be), (e a)(b c d)}.

Из теории групп известно, что преобразование подобия $\rho^{-1}Q\rho$ оставляет неизменной длину циклов циклического разложения перестановки Q. Все подстановки множества R-вращаемых алфавитов имеют одинаковое циклическое разложение (это можно назвать «основной теоремой роторного шифрования»). Не так обстоит дело для сдвигаемых P-алфавитов.

В нашем случае разбиение, принадлежащее циклическому разложению, имеет вид $3+2$. В примере разд. 7.2.2 разбиение имеет вид $10+8+6+1+1$, принадлежа циклическому разложению

(a n f r l b e o g k) (c w u q j t p h) (d y x v s m) (i) (z).

7.2.5. Число различных сопутствующих алфавитов точно равно N^n для случая (a), для случая (b) оно лежит между 1 и N^n в зависимости от R. Оно равно 1, если P является тождественной перестановкой, и равно N^n, если $\rho^j P \neq P\rho^j$ для $j = 1, 2, \ldots, N^n - 1$.

Для малых значений N^n имеется только несколько «роторных» R, удовлетворяющих этим условиям. Для $N^n = 4$ и $\rho = (\mathrm{a\ b\ c\ d})$ имеется ровно четыре максимальных «роторных» множества:

$$\{(\mathrm{a\ b}),\ (\mathrm{b\ c}),\ (\mathrm{c\ d}),\ (\mathrm{d\ a})\},$$
$$\{(\mathrm{a\ c\ b\ d}),\ (\mathrm{b\ d\ c\ a}),\ (\mathrm{c\ a\ d\ b}),\ (\mathrm{d\ b\ a\ c})\};$$
$$\{(\mathrm{a\ c\ b}),\ (\mathrm{b\ d\ c}),\ (\mathrm{c\ a\ d}),\ (\mathrm{d\ b\ a})\},$$
$$\{(\mathrm{a\ b\ c}),\ (\mathrm{b\ c\ d}),\ (\mathrm{c\ d\ a}),\ (\mathrm{d\ a\ b})\};$$

для $N^n = 3$ и $\rho = (\mathrm{b\ c})$ существует только одно «роторное» множество: $\{(\mathrm{ab}),\ (\mathrm{bc}),\ (\mathrm{ca})\}$. Для $N^n = 2$ вообще нет «роторных» множеств из двух циклов.

7.3. Роторные шифровальные машины

С появлением электрических печатающих устройств на первый план вышли электромеханические шифровальные машины. Для реализации фиксированной подстановки P в виде электрической схемы можно воспользоваться коммутатором с N разъемами на входе для символов открытого текста и N разъемами на выходе для символов шифротекста, связанных внутри N проводами, рис. 47 (a).

Чтобы получить электромеханическую реализацию множества $\{P\rho^i\}$ сдвигаемых P-алфавитов, после выходных разъемов коммутатора размещают скользящие контакты; еще лучше обеспечить скольжение коммутатора по контактной панели, как показано на рис. 47 (b). В любом случае, необходимы гибкие проводники, что приводит к механическим поломкам.

Этого можно избежать, если между неподвижными входами и выходами поместить скользящий коммутатор. В этом случае гибкие провода больше не нужны. Подобная схема показана на рис. 47 (c). В дублировании контактов нет никакой надобности, если использовать подвижный переключающий барабан — ротор (нем. *Walze*). Таким образом, получается реализация множества

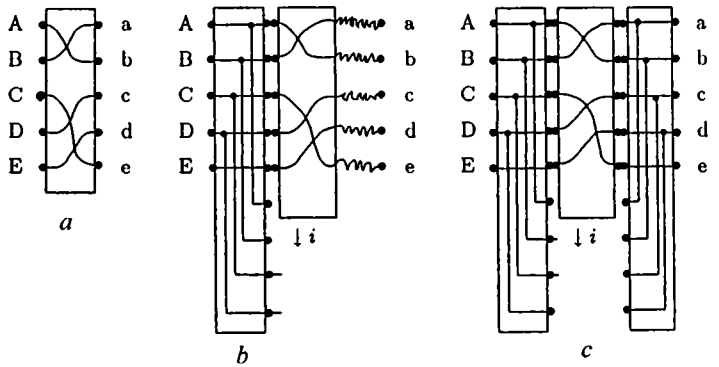


Рис. 47. Фиксированная подстановка и сдвиг, реализованные электрическими схемами

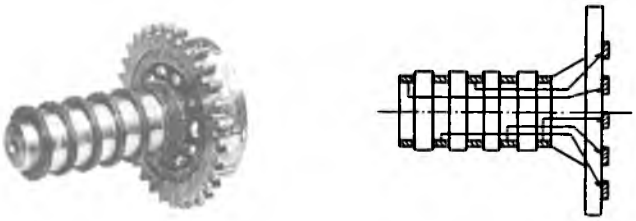


Рис. 48. «Полу-ротор» Арвида Дамма (1919 г.)

алфавитов $\{\rho^{-i}P\rho^{i}\}$, которая и дает название R-вращаемым (стандартным) алфавитам.

Использование скользящих колец с переключающим барабаном обеспечивает реализацию сдвигаемых P-алфавитов $\{P\rho^{i}\}$. Ранее оно получило название «полу-ротор» (рис. 48).

7.3.1. Истоки. Идея электрического контактного ротора зародилась до 1920 г. независимо у четырех человек. Как пишет Кан, из материалов патентного бюро видно (свидетельство 77716 бюро патентов США), что американец Эдвард Хью Хеберн (1869–1952 гг.), который в 1915 г. посимвольно соединил 26-ю проводами два электрических печатающих устройства, в 1917 г. сделал первый эскиз ротора для механической смены соединений и, таким образом, получил возможность использовать 26 алфавитов. Он подал заявку на патент США только в 1921 г. и получил его (№ 1510441) в 1924 г. Три других заявки были поданы раньше: немцем Артуром Шербиусом (германский патент 416219; заявка подана 23 февраля 1918 г.), почти в это же время, голландцем Хьюго Александером Кохом (1870–1928 гг.), зарегистрировавшим патент 7 октября 1919 г., и шведом Арвидом Герхардом Даммом, который зарегистрировал свой патент 10 октября 1919 г. (полу-роторы).

Э. Х. Хеберн *А. Г. Дамм* *А. Шербиус*

Ни один из этих изобретателей не нажил ни состояния, ни счастья. С Хеберном обошлись очень скверно ВМС США в 1934 г. и, позже, правительство США; в 1941 г. он уступил свой патент ИБМ. Хеберн имел небольшой доход, когда скончался от сердечного приступа в возрасте 82 лет. Кох умер уже в 1928 г.; к тому времени компания Шербиуса скупила его патенты. Сам Шербиус (30.10.1878 г.–13.05.1929 г.) погиб при несчастном случае; его фирма *Chiffriermaschinen Aktiengesellschaft* (находившаяся в Берлине по адресу В 36, Штиглитцштрассе, 2) позже была переименована в Heimsoeth & Rinke и просуществовала до 1945 г. Дамм был, как говорится, *дамским угодником* (*homme galant*) и умер в 1928 г., его фирма перешла в руки Бориса Хагелина (2.06.1892 г.–7.09.1983 г.), который в 1935 г. отказался от полу-ротора и в

1939 г. переименовал компанию в *Aktiebolaget Cryptograph*. Дамм со своими парами полу-роторов с 5-ю штифтами, которые он использовал для шифра Полибиева типа, также оказался не у дел. В патентах Шербиуса в демонстрационных целях использовались роторы с 10-ю штифтами, подходящие для шифрования числовых кодов.

В приложении к своему патенту 1918 г. Шербиус рассматривал кратные роторы, числом до десяти, используемые последовательно. Аналогичным образом, Хеберн использовал пять роторов (два из которых имели фиксированное положение), а Шербиус в первой коммерческой модели ЭНИГМА A и в ЭНИГМА B 1923 г. использовал четыре ротора (нем. *Durchgangsräder*), см. вклейку К. В последнем случае имеется множество сопутствующих алфавитов $\{R_{(i_1,i_2,i_3,i_4)}\}$ с

$$R_{(i_1,i_2,i_3,i_4)} = \rho^{-i_1} R_N \rho^{i_1-i_2} R_M \rho^{i_2-i_3} R_L \rho^{i_3-i_4} R_K \rho^{i_4},$$

которое при соответствующем выборе R_K, R_L, R_M, R_N состоит из $26^4 = 456976$ элементов.

7.3.2. Отражение. Несколькими годами позже, когда в 1926 г. появилась коммерческая ЭНИГМА, коллега Шербиуса Вилли Корн (заявка от 21 марта 1926 г., германский патент 452194) предложил разумную идею, состоящую в добавлении в конструкцию отражателя сигналов (нем. *Umkehrwalze*, в Блетчли Парк часто ошибочно писали *Umkerwaltz* и произносили «анкл волтер»). Пользуясь теперь лишь тремя сменными роторами R_L, R_M, R_N и соответствующей взаимообратной подстановкой U (которая требует четности N), мы имеем (рис. 49) множество $\{P_{(i_1,i_2,i_3)}\}$, где

$$P_{(i_1,i_2,i_3)} = S_{(i_1,i_2,i_3)} U S^{-1}_{(i_1,i_2,i_3)} \quad \text{и}$$

$$S_{(i_1,i_2,i_3)} = \rho^{-i_1} R_N \rho^{i_1-i_2} R_M \rho^{i_2-i_3} R_L \rho^{i_3}.$$

Все элементы этого множества являются теперь взаимообратными подстановками. Считается преимуществом, когда процедуры шифрования и расшифровки идентичны и отсутствует необходимость в переключателе. Роторная схема, однако, имела следствием то, что никакая буква не могла быть зашифрована сама собой. Это должно привести в конечном итоге к большой угрозе безопасности шифрования (разд. 11.2.4, 14.5.1 и 19.7.2).

Дальнейшая судьба Корна не известна. Последний патент, который он получил в Германии, датирован 5 марта 1930 г.; после него следы Корна в Берлине теряются.

Стоимость ЭНИГМЫ доходила в то время до 600 рейхсмарок ($140). На рис. 49 схематично показано течение электрического тока для буквы открытого текста *e* и соответствующей буквы криптотекста M, S обозначает статор. В ЭНИГМЕ C отражатель U устанавливался в двух фиксированных положениях, но в коммерческой ЭНИГМЕ D 1927 г. он мог поворачиваться подобно трем обычным роторам; внешне он напоминал дополнительный четвертый ротор — отражательный ротор U.

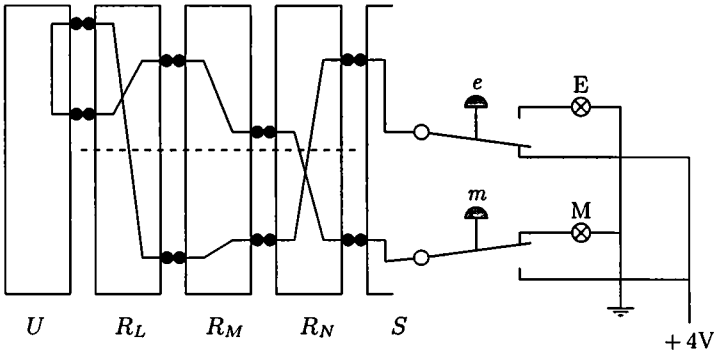


Рис. 49. Электрический ток в 3-роторной ЭНИГМЕ, проходящий от клавиши e до лампочки M

Таким образом, мы имеем семейство $\{P_{(i_1,i_2,i_3,i_4)}\}$ соответствующих инволюций, где

$$P_{(i_1,i_2,i_3,i_4)} = S_{(i_1,i_2,i_3,i_4)} U S^{-1}_{(i_1,i_2,i_3,i_4)} \quad \text{и}$$

$$S_{(i_1,i_2,i_3,i_4)} = \rho^{-i_1} R_N \rho^{i_1-i_2} R_M \rho^{i_2-i_3} R_L \rho^{i_3-i_4}.$$

Коммерческая ЭНИГМА D (Уэлчман: «машина-лампочка») первоначально комплектовалась резервными роторами; машина широко использовалась и поставлялась в Швецию, Голландию, Англию, Японию, Италию, Испанию, США и была куплена польским Бюро шифров (*Biuro Szyfrów*). Ее преемница ЭНИГМА K была поставлена в 1938–1940 гг. швейцарской армии (в США имела кодовое имя INDIGO). В более поздних моделях для Рейхсвера, отражающий ротор был вновь зафиксирован ($i_4 = 0$).

15 июля 1928 г. польское Бюро шифров впервые перехватило радиосообщения *Рейхсвера*, зашифрованные ЭНИГМОЙ. Немецкий военный флот (*Reichsmarine*) в 1925 г. начал эксперименты с 28-контактной 3-роторной ЭНИГМОЙ (*Funkschlüssel C*, 1926) с упорядоченной в алфавитном порядке клавиатурой включающей дополнительные буквы Ä, Ö, Ü (буква X пропускалась). Отражатель теперь можно было вставить в четырех фиксированных положениях, обозначаемых α, β, γ, δ. В 1933 г. была незначительно модифицирована *Funkschlüssel C*; проводилось тестирование версии с буквами Ä, Ü. *Рейхсверовские* модели шифромашин ЭНИГМА G (1928 г.) и ЭНИГМА I (1930) снова имели только 26 контактов и клавиатуру стандартной немецкой пишущей машинки; соединение с контактами статора производилось в алфавитном порядке. Отражатель имел одно фиксированное положение. Имелась также ЭНИГМА II с пишущей машинкой.

Армейская ЭНИГМА I (ЭНИГМА-один; введена в эксплуатацию 1 июня 1930 г.), которая позже стала общей ЭНИГМОЙ Вермахта, была защищена коммутационной панелью, которая предусматривалась для дополнитель-

Рис. 50 a. Трехроторная ЭНИГМА *Вермахта* (1937 г.)

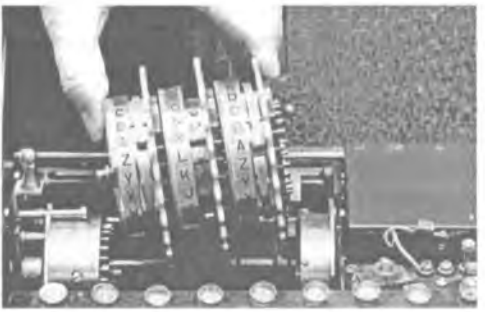


Рис. 50 b. Три сменных ротора ЭНИГМЫ *Вермахта* (1937 г.)

ной (необязательно взаимообратной) подстановки T соединений, называемой *Steckerverbindung* (перекрестное подключение). Такая конструкция приводит к уравнению ЭНИГМЫ

$$c_i = p_i T S_i U S_i^{-1} T^{-1} \quad (= p_i T U_i T^{-1})$$

между буквами открытого текста p_i и буквами c_i; имеется изоморфизм (разд. 2.6.3) между $c_i T S_i$ и $p_i T S_i$, так как

$$c_i T S_i = p_i T S_i U.$$

7.3.3. ЭНИГМА Вермахта. В 1934 г. флот и армия согласились на общую версию ЭНИГМЫ. Произошло это под давлением полковника Эриха Фельгебеля (1886–1944 гг.), будущего (с 1939 г.) генерал-майора и главы управления связи OKW.

Три ротора ЭНИГМЫ военно-морского флота (*Funkschlüssel M*) могли выбираться из набора, соответственно, в пять (1934 г.), семь (1938 г.) или восемь (1939 г.) роторов (кроме того, они могли переставляться). Роторы отмечались римскими цифрами I, ..., VIII. До 12 декабря 1938 г. армия использовала только три из пяти возможных роторов. ВВС с 1935 г. также приняли на вооружение ЭНИГМУ Вермахта (рис. 50 a) для нового *Luftnachrichtentruppe*. Три сменных ротора показаны на рис. 50 b.

Ведомство железных дорог (*Deutsche Reichsbahn*), почта (*Deutsche Reichspost*) и полиция использовали менее надежные устаревшие модели без коммутационной панели, хотя, например, сообщения о железнодорожных составах в России могли бы выдать противнику много ключей.

Первый отражатель в ЭНИГМЕ Вермахта, маркированный «A», 1 ноября 1937 г. был заменен на версию «B». В 1941 г. появилась версия отражателя «C». «Съемный» отражатель «D» с возможностью перемонтажа первый раз использовали 2 января 1944 г. в Норвегии. Коммутация отражателя изменялась три раза в месяц (рис. 51 b).

В 1944 г. в комплекте ЭНИГМЫ Вермахта появилось специальное устройство, блок *Uhr* (вклейка M), для замены топологии коммутационной панели невзаимообратной подстановкой, которую также можно было легко менять поворотом ручки. Несмотря на дополнительную защиту, которую обеспечивало это устройство, оно широко не использовалось.

Роторы в ЭНИГМЕ могли устанавливаться в произвольном порядке. До конца 1935 г. порядок роторов (нем. *Walzenlage*) и перекрестное подключение менялись раз в три месяца. С 1 января 1936 г. они стали меняться каждый месяц; с 1 октября 1936 г. — каждый день. Позднее, во время Второй мировой войны, замена производилась каждые восемь часов. Вопрос, почему не ранее?

Военно-морской флот сомневался в стойкости ЭНИГМЫ и 1 февраля 1942 г. ввел для ключевой сети связи TRITON новую версию *Funkschlüssel M4* (вклейка I) с четвертым ротором, промаркированным β и поэтому называемым *Griechenwalze* («греческий ротор»). Дополнительный ротор можно было поворачивать, но нельзя было двигать в сеансе шифрования. ЭНИГМА

VIII. Beiſpiel.

17. Gültiger Tagesſchlüſſel:

(Ausſchnitt aus der für die Verſchlüſſelung des Klartextes in Betracht kommenden Schlüſſeltafel, z. B. ».....« Maſchinenſchlüſſel für Monat Mai«)

Datum	Walzenlage	Ringſtellung	Grundſtellung
4.	I III II	16 11 13	01 12 22

Steckerverbindung	Kenngruppen-Einſatzſtelle ..... Gruppe	Kenngruppen
CO DI FR HU JW LS TX	2	adq nuz opw vxz

Рис. 51 a. Армейская инструкция по эксплуатации ЭНИГМЫ Вермахта, датированная 8 июля 1937 г. Установка положения колец и порядка роторов в ВМФ были прерогативой офицера

Geheime Kommandosache! Jeder einzelne Tagesschlüssel ist geheim! Unmittelbar in Empfang berühren! № 000082

Luftwaffen-Maschinen-Schlüssel Nr. 2744

Achtung! Schlüsselmittel dürfen nicht unversehrt in Feindeshand fallen. Bei Gefahr restlos und frühzeitig vernichten.

Tag	Walzenlage	Ringstellung	Steckerverbindungen	Kenngruppen
2744	31 III V IV	17 11 04	TW BI UY GF CK JQ DL RV EM AH	NS FO	kim pwb	sbx caw
2744	30 I IV V	08 17 21	LS DH MT EO AP UZ FQ WY BK GR	CI JN	uaq omn	ume duf
2744	29 V II III	11 14 05	DO JW GN IV PZ BM HU AL FR KX	EQ GT	don cqo	xum bpg
2744	28 II IV V	02 20 16	NT HK BW EP LQ AU OY FJ CX GI	DS MR	lui pyg	sby dtq
2744	27 III V IV	18 13 22	HM GV KZ AI DQ HR ES BL GU FT	CP JY	cmy fqr	scl bur
2744	26 I III II	24 10 01	GW AQ MO PV FS DI RU JZ BN EH	KT GL	kbj yaq	udm cns
2744	25 IV I III	04 25 23	LT DR QX AO IN EU BJ KP FW CM	SS HO	kqs yar	vdb coa
2744	24 V III I	09 19 06	GL MY OR HN JX DT AP FU IQ BO	EW KS	cms aoj	zod auh
2744	23 IV I V	15 03 19	IT DV KQ AJ MU EX KO OS PY LN	BP GZ	kra yas	xun cob
2744	22 I V III	12 26 07	EY JL AK NV PX CT HP MX BQ GS	DW IO	jdm uhf	xuo bph
2744	21 III IV II	15 09 12	JP DY QS HU AE NW CU IK PX BR	MV GO	jpf aok	iya btx
2744	20 IV II I	02 22 05	HT NP AM DX GJ KQ BS OV ER CW	IU FL	boy wac	uow cse
2744	19 V I II	08 19 17	GM OX BT QG DP HJ PK SW AN EL	CX IR	xje wad	unj ctd
2744	18 III IV I	11 21 01	KW IP DM SV JR CX EN AZ QT BU	FH GY	kpn rsi	vcm bpo
2744	17 I V II	18 23 14	BV HW AR NX DS PT CZ PI LY EJ	GK MQ	kdz crq	vcn cod
2744	16 III IV V	16 04 07	LU CV PM KR BY GN QW DJ FS AO	EI HX	lgx jri	uob aur
2744	15 V III IV	24 13 10	HZ NQ AD TV IX KM BQ LO CE RY	JU FP	wpt vhy	xoe aua
2744	14 I IV II	06 20 25	FN OY CJ IW LP AS DK GQ MO BZ	ET HR	wog hxi	zxi bpi
2744	13 III II I	03 26 18	KR IZ AT NV BH MP CG OT ES DF	UW LQ	lqv iqb	zsy coe
2744	12 II IV III	04 11 15	DT JV HS OI AY KU EN PQ LR BW	MP GO	zic myt	zof dtr
2744	11 V I IV	16 07 02	JS PW AV QX DN IZ KM CG EQ FL	HY BR	inf sbm	krs dug
2744	10 IV III II	20 12 14	FS CQ JO FR AW HV EZ KN DU GT	IL BY	ink acu	zxj cnu
2744	9 III II V	06 18 10	MK TZ MX LW GQ AD NY BE CS JP	RV IO	efm pmi	enw cof
2744	8 V I III	01 21 17	GU SW BP RX EV OT LQ CH IP KY	JM NZ	imy rjw	tjm cog
2744	7 II V I	25 08 23	CX AS DV KT HU LW GP EY MR FQ	IN OS	inv rkc	snx bpj
2744	6 IV II V	13 26 03	DV LP NQ GZ OS FK EW MR IT HX	UY BJ	yvu hsb	swq aut
2744	5 III I II	24 19 22	SY EK NZ OR CG JM QU PV BI LW	TX DP	aeu iqe	awr auv
2744	4 II IV I	17 05 09	BD GV AX KP EM FN CW RU HO JT	IL QS	sfj hxj	zxk dpt
2744	3 V II IV	20 16 11	JT NW DU EO KV BY PS HQ IM LX	OP CR	cix sbn	sxa buk
2744	2 II III V	14 03 19	RW CQ GI AS EJ MS CU DH PY BP	LV TX	ljs jre	spq coh
2744	1 III I IV	18 24 15	NP JV LY IX KQ AO DZ GR PT EM	OS HW	plf dgw	tjn cnv

Рис. 51 b. Журнал установок № 2744 для ЭНИГМЫ Люфтваффе (вероятно, на август или сентябрь 1944 г.)

с четырьмя роторами сначала использовалась только на немецких подводных лодках в Атлантике. С 1 июля 1943 г. стал применяться дополнительный ротор γ. Чтобы гарантировать совместимость новой 4-х роторной ЭНИГМЫ со старой 3-х роторной, старый отражатель «B» или «C» был разделен на фиксированный тонкий отражательный диск «B dünn» или «C dünn» и поворачиваемый дополнительный ротор β или γ, соответственно.

Другое изобретение сделал Пауль Бернштейн для ЭНИГМЫ A: кольцо, смонтированное по ободу ротора, которое позволяло считывать позиции ротора (алфавитное кольцо, нем. *Sperr-Ring*), было сделано подвижным, а его позиция относительно сердечника — кольцевая установка (нем. *Ringstellung*), могла быть зафиксирована штырьком.

На рис. 51 a показан раздел инструкции по эксплуатации ЭНИГМЫ с ежедневным порядком замены ротора, кольцевыми установками и параметрами перекрестного подключения. *Grundstellung* (базисные установки) характеризует расположение трех роторов в начале процесса зашифрования. *Kenngruppen* не имел никакого криптологического значения. На рис. 51 b изображен журнал установок ЭНИГМЫ Люфтваффе от 1944 г., указывающий, что в 1944 г. съемный отражатель изменялся каждые 10 дней, а параметры перекрестного переключения изменялись каждые 8 часов.

Армия использовала ЭНИГМЫ начиная с уровня полков и выше. Оценки количества эксплуатируемых ЭНИГМ колеблются от 20 000 (оценка польской стороны) до 200 000 (Джонсон; явно завышенная оценка) к 1938 г., 40 000 на момент начала войны, приближаясь в целом к 100 000. После Второй мировой войны страны-победители продали захваченные ЭНИГМЫ, которые в то время повсеместно считались достаточно безопасными, развивающимся странам.

7.3.4. TYPEX. В Англии во Вторую мировую войну также были созданы роторные машины: шифратор TYPEX был улучшенным вариантом ЭНИГМЫ (вместо коммутационной панели применялась входная подстановка, производимая двумя фиксированными роторами, подстановка не была взаимообратной, разд. 22.2.7). В США под влиянием Уильяма Фридмана (1891–1969 гг.) и на основе разработок Хеберна в начале 1930-х годов была запущена независимая линия роторных машин, приведшая в 1933 г. к шифратору M-134-T2, затем к M-134-A (SIGMYC) и в 1936 г. к армейской машине M-134-C (SIGABA), называемой на флоте CSP889 (ECM Mark II). Немцы, очевидно, не преуспели во взломе SIGABA, которая была сделана Франком Роулеттом (1908–1998 гг.), помощником Фридмана с апреля 1930 г.

Интересный послевоенный вариант ЭНИГМЫ с семью роторами и фиксированным отражателем был сконструирован и выставлен на продажу итальянской компанией Ottico Meccanica Italiana (OMI) из Рима. Швейцарская армия использовала с 1946 г. вариант ЭНИГМЫ NEMA с четырьмя активными роторами, который был сконструирован фирмой Zeilweger AG, из Устера.

7.3.5. Подстановки ЭНИГМЫ.

7.3.5.1. Статор, три ротора и отражательный ротор ЭНИГМЫ D выполняют следующие подстановки (Кифер А. Девур, Луи Кру):

вход	a b c d e f g h i j k l m n o p q r s t u v w x y z
статор	J W U L C M N O H P Q Z Y X I R A D K E G V B T S F
выход ротора 1	L P G S Z M H A E O Q K V X R F Y B U T N I C J D W
выход ротора 2	S L V G B T F X J Q O H E W I R Z Y A M K P C N D U
выход ротора 3	C J G D P S H K T U R A W Z X F M Y N Q O B V L I E
выход отражателя	I M E T C G F R A Y S Q B Z X W L H K D V U P O J N

7.3.5.2. Статор, восемь роторов и отражательный ротор ЭНИГМЫ Вермахта выполняют следующие подстановки (Ральф Эрскин, Фроде Вейруд, Гейнц Ульбрихт):

вход	a b c d e f g h i j k l m n o p q r s t u v w x y z
статор	A B C D E F G U I J K L M N O P Q R S T U V W X Y Z
выход ротора I	E K M F L G D Q V Z N T O W Y H X U S P A I B R C J
выход ротора II	A J D K S I R U X B L H W T M C Q G Z N P Y F V O E
выход ротора III	B D F H J L C P R T X V Z N Y E I W G A K M U S Q O
выход роторам IV	E S O V P Z J A Y Q U I R H X L N F T G K D C M W B
выход ротора V	V Z B R G I T Y U P S D N H L X A W M J Q O F E C K
выход ротора VI	J P G V O U M F Y Q B E N H Z R D K A S X L I C T W
выход ротора VII	N Z J H G R C X M Y S W B O U F A I V L P E K Q D T
выход ротора VIII	F K Q H T L X O C B J S P D Z R A M E W N I U Y G V
выход отражателя A	E J M Z A L Y X V B W F C R Q U O N T S P I K H G D
выход отражателя B	Y R U H Q S L D P X N G O K M I E B F Z C W V J A T
выход отражателя C	F V P J I A O Y E D R Z X W G C T K U Q S B N M H L
выход ротора Beta	L E Y J V C N I X W P B Q M D R T A K Z G F U H O S
выход отражателя B dünn	E N K Q A U Y W J I C O P B L M D X Z V F T H R G S
выход ротора Gamma	F S O K A N U E R H M B T I Y C W L Q P Z X V G J D
выход отражателя C dünn	R D O B J N T K V E H M L F C W Z A X G Y I P S U G

Замечание: подстановка Beta, с последующими подстановками B dünn и Beta^{-1}, совпадает с подстановкой B, например: Beta(a) = L, B dünn(L) = O, Beta^{-1}(O) = y; таким образом, B(a) = y.

7.3.5.3. Роторы ЭНИГМЫ D и ЭНИГМЫ Вермахта имеют следующие циклические разбиения подстановок, допускающие простую идентификацию:

статор	13 + 7 + 3 + 2 + 1	ротор I	10 + 4 + 4 + 3 + 2 + 2 + 1
ротор 1	25 + 1	ротор II	8 + 7 + 3 + 2 + 2 + 2 + 1 + 1
ротор 2	17 + 7 + 2	ротор III	17 + 8 + 1
ротор 3	19 + 6 + 1	ротор IV	22 + 2 + 2
		ротор V	11 + 9 + 6
		ротор VI	14 + 8 + 4
		ротор VII	26
		ротор VIII	17 + 3 + 3 + 3

7.3.5.4. Ротор I ЭНИГМЫ *Вермахта* дает следующие вращаемые алфавиты (с кольцевой установкой A для $i = 0$):

установка кольца	i	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
A		0	E	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P	A	I	B	R	C	J
Z		1	K	F	L	N	G	M	H	E	R	W	A	O	U	P	X	Z	I	Y	V	T	Q	B	J	C	S	D
Y		2	E	L	G	M	O	H	N	I	F	S	X	B	P	V	Q	Y	A	J	Z	W	U	R	C	K	D	T
X		3	U	F	M	H	N	P	I	O	J	G	T	Y	C	Q	W	R	Z	B	K	A	X	V	S	D	L	E
W		4	F	V	G	N	I	O	Q	J	P	K	H	U	Z	D	R	X	S	A	C	L	B	Y	W	T	E	M
V		5	N	G	W	H	O	J	P	R	K	Q	L	I	V	A	E	S	Y	T	B	D	M	C	Z	X	U	F
⋮		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

7.4. Стандартные сдвигаемые алфавиты: Виженер и Бофорт

Используя обозначения разд. 7.2.1 и выбирая в (a′) $P=\rho$, получаем $\{\rho^i\rho\colon i\in\mathbb{N}\}$, множество горизонтально сдвигаемых стандартных алфавитов, которое совпадает (см. (a^0)) с множеством $\{\rho^i\colon i\in\mathbb{N}\}$ степеней стандартного алфавита и, аналогично (см. (a″)), с множеством вертикально продолженных стандартных алфавитов. Этот случай был исследован бенедиктинским аббатом Иоганнесом Гейденбергом из Триттенхейма на реке Мозель и был латинизирован Тритемием (1462–1516 гг.) в пятой книге его «Полиграфии» в стандартную упорядоченную таблицу («tabula recta», рис. 52, фр. *table réguliere*). Для механизации работы с нею можно воспользоваться диском

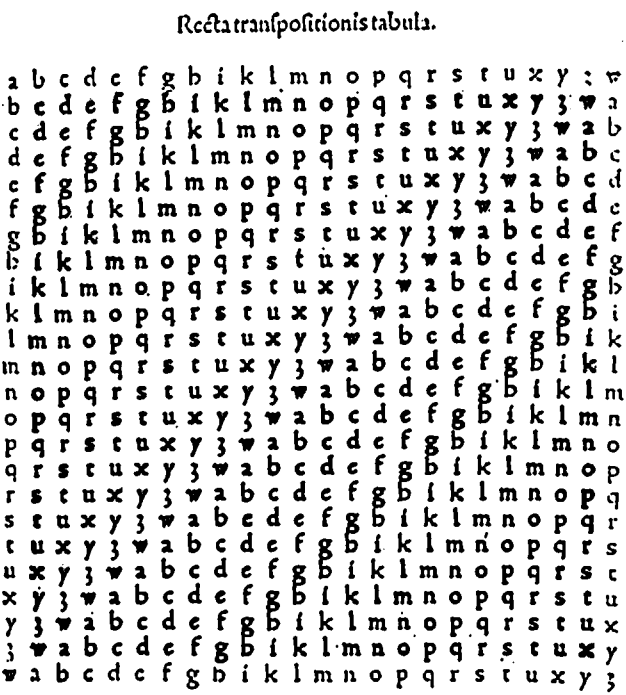


Рис. 52. «Tabula recta» Тритемия (оригинал в Мюнхенской городской библиотеке)

Альберти с подвижным окном, на внутренней окружности которого нанесен стандартный алфавит; а на внешнем диске располагается цикл ρ стандартного алфавита для формирования его степеней. В литературе говорят о шаге шифрования ВИЖЕНЕРа. Более уместно было бы дать ему имя Тритемия. Литература XIX в. была несправедлива к Виженеру, связав с его именем только сдвигаемые *стандартные* алфавиты, в то время как фактически его предложение не ограничивалось этим: Виженер записывал в заголовке таблицы *tabula recta* смешанный алфавит: очевидно, это было эквивалентно диску Альберти. Реализация множество $\{\rho^i\colon i\in\mathbb{N}\}$ электромеханическим способом возможна с помощью полу-ротора (рис. 48).

7.4.1. Виженер. Очевидно, шаг шифрования ВИЖЕНЕРа является линейной подстановкой: цикл ρ определяет линейный циклический квазипорядок на V^n и, таким образом, сложение *по модулю* N^n (гл. 5); множество алфавитов $\{\rho^i\colon i\in\mathbb{N}\}$ соответствует сложению чисел сдвига A_i (фр. *nombre de décalage*):

$$\{A_i\colon i\in\mathbb{Z}_{N^n}\}\quad\text{с}\quad A_i\colon A_i(x)\overset{N^n}{\simeq}x+i.$$

Шаг расшифрования $A_i^{-1}\colon A_i^{-1}(y)\overset{N^n}{\simeq}y-i$ представляет собой вычитание номера сдвига. Преобладающим является случай $n=1$. Такие «криптографические уравнения» использовались Чарльзом Бэббиджем приблизительно в 1846 г. (Британский музей, доп. реестр 37205, том 59), однако он не опубликовал их.

Шифрование ВИЖЕНЕРа появляется иногда в неявной форме: в 1913 г., когда Вильсон был президентом США, государственный департамент и армия США ввели вариант шифра Виженера под названием «Larrabee», используя двадцать шесть карточек: на каждой изображался стандартный алфавит открытого текста Z_{26} и алфавит шифротекста, получаемый прибавлением номера сдвига.

7.4.2. Эйрауд. Другое семейство сопутствующих шагов шифрования, содержащее умножение, предложено в разд. 5.6 (шаг шифрования ЭЙРАУДа):

$$\{C_q\colon q\in\mathbb{Z}_{N^n}\wedge\text{НОД}(q,N^n)=1\}\quad\text{с}\quad C_q\colon C_q(x)\overset{N^n}{\simeq}q\cdot x.$$

Шаг расшифрования (с q' таким, что $q\cdot q'\overset{N^n}{\simeq}1$), есть $C_q^{-1}\colon C_q^{-1}(x)\overset{N^n}{\simeq}q'\cdot x$.

Это семейство является семейством алфавитов с выбыванием (фр. *alphabets chevauchants*, Эйрауд; см. разд. 5.6), таковыми они являются лишь при $|q|\neq 1$.

Наиболее общей линейной подстановкой в кольце $\mathbb{Z}_{N^n}$ является суперпозиция шагов шифрования ВИЖЕНЕРа и ЭЙРАУДа[1]):

$$T_{q,i}\colon T_{q,i}(x)\overset{N^n}{\simeq}q\cdot x+i=A_i(C_q(x)).$$

[1])Эйрауд описал метод, порождающий это наиболее общее семейство: после того как получены два цикла символов открытого текста и символов шифротекста, для каждой допустимой пары q, q' выполняется следующее: один цикл записывается в q строках, другой — в q' строках; в обоих случаях циклы периодически продолжаются. Когда определена начальная пара, она определяет i. Читая построчно блок открытого текста и по столбцам

7.4.3. Бофорт. Случай фиксированного $q = q' = N^n - 1$ порождает семейство

$$\{B_i : i \in \mathbb{Z}_{N^n}\} \quad \text{с} \quad B_i : B_i(x) \overset{N^n}{\simeq} i - x.$$

В литературе в этом случае говорят о шифровании БОФОРТа, хотя оно изучалось уже Джованни Сестри в 1710 г. и было повторно открыто в 1857 г. адмиралом сэром Фрэнсисом Бофортом (1774–1857 гг.), который известен тем, что разработал шкалу скоростей ветра. Шаг расшифрования

$$B_i^{-1} : B_i^{-1}(y) \overset{N^n}{\simeq} i - y$$

совпадает с шагом шифрования, так что шаг шифрования БОФОРТа является взаимообратным:

x является неподвижной точкой B_i тогда и только тогда, когда $2 \cdot x \overset{N^n}{\simeq} i$.

В классических случаях шагов шифрования ВИЖЕНЕРа и БОФОРТа, естественно, $n = 1$. Де Виарис[2)] опубликовал в 1888 г. интерпретацию шагов шифрования ВИЖЕНЕРа и БОФОРТа как сложение и вычитания *по модулю* N, после того, как Керкхоффс в 1883 г. (не зная о штудиях Бэбби-

---

блок шифротекста, разместив их рядом, получим шаг шифрования; считывая по столбцам блок открытого текста и построчно блок шифротекста, получим шаг расшифрования.

Пример: $n = 1$, $N^n = 26$, $q = 3$, $q' = 9$; начальная пара: **c-P**.

цикл открытого текста e v i t z i s c o u r a n d b f g h j k m p q w x y
цикл шифротекста S E C U R I T Y A B D F G H J K L M N O **P** Q V W X Z

e	v	i	S	E	C	U	R	I	T	Y	A		e	v	i	S	E	C	U	R	I	T	Y	A
t	z	i	B	D	F	G	H	J	K	L	M		t	z	i	B	D	F	G	H	J	K	L	M
s	**c**	o	N	O	**P**	Q	V	W	X	Z	S		s	c	o	N	O	**P**	Q	V	W	X	Z	S
u	r	a	E	C	**U**	R	I	T	Y	A	B		u	r	a	E	C	U	R	I	T	Y	A	B
n	d	b	D	F	**G**	H	J	K	L	M	N		n	d	b	D	F	G	H	J	K	L	M	N
f	g	h	O	P	**Q**	V	W	X	Z	S	E		f	g	h	O	P	Q	V	W	X	Z	S	E
j	k	m	C	U	**R**	I	T	Y	A	B	D		j	k	m	C	U	R	I	T	Y	A	B	D
p	q	w	F	**G**	**H**	J	K	L	M	N	O		p	q	w	F	G	H	J	K	L	M	N	O
x	y	e	P	Q	**V**	W	X	Z	S	E	C		x	y	e	P	Q	V	W	X	Z	S	E	C
v	i	t	U	R	**I**	T	Y	A	B	D	F		v	i	t	U	R	I	T	Y	A	B	D	F
z	l	s	G	H	**J**	K	L	M	N	O	P		z	l	s	G	H	J	K	L	M	N	O	P
⋮			⋮										⋮			⋮								

шаг шифрования c o u r a n d b f g h j k m p q w x **y** e v i t z l s
P U G Q R H V I J W T K X Y L Z A M S B N E D O C F

шаг расшифрования P Q V W X Z S E C U R I T Y A B D F G H J K L M N O
c r d g k q **y** i l o a b h m w e t s u n f j p x v z

в циклической записи (a r q z o u g w)(b e i)(c p l)(d v n h t)(f j k x m y s).

[2)] Маркиз Гаэтан Генри Леон Виарис (1847–1901 гг.), французский криптолог. Де Виарис примерно в 1885 г. изобрел одну из первых печатающих шифровальных машин — согласно Кану, самые первые были изобретены, возможно, ранее 1874 г. Эмилем Виньи и Джозефом Госсеном.

джа) указал на математические соотношения между шагами ВИЖЕНЕРа и БОФОРТа.

При любой фиксированной позиции сдвига шаг шифрования ВИЖЕНЕРа превращается в шаг шифрования ЦЕЗАРЯ, в то время как шаг шифрования БОФОРТа превращается в шаг шифрования ЦЕЗАРЯ для перевернутого алфавита.

7.4.4. Обратный Виженер, обратный Бофорт. В английской литературе шаг шифрования «обратный ВИЖЕНЕР»

$$E_i\colon E_i(x) \overset{N^n}{\simeq} -i+x=-(i-x)$$

также носит имя «вариант БОФОРТа», а во Франции — «variante à l'alleniande». Он был предложен в 1858 г. Льюисом Кэрроллом и описан де Виарисом.

Тривиальный взаимообратный шаг шифрования «обратный БОФОРТ»

$$F_i\colon F_i(x) \overset{N^n}{\simeq} -i-x,$$

также описанный де Виарисом, был повторно рассмотрен в 1972 г. Оле Иммануэлем Франксеном.

«Вариант», приписываемый Каспаром Шоттом в «*Schola steganographia*» (1665 г.) графу Гронсфельду, является не более чем литеральным «Виженером», использующим только десять алфавитов, которые обозначались цифрами 0 ... 9. Криптографически это приносит одни только недостатки. Жюль Верн описал его в своем романе «*Гигантский плот*» (*The Giant Raft*) в 1881 г. В 1892 г. группа французских анархистов использовала его, и шифр был взломан Этьеном Базерье.

7.4.5. Порта. Семейство из одиннадцати взаимообратных подстановок ($V=Z_{22}$) использовалось уже в 1563 г. Джованни Баттистой Порта (рис. 53). Эти алфавиты называются сдвигаемыми взаимообратными алфавитами, они обозначались омофонно 22 ключевыми символами, расположенными попарно. Подобное расположение с десятью алфавитами ($V=Z_{20}$) использовалось в 1589 г. Дж. Б. и М. Ардженти (рис. 69). Для алфавита с четным числом $N=2\nu$ букв, существует ν таких подходящих взаимообратных алфавитов, мы назовем их шагами шифрования Порта.

7.5. Несвязанные алфавиты

Порта лишил себя славы изобретателя общей многоалфавитной подстановки, основанной на множестве из θ ($\theta\leqslant(N^n)!$) «взаимно несвязанных» смешанных алфавитов, т. е. алфавитов, которые не связаны друг с другом каким либо простым алгебраическим способом вроде сдвига или преобразования подобия (Кан: «Порядок букв в таблице может быть произвольным, если он обеспечивает присутствие всех букв»).

7.5.1. Перемешивание. Хотя Порта описал в своей «*De furtivis literdrum notis*» случай многократно перемешанных алфавитов (фр. *alphabete indépen-*

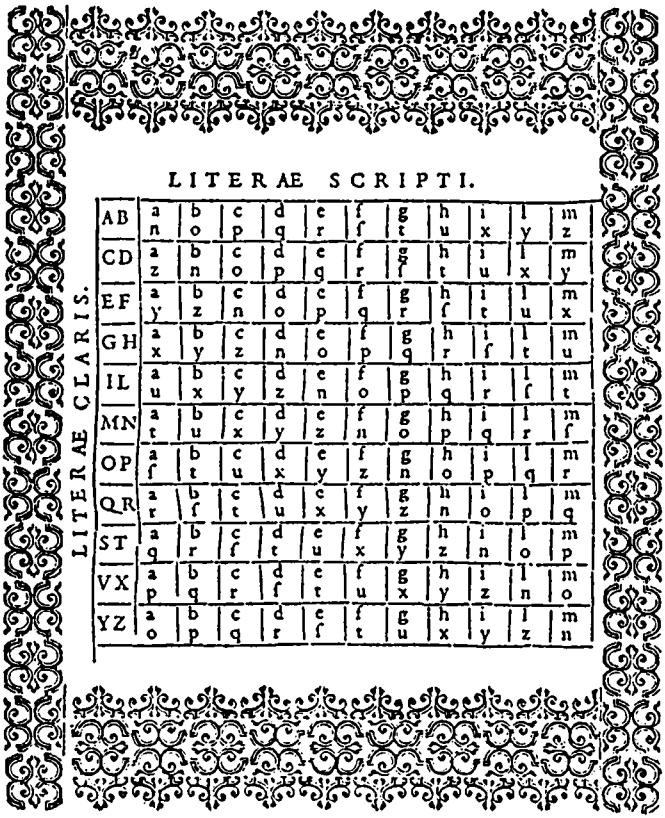


Рис. 53. Одиннадцать взаимообратных алфавитов для многоалфавитного шифрования (Порта, 1563)

dants, нем. *unabhängige Alphabete*), он не привел других примеров, кроме сдвигаемых взаимообратных алфавитов, наподобие алфавита, изображенного на рис. 53. Эйрауд склоняется к тому, чтобы отдать славу первопроходца в этой области исключительно французу Виженеру. Аналогично, Луиджи Сакко, автор превосходного Manuale di crittografia (3-е изд, Рим, 1947 г.), благоволит Италии (Кан: «пытаясь доказать, что все итальянское было первым»). Чарльз Дж. Мендельсон, которого нельзя обвинить в протекционизме, хвалит Порта как «выдающегося криптографа Возрождения». Когда мы имеем дело с наиболее общими перестановками, мы будем говорить о семействе шагов шифрования перемешиванием.

Таблица для общей многоалфавитной подстановки с несвязанными алфавитами могла бы походить на следующую (отметим правило построения ее из пароля, а именно

пароли служат для выбора метода из класса методов и ключей, в особенности, для выбора шифрования и...

passwords serve to select a method from a class of methods and keys especially to select encryptions e...)

	c	h	a	p	t	e	r	l	v	n	b	d	f	g	i	j	k	m	o	q	s	u	w	x	y	z
C	**P**	**A**	**S**	**W**	**O**	**R**	**D**	**E**	**V**	**T**	B	C	F	G	H	I	J	K	L	M	N	Q	U	X	Y	Z
R	N	P	Q	R	U	V	W	X	Y	Z	**O**	**S**	**E**	**L**	**C**	**T**	**A**	**M**	**H**	**D**	B	F	G	I	J	K
Y	**L**	**S**	**E**	**T**	B	D	G	H	I	J	K	N	P	Q	U	V	W	X	Y	Z	**F**	**R**	**O**	**M**	**A**	**C**
P	V	W	X	Z	**H**	**O**	**D**	**S**	**A**	**N**	**K**	**E**	**Y**	**P**	B	C	F	G	I	J	L	M	Q	R	T	U
T	F	G	H	J	K	M	P	Q	R	U	V	W	X	Z	**E**	**C**	**I**	**A**	**L**	**Y**	**T**	**O**	**S**	**N**	B	D
O	**Y**	**P**	**T**	**I**	**O**	**N**	**S**	**E**	A	B	D	F	G	H	J	K	L	M	Q	U	V	W	X	Z	**C**	**R**
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

7.5.2. Грипенштерн. Общая многоалфавитная подстановка с несвязанными алфавитами используется в небольшом известном шифровальном устройстве 1786 г., сконструированном шведским бароном Фредриком Грипенштерном (1728–1804 гг.) для шведского короля Густава III (рис. 54), и восстановленном фирмой Crypto AG (Швейцария, Цуг) по документам, обнаруженным Свеном Вастрёмом в городском архиве Стокгольма. Устройство имело 57 дисков, используемых для выполнения различных (фиксированных) подстановок $Z_{26} \to Z_{10}^2$. Рассматриваемая как многобуквенная подстановка длины 57, перестановка дисков давала семейство из огромного числа $57! \approx 4.05 \cdot 10^{76}$ алфавитов. Даже при использовании устройства с неизменным порядком дисков для сообщения в несколько сотен символов или для нескольких таких сообщений эта система шифрования превосходила все, что применялось еще где-либо в то время.

Рис. 54. Шифровальное устройство барона Фредрика Грипенштерна (Crypto AG)

Аналогично, в 1799 г. римский католический священник Иоганн Баптист Андрес (1770–1823 гг.) описал использование таблицы с 26 несвязанными смешанными алфавитами, выбираемыми периодически в соответствии с ключом.

В 1915 г. шведский изобретатель Арвид Дамм придумал устройство, отчасти соответствующее этому описанию, используя множество сменных полос с несвязанными смешанными алфавитами, размещенными на барабане параллельно его оси (A-21: рис. 55). Рядом с полосами находилась прямая линейка для алфавита открытого текста; после каждого шага шифрования барабан смещался на один шаг. Линейка с алфавитом открытого текста могла находиться в двух положениях, которые сменялись достаточно часто.

Рис. 55. Шифратор A-21 (1915) Арвида Дамма (AB Cryptograph, Стокгольм)

Многоалфавитное шифрование с несвязанными алфавитами использовалось в Первую мировую войну немцами для передачи радиосообщений диверсионной группе в Северной Африке (система «für GOD»), а во Вторую мировую войну ВВС США для связи «земля-воздух» (SYKO). В обоих случаях это было ненадежно. SYKO состоял из тридцати взаимообратных алфавитов, напечатанных на карточках: та же самая старая идея «Larrabee» (разд. 7.4.1). Алфавиты использовались в циклическом порядке — шифровальщик, используя указатель, определял начальный алфавит на целый день. Это было слишком долго и было причиной слабости в целом хорошей системы.

Томас Джефферсон (1743–1826)

7.5.3. Мультиплексность. Циклические перестановки как частный случай многоалфавитного шифрования с несвязанными алфавитами, имели классическую реализацию в виде специального шифровального цилиндра, использованного Джефферсоном (между 1790 г. и 1800 г.) и повторно изобретенного в 1891 г. Базерье. Циклические подстановки представлялись одна за другой как цикл на ободе тонкого цилиндра (фр. *rondelle*). Джефферсон упорядочивал 36 таких цилиндров (каждый с перемешанным алфавитом Z_{26}) в длинный ци-

1	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	z
2	b	c	d	f	g	h	j	k	l	m	n	p	q	r	s	t	v	x	z	a	e	i	o	u	y
3	a	e	b	c	d	f	g	h	i	o	j	k	l	m	n	p	u	y	q	r	s	t	v	x	z
4	z	y	x	v	u	t	s	r	q	p	o	n	m	l	k	j	i	h	g	f	e	d	c	b	a
5	y	u	z	x	v	t	s	r	o	i	q	p	n	m	l	k	e	a	j	h	g	f	d	c	b
6	z	x	v	t	s	r	q	p	n	m	l	k	j	h	g	f	d	c	b	y	u	o	i	e	a
7	a	l	o	n	s	e	f	t	d	p	r	i	j	u	g	v	b	c	h	k	m	q	x	y	z
8	b	i	e	n	h	u	r	x	l	s	p	a	v	d	t	o	y	m	c	f	g	j	k	q	z
9	c	h	a	r	y	b	d	e	t	s	l	f	g	i	j	k	m	n	o	p	q	u	v	x	z
10	d	i	e	u	p	r	o	t	g	l	a	f	n	c	b	h	j	k	m	q	s	v	x	y	z
11	e	v	i	t	z	l	s	c	o	n	r	a	n	d	b	f	g	h	j	k	m	p	q	x	y
12	f	o	r	m	e	z	l	s	a	i	c	u	x	b	d	g	h	j	k	n	p	q	t	v	y
13	g	l	o	i	r	e	m	t	d	n	s	a	u	x	b	c	f	h	j	k	p	q	v	y	z
14	h	o	n	e	u	r	t	p	a	i	b	c	d	f	g	j	k	l	m	q	s	v	x	y	z
15	i	n	s	t	r	u	e	z	l	a	i	b	c	d	f	g	h	k	m	o	p	q	v	x	y
16	j	a	i	m	e	l	o	g	n	f	r	t	h	u	b	c	d	k	p	q	s	v	x	y	z
17	k	y	r	i	e	l	s	o	n	a	b	c	d	f	g	h	j	m	p	q	t	u	v	x	z
18	l	h	o	m	e	p	r	s	t	d	i	u	a	b	c	f	g	j	k	n	q	v	x	y	z
19	m	o	n	t	e	z	a	c	h	v	l	b	d	f	g	i	j	k	p	q	r	s	u	x	y
20	n	o	u	s	t	e	l	a	c	f	b	d	g	h	i	j	k	m	p	q	r	v	x	y	z

Рис. 56. Двадцать циклов Базерье

линдр; Базерье использовал 20 цилиндров (каждый с перемешанным алфавитом Z_{25}), как показано на рис. 19. Фридман назвал эти семейства несвязанных циклов «мультиплексными системами».

Четырнадцать циклов из двадцати, использованных Базерье (они показаны на рис. 56), определялись началами причудливых стихов, наподобие паролей:

Allons enfants de la patrie, le jour de gloire est arrivé
Bienheureux les pauvres d'esprit, le royaume des Cieux
Charybde et Scilla
Dieu protège la France
Évitez les courants d'air
Formez les faisceaux
Gloire immortelle de nos a ... eux
Honneur et Patrie
Instruisez la jeunesse
J'aime l'oignon frit à l'huile
Kyrie eleison
L'homme propose et Dieu dispose
Montez à cheval
Nous tenons la clef

Базерье не смог убедить французский генеральный штаб (*état-major général*) принять его изобретение — де Виарис (разд. 14.3.1) показал, как взла-

мывать сообщения, зашифрованные цилиндром при известных алфавитах — реалистичном предположении для военных в боевой обстановке (разд. 11.2.3). Очевидно, Базерье не знал, что Джефферсон намного раньше его высказал ту же самую идею, и наиболее вероятно — он умер в 1931 г. в возрасте 85 лет — он не узнал о запоздалом принятии его предложения в 1922 г. армией США. Устройство с тринадцатью цилиндрами было предложено в 1900 г. итальянским полковником Оливье Дукросом.

Цилиндры Джефферсона и Базерье позволяли считывать зашифрованный текст (рис. 21) не только в следующей, но и в произвольно выбранной i-й строке («i-й образующей»). Таким образом, шифрование было полифоническим. Получатель, установив шифротекст, просто искал строку, которая бросалась в глаза. Для неуполномоченного же дешифрования, эта сложность была не столь велика, как наивно можно было бы предполагать (разд. 14.3.1).

Обычно порядок цилиндров оставался неизменным для одного или нескольких сообщений, или в течение определенного периода времени, например, дня.

Вместо цилиндров можно использовать ленты с продублированным алфавитом. Такое устройство шифрования было предложено в 1893 г. французом Артуром Ж. Эрманом и распространено в 1914 г. капитаном, позже полковником Паркером Хиттом со ссылкой на Базерье. Хитт поначалу не имел никакого успеха. Тем временем, в 1917 г. морской лейтенант Рассел Вильсон также изобрел ленточное устройство NCB (шифратор Военно-морского флота), которое использовалось в ВМФ США по крайней мере до 1935 г. Армия США в конце концов стала использовать цилиндр Джефферсона; известная шифровальная машина M-94 была принята на вооружение в 1922 г. под влиянием Фридмана после существенных улучшений, сделанных под руководством полковника Мауборна[3)], тогдашнего главы отдела исследований и технического обслуживания корпуса связи. Она имела 25 тонких алюминиевых цилиндров размером с серебряный доллар, вращающихся на шпинделе длиной 110 мм. В 1943 г. M-94 (вклейка D, рис. 57) была объявлена устаревшей. В это время в достаточном количестве появились шифраторы M-209s (разд. 4.4.8).

Рис. 57. Шифровальщик, собирающий M-94

[3)] Это был тот самый Мауборн, который в 1918 г. усовершенствовал последовательное шифрование Вернама путем введения неограниченных и хаотичных ключей («одноразовых ключей»). См. 8.8.2.

Дж. Вернам

Паркер Хитт

Дж. Мауборн

В 1934 г. был принят на вооружение шифратор М-138, ленточная версия; он располагал сотней лент, из которых одновременно использовались тридцать. Улучшенная модель М-138-А с 1938 г. служила армейским офицерам и дипломатам. Шифратор считался настолько безопасным по отношению к взлому, что радиосообщения Рузвельта Черчиллю сразу после Атлантической конференции посылались зашифрованными М-138-А. В то время как японцы, по-видимому, не обладали достаточной квалификацией для взлома американского ленточного шифра, немцы сделали это: Ганс Рорбах в 1944 г. (разд. 14.3.6) взломал его, не имея доступа к алфавитам. (Его успех длился недолго, вскоре США внесли изменения в устройства SIGTOT, машины типа Вернама.) На вклейке Е показана модификация шифратора М-138-Т4 с 25 лентами, используемыми одновременно.

Система «О-2» Госдепартамента США, которую взломал Рорбах, из имеющихся пятидесяти лент одновременно использовала тридцать, которые были разбиты на две группы по пятнадцать лент. Здесь был риск: общее количество имеющихся лент должно значительно превосходить период, т. е. число одновременно используемых лент.

Между прочим, ВМФ США использовал в качестве дополнительного шифратора устройство шифрования CSP 642, также работавшее с тридцатью лентами. Японцы захватили несколько из таких устройств и предприняли огромные усилия по взлому сообщений, успеха они не добились — возможно они не изучали методов де Виариса и Фридмана (разд. 14.3).

Для цилиндровых и ленточных шифровальных устройств мы будем следовать Фридману и называть их шаг шифрования МУЛЬТИПЛЕКСНЫМ. Эти шаги специфичны — они являются полностью циклическими перемешанными шагами шифрования.

7.5.4. Свойство «латинского квадрата». В частном случае $\theta \leqslant N$ можно потребовать, чтобы N перемешиваемых алфавитов, каждый из которых состоит из N символов, записанных построчно, имели бы следующее свойство: ни в каком столбце символы не встречаются более одного раза (Эйрауд: «алфавиты действительно не параллельны»). В случае $\theta = N$ такие алфавиты образуют «латинский квадрат», т. е. в каждом столбце каждый символ встречается только один раз. Это требование упоминалось уже в *Geheimschreibekunst* Иоганна Баптиста Андреса в 1799 г. (см. разд. 7.5.2), который также дал пример своей таблицы. *Tabula recta* тривиально удовлетворяет этому требованию, но ее алфавиты не несвязаны.

Это требование может быть постулировано для перемешанных алфавитов, принадлежащих циклам мультиплексной системы (разд. 7.5.3), предотвращая

тем самым атаку де Виариса (разд. 14.3.1). Циклы, получаемые из мнемонических паролей, вряд ли можно квалифицировать подобным образом; на самом деле, перемешанные алфавиты, принадлежащие к циклам Базерье (табл. 2 а), демонстрируют своеобразный результат: большинство столбцов имеет один или два символа, встречающихся часто. Ясно, что отсутствие многих букв облегчает взлом шифра. Алфавиты, принадлежащие циклам Базерье не могут порождать латинский квадрат.

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	z
1	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	z	a
2	e	c	d	f	i	g	h	j	o	k	l	m	n	p	u	q	r	s	t	v	y	x	z	b	a
3	e	c	d	f	b	g	h	i	o	k	l	m	n	p	j	u	r	s	t	v	y	x	z	q	a
4	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y
5	j	y	b	c	a	d	f	g	q	h	e	k	l	m	i	n	p	o	r	s	z	t	v	u	x
6	z	y	b	c	a	d	f	g	e	h	j	k	l	m	i	n	p	q	r	s	o	t	v	u	x
7	l	c	h	p	f	t	v	k	j	u	m	o	q	s	n	r	x	i	e	d	g	b	y	z	a
8	v	i	f	t	n	g	j	u	e	k	q	s	c	h	y	a	z	x	p	o	r	d	l	m	b
9	r	d	h	e	t	g	i	a	j	k	m	f	n	o	p	q	u	y	l	s	v	x	z	b	c
10	f	h	b	i	u	n	l	j	e	k	m	a	q	c	t	r	s	o	v	g	p	x	y	z	d
11	п	f	o	b	v	g	h	j	t	k	m	s	p	d	u	q	x	a	c	z	r	i	y	e	l
12	i	d	u	g	z	o	h	j	c	k	n	s	e	p	r	q	t	m	a	v	x	y	b	f	l
13	u	c	f	n	m	h	l	j	r	k	p	o	t	s	i	q	v	e	a	d	x	y	b	z	g
14	i	c	d	f	u	g	j	o	b	k	l	m	q	e	n	a	s	t	v	p	r	x	y	z	h
15	j	c	d	f	z	g	h	k	n	b	m	a	o	s	p	q	v	u	t	r	e	x	y	i	l
16	i	c	d	k	l	r	n	u	in	a	p	o	e	f	g	q	s	t	v	h	b	x	y	z	j
17	b	c	d	f	l	g	h	j	e	m	y	s	p	a	n	q	t	i	o	u	v	x	z	r	k
18	b	c	f	i	p	g	j	o	u	k	n	h	e	q	m	r	v	s	t	d	a	x	y	z	l
19	c	d	h	f	z	g	i	v	j	k	p	b	o	t	n	q	r	s	u	e	x	l	y	m	a
20	c	d	f	g	l	b	h	i	j	k	m	a	p	o	u	q	r	v	t	e	s	x	y	z	n

Таблица 2 а. Двадцать перемешанных алфавитов, соответствующих циклам Базерье

Обычно в первой строке и первом столбце латинского квадрата расставляется стандартный алфавит из N символов. Тогда при $N = 2$ и $N = 3$ имеются только тривиальные решения для таблицы «*tabula recta*». При $N = 4$ кроме «*tabula recta*», имеется еще три из таких «приведенных» латинских квадратов:

a b c d	a b c d	a b c d	a b c d
b c d a	b d a c	b a d c	b a d c
c d a b	c a d b	c d b a	c d a b
d a b c	d c b a	d c a b	d c b a

Количество латинских квадратов быстро растет с ростом N: 56 при $N = 5$, 9408 при $N = 6$, 16 942 080 при $N = 7$, 535 281 401 856 при $N = 8$. При $N = 9$ имеется уже 377 597 570 964 258 816 приведенных латинских квадратов, как подсчитали С. Э. Баммель и Дж. Ротштейн в 1975 г., в то время как общее

количество всех приведенных квадратов с девятью алфавитами из девяти символов составляет $(8!)^8 \approx 6.98 \cdot 10^{36}$.

Приведем два примера латинских квадратов для $N = 10$ с алфавитом $Z_{10} = \{0, 1, 2, \ldots, 9\}$:

0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9
1 5 7 2 8 9 0 3 4 6	1 4 3 2 0 9 8 5 6 7
2 4 6 1 3 8 9 0 5 7	2 6 5 4 3 0 9 8 7 1
3 0 5 7 2 4 8 9 6 1	3 8 7 6 5 4 0 9 1 2
4 9 0 6 1 3 5 8 7 2	4 9 8 1 7 6 5 0 2 3
5 8 9 0 7 2 4 6 1 3	5 0 9 8 2 1 7 6 3 4
6 7 8 9 0 1 3 5 2 4	6 7 0 9 8 3 2 1 4 5
7 6 1 8 9 0 2 4 3 5	7 2 1 0 9 8 4 3 5 6
8 3 4 5 6 7 1 2 9 0	8 3 4 5 6 7 1 2 9 0
9 2 3 4 5 6 7 1 0 8	9 5 6 7 1 2 3 4 0 8

В таблице 2b показаны алфавиты, принадлежащие 25 циклам шифровальной машины M-94; они образуют почти латинский квадрат для $N = 26$.

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
1	b	c	e	j	i	v	d	t	g	f	z	r	h	a	l	w	k	x	p	q	y	u	n	s	m	o
2	c	a	d	e	h	i	z	f	j	k	t	m	o	p	u	q	x	w	b	l	v	y	s	r	g	n
3	d	g	z	k	p	y	e	s	n	u	o	a	j	x	m	h	r	t	c	v	b	w	l	f	q	i
4	e	i	b	c	d	g	j	l	f	h	m	k	r	w	q	t	v	u	a	n	o	p	y	z	x	s
5	f	r	y	o	m	n	a	c	t	b	d	w	z	q	p	i	u	h	l	j	k	x	e	g	s	v
6	g	j	i	y	t	k	p	w	x	s	v	u	e	d	c	o	f	n	q	a	r	m	b	l	z	h
7	h	n	f	u	z	m	s	x	k	e	p	c	q	i	g	v	t	o	y	w	l	r	a	j	d	b
8	i	w	v	x	r	z	t	p	h	o	c	q	g	s	b	j	e	y	u	d	m	f	k	a	n	l
9	j	x	r	s	f	h	y	g	v	d	q	p	b	l	i	m	o	a	k	z	n	t	c	w	u	e
10	k	d	a	f	l	j	h	o	c	g	e	b	t	m	n	r	s	q	v	p	x	z	i	y	w	u
11	l	e	g	i	j	b	k	u	z	a	r	t	s	o	h	n	p	f	x	m	w	q	d	v	c	y
12	m	y	u	v	w	l	c	q	s	t	x	h	n	f	a	z	g	d	r	b	j	e	o	i	p	k
13	n	m	j	h	a	e	x	b	l	i	g	d	k	c	r	f	y	p	w	s	z	o	q	u	v	t
14	o	l	t	w	g	a	n	z	u	v	j	e	f	y	d	k	h	s	m	x	q	i	p	b	r	c
15	p	v	x	r	n	q	u	i	y	z	s	j	a	t	w	b	d	l	g	c	e	h	f	o	k	m
16	q	t	s	**e**	o	p	i	d	m	n	f	x	w	u	k	y	j	v	h	g	**b**	l	z	c	**a**	r
17	r	k	w	p	u	t	q	e	b	x	l	n	y	v	f	c	i	m	z	h	s	a	g	d	o	j
18	s	o	n	m	q	u	v	a	w	r	y	g	c	e	z	l	b	k	d	f	i	j	x	h	t	p
19	t	s	m	z	k	x	w	v	r	y	u	f	i	g	j	d	a	b	e	o	p	c	h	n	l	q
20	u	p	k	g	s	c	f	j	o	w	a	y	d	h	v	e	l	z	n	r	t	b	m	q	i	x
21	v	f	l	q	y	s	o	r	p	m	h	z	u	k	x	a	c	g	j	i	d	n	t	e	b	w
22	w	h	o	l	b	d	m	k	e	q	n	i	x	r	t	u	z	j	f	y	c	s	v	p	a	g
23	x	z	p	t	v	o	b	m	q	c	w	s	l	j	y	g	n	e	i	u	f	d	r	k	h	a
24	y	q	h	a	c	r	l	n	d	p	b	o	v	z	s	x	w	i	t	e	g	k	u	m	j	f
25	z	u	q	n	x	w	r	y	a	l	i	v	p	b	e	s	m	c	o	k	h	g	j	t	f	d

Таблица 2b. Почти латинский квадрат, образуемый алфавитами Мауборна для M-94. Циклическая перестановка из трех символов, выделенных жирным шрифтом в строке 16, превращает его в правильный латинский квадрат

Для чего Майборн предусмотрел три исключения в алфавите строки 16, неизвестно.

Заметим также, что вращаемые алфавиты (разд. 7.2.2) — в отличие от сдвинутых алфавитов (разд. 7.2.1) — обычно не образуют латинский квадрат.

Простые арифметические формулы для числа $l(N)$ приведенных латинских квадратов из N строк и столбцов были получены недавно. Эрдёш (1913–1996 гг.) и Каплански в 1946 г. предположили, что

$$l(N) \asymp N \cdot (N!)^{N-2}/e^{N\cdot(N-2)/2} \qquad (l(9) < 1.73 \cdot 10^{24}).$$

Для $N \leqslant 9$ хорошая эмпирическая верхняя граница задается неравенством

$$l(N) \leqslant \sqrt{((N-1)!)^{N-1}} \qquad (l(9) < 2.64 \cdot 10^{18}).$$

Имеется очень грубая нижняя граница (Хейзе):

$$l(N) \geqslant 2! \cdot 3! \cdot 4! \cdot \ldots \cdot (N-2)! \qquad (l(9) > 1.25 \cdot 10^{11}).$$

Заметим, что $l(9) \approx 3.78 \cdot 10^{17}$. Для $l(26)$ указанные оценки дают границы $10^{243} < l(26) < 10^{499}$. Имеются обоснованные надежды на то, что истинное значение ближе к верхней границе.

Глава 8

Многоалфавитное шифрование: ключи

Никакое сообщение не является надежно зашифрованным, если ключевая последовательность не сравнима по длине с самим сообщением.

Паркер Хитт, 1914 *г.*

8.1. Ранние методы с периодическими ключами

8.1.1. Альберти. Самое раннее упоминание о многоалфавитном шифровании можно найти в 25-страничном трактате «*De cifris*» Леона Баттиста Альберти (1404–1472 гг.), который он написал в 1466 или 1467 гг. для своего друга Леонардо Дато, папского секретаря. Латинский оригинал трактата воспроизвел Алоис Мейстер в работе «*Die Geheimschrift im Dienste der Papstlichen Kurie*» (стр. 125–141), Падеборн, Шенинг, 1906 г. (итальянский перевод рукописи «*Trattati in cifra*», датируемой приблизительно 1470 г.). Альберти был не только архитектором, живописцем, композитором и органистом, но также и крупным филологом эпохи Возрождения. Он мог взломать шифр простой замены и знал, как избежать этого. Альберти предложил изменять алфавит подстановки после каждых трех или четырех слов, «вводя новые значения для букв шифра». Для этой цели он изобрел специальное устройство — поворачиваемый шифровальный диск (рис. 26), который обеспечивал достаточно большое количество алфавитов подстановки. На три-четыре слова приходилось в среднем 18 символов. Таким образом, Альберти неосознанно придерживался минимального расстояния единственности Шеннона (разд. 12.6) для простой подстановки. Это было большим прогрессом по сравнению с типичным для того времени использованием омофонов: если в омофонной простой подстановке $Z_{25} \rightarrow Z_{10}^2$ биграммы 89, 43, 57 и 64 могли бы обозначать символ /a/, то теперь на эту роль годится любая биграмма. Конечно, отправитель и получатель должны были иметь идентичные диски.

В последующей литературе авторы расходятся в том, как выполнялась замена алфавита с помощью диска Альберти. Криптологи итальянец Сакко и

француз Эйрауд дают следующее объяснение: определенная буква, например, /b/ объявляется указателем (фр. *index*). Отправитель перед каждой частью текста, которая должна быть зашифрована с новым алфавитом, вставляет произвольную цифру от 1 до 4. При каждом прохождении цифры определяется новая позиция поворачиваемого диска путем совмещения указателя с соответствующим шифросимволом для этой цифры. Получатель знает, что при появлении цифры в расшифрованном открытом тексте он должен повернуть диск и выставить символ шифротекста напротив указателя. Заметим, что эта процедура является первым образцом скрытной ключевой связи, получившей важное значение в современных шифровальных машинах.

Кан видит главное использование цифр в диске Альберти в перешифровании введенного им же кватернарного кода; коды 336 двух-, трех- и четырехзначных групп из цифр 1 ... 4 вставлялись в буквенный текст. Альберти упоминал также об упорядочении кода для шифрования в словах и для расшифрования групп. Был ли этот ранний двухчастный код полностью перемешанным, как это было сделано Россиньолем для повышения защищенности, неясно.

Введя многоалфавитное шифрование и перешифрование кода, Альберти вполне заслужил право называться отцом современной криптологии и это не будет проявлением непочтительности к архитектору, построившему Палаццо Питти, создателю церквей св. Андрея в Мантуе, св. Марии Новелле во Флоренции и Темпио Малатестьяна в Римини. Эти архитектурные работы прославили его имя, а его вклад в криптологию на долгое время был предан забвению.

8.1.2. Тритемий. В то время как Альберти изменял алфавит после каждых трех-четырех слов, Тритемий в пятом томе своей «*Полиграфии*» (1508–1518 гг.) предложил после каждого символа переходить к следующему алфавиту. Однако делал это он согласно регулярному, периодическому — строка за строкой — прохождению своей *tabula recta*. Таким образом, его метод был заметно слабее метода Альберти.

С другой стороны, у него использовались все имеющиеся алфавиты прежде, чем какой-либо алфавит появится повторно. Следуя Кану, подобную систему называют шифросистемой с «последовательным ключом» — чтобы не путать с выражением «бегущий ключ» (разд. 2.3.3), введенным Фридманом. Последовательное шифрование часто используется в современных шифровальных машинах, однако при этом применяется гораздо больше пары дюжин алфавитов. Подробнее этот вопрос обсуждается в разд. 8.4.3.

Шифрование Тритемия было фиксированной шифросистемой с периодом 24; его можно рассматривать как одноалфавитное многосимвольное шифрование ширины 24. Трудно было поверить, что такой метод будет профессионально использоваться в XX в. Фактически, он применялся даже с периодом 3 (с последующей простой столбцовой перестановкой) в 1914 г. немцами на западном фронте. Французы называли его шифром *ABC* (в современной терминологии периодическое шифрование ВИЖЕНЕРа со специальным ключом *ABC*, и, конечно же, любили его (разд. 2.1.1).

Порта в 1602 г. указал возможную атаку на шифрование Тритемия: если открытый текст содержит три последовательные буквы алфавита, наподобие *pon* в *pondus*, то они дают тройную букву в шифротексте.

Ни Альберти, ни Тритемий не использовали ключевые слова в многоалфавитном шифровании. Джованни Баттиста Белазо был первый (1553 г.), кто обозначал последовательные шаги шифрования символами A, B, ..., Z и использовал ключевую последовательность для выбора алфавитов — путем поворота шифровального диска или выбора строки из таблицы. Его ключевая последовательность состояла из довольно длинных фраз наподобие *OPTARE MELIORA* и *VIRTUTI OMNIA PARENT*, в случае необходимости, повторяемых. Такой повторенный ключ ведет к периодическому многоалфавитному (односимвольному) шифрованию (разд. 2.3.3). Если используется только стандартный алфавит, то комбинаторная сложность Z этого метода равна N^d для ключей из d символов, а для многосимвольного случая n она равна $Z = (N^n)^d = N^{nd}$.

8.2. «Двойной ключ»

8.2.1. Порта. В 1563 г., через десять лет после того как Белазо предложил использовать ключ, применение ключа было объединено с применением перемешанного алфавита и алфавитов, получаемых из него сдвигами. Этот решительный шаг сделал 28-летний Джованни Баттиста Порта (1535–1615 гг.) в своей книге «*De furtivis literarum notis*», которую Кан охарактеризовал как «выдающуюся, свежую, обладающую шармом и способностью научить».

Так как пароль, определяющий перемешанный алфавит, уже назывался ключом, стало общепринятым говорить о «двойном шифре» (во французской терминологии, действующей и по сей день, *substitution à double clef*), но сегодня это уже устарело. Кан пишет: «Живарж даже тогда [1920-е гг.] называл многоалфавитные системы почти путающим термином «двойная подстановка», который абсолютно ничего не говорит о системе». Живарж говорит о *clef principale* применительно к термину «актуальный ключ». Пароль он иногда называет «вторым ключом».

Комбинаторная сложность этого метода для ключей из d символов и с произвольным перемешанным алфавитом в случае простого (односимвольного) шифрования равна $N! \cdot N^{d-1}$, для многосимвольного случая ширины n она равна $(N^n)! \cdot (N^n)^{d-1}$.

Вместо диска можно использовать таблицу. Для диска Альберти, изображенного на рис. 26 (с $\{\rho^{-i}P\colon i \in \mathbb{N}\}$), она будет выглядеть так:

	a	b	c	d	e	f	g	i	l	m	n	o	p	q	r	s	t	v	x	z	1	2	3	4
0	D	L	G	A	Z	E	N	B	O	S	F	C	H	T	Y	Q	I	X	K	V	P	&	M	R
1	R	D	L	G	A	Z	E	N	B	O	S	F	C	H	T	Y	Q	I	X	K	V	P	&	M
2	M	R	D	L	G	A	Z	E	N	B	O	S	F	C	H	T	Y	Q	I	X	K	V	P	&
3	&	M	R	D	L	G	A	Z	E	N	B	O	S	F	C	H	T	Y	Q	I	X	K	V	P
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Шаг расшифрования получится очевидным образом, если поместить перемешанный алфавит символов шифротекста в строку $i = 0$ наверху таблицы *tabula recta* символов открытого текста; в нашем примере с $\{P^{-1}\rho^i : i \in \mathbb{N}\}$

	D	L	G	A	Z	E	N	B	O	S	F	C	H	T	Y	Q	I	X	K	V	P	&	M	R
0	a	b	c	d	e	f	g	i	l	m	n	o	p	q	r	s	t	v	x	z	1	2	3	4
1	b	c	d	e	f	g	i	l	m	n	o	p	q	r	s	t	v	x	z	1	2	3	4	a
2	c	d	e	f	g	i	l	m	n	o	p	q	r	s	t	v	x	z	1	2	3	4	a	b
3	d	e	f	g	i	l	m	n	o	p	q	r	s	t	v	x	z	1	2	3	4	a	b	c
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

8.2.2. Виженер. Имеется связь *tabula recta* с произвольной подстановкой (см. разд. 7.4), предложенной в 1585 г. Виженером. Предложив использование ключей, он получил всю мощь шагов шифрования Альберти. Заметим также, что Виженер осознавал важность выбора достаточно длинных ключевых слов для повышения трудности криптоанализа.

Блез де Виженер родился 5 апреля 1523 г. в Сент-Пурси, «на полпути между Парижем и Марселем», как с американским простодушием пишет Кан. Он принимал участие в заседании Рейхстага в Вормсе в качестве секретаря, а последующие скитания по Европе с дипломатическими миссиями расширили его опыт. Остаток своей жизни Виженер служил у герцога Неверского. Он читал Тритемия, Белазо, Кардано и Порта и держал в руках рукопись Альберти. С 1570 г. в возрасте 47 лет он полностью сконцентрировался на писательстве; до своей смерти в 1596 г. он писал обо всем на свете, даже о кометах (*Traicté des Comètes*). В 1570 г. Виженер женился на Мери Варэ, которая была намного моложе его. В 1585 г. в возрасте 62 лет он написал свой *Traicté des Chiffres*, «несмотря на отвлечение на годовалую дочку» — пишет Кан. Книга по криптологии имела более 600 страниц и ее содержание выходило далеко за пределы криптографии: здесь были японские идеограммы, алхимия, магия, каббалистика, рецепты создания золота, но было также и достоверное, точное отражение состояния криптологии того времени. Обсуждая многоалфавитное шифрование, он следовал Альберти и Тритемию в использовании алфавитов, получаемых сдвигами, и помечал строки символами ключа, как делали это Белазо и Порта для своих взаимообратных алфавитов. В целом, он придал простой многоалфавитной подстановке ее современную форму.

8.2.3. «Тройной ключ» (фр. *triple clef*) получается, если два первичных алфавита объединяются с помощью итераций ключевых подстановок, например, если для данных множеств P_1, P_2 мы применим случай (а) из разд. 7.2.1, т. е. рассмотрим множество алфавитов $\{P_1\rho^i P_2 : i \in \mathbb{N}\}$ (разд. 19.5.3). Виженер пришел к тройному ключу, обозначив шаги шифрования со смешанным алфавитом, называемые сейчас его именем, символами ключа.

8.3. Шифрование Вернама

Современные каналы связи работают с двоичным алфавитом $Z_2 = \{O, L\}$ или $\mathbb{Z}_2 = \{0, 1\}$. Шифрование символов международного телетайпного кода

CCITT № 2 можно рассматривать как многосимвольное двоичное кодирование с $N = 2$ и $n = 5$; для шифрования байтов, т. е. 8-битовых символов, которые в современных компьютерах являются базисными элементами для представления данных, мы сталкиваемся со случаем $N = 2$ и $n = 8$; для блоков в 8 байтов имеем $N = 2$ и $n = 64$.

Нетрудно видеть, что для $\mathbb{Z}_2^5$ существует 32 различных шага шифрования Виженера, а для $\mathbb{Z}_2^8$ — 256; для выполнения шифрования как сложения по модулю 32 или по модулю 256 требуется циклический сумматор с 5 или 8 битами. Соответствующая двоичная электрическая схема (для $n = 5$) показана на рис. 58. Мощные современные микропроцессоры выполняют 64-разрядное сложение и могут непосредственно зашифровывать октограммы байтов.

8.3.1. Поразрядное шифрование. С другой стороны, шаг шифрования Виженера может быть выполнен последовательно бит за битом. Этот предельный случай поразрядного шифрования станет особенно важным позже. Если поразрядное двоичное шифрование $\mathbb{Z}_2 \to \mathbb{Z}_2$ рассматривать как отображение, то оно является перестановкой двух символов 0 и 1;

$$\text{тождественная подстановка } O\colon \begin{matrix} 0 \mapsto 0 \\ 1 \mapsto 1 \end{matrix} \quad \text{и инверсия } L\colon \begin{matrix} 0 \mapsto 1 \\ 1 \mapsto 0 \end{matrix}$$

являются единственными шагами шифрования (шаги шифрования Вернама). Они совпадают с шагами Виженера $+0$ и Бофорта $+1$. Это шифрование является взаимообратным несобственным преобразованием. Значение $|M| = 2$ является самым маленьким числом, которое допускает многоалфавитное шифрование. Таким образом, ключ шифрования Вернама генерируется конечным (O, L)-словом, которое периодически повторяется или же является бесконечной (O, L)-последовательностью, наподобие $O\,L\,L\,O\,L\,O\,O\,L\,L\,O\ldots$ Так как в $\mathbb{Z}_2$ тождественная перестановка O может интерпретироваться как сложение с 0, а отражение L — как сложение с 1, то шифрование $\mathbb{Z}_2 \to \mathbb{Z}_2$ является линейным преобразованием. Сложение в $\mathbb{Z}_2$ (сложение *по модулю* 2), часто обозначаемое символом $\oplus$, совпадает с булевской операцией $\leftarrow|\rightarrow$ (ис-

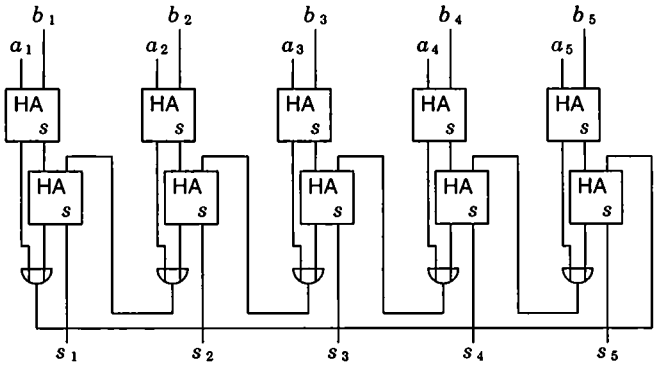


Рис. 58. Схема суммирования построенная из полусумматоров HA

ключающее ИЛИ). Используется также термин «двоичное сложение без переноса», так как результат операции получается как сумма на выходе полусумматора.

8.3.2. Вернам. Использовать для реализации этих двух шагов шифрования электрические реле предложил в 1917 г. (раньше, чем Лестер С. Хилл) молодой служащий AT&T в Нью-Йорке Джильберт С. Вернам (1890–1960 гг.).

Вернам сконструировал блок двоичного шифрования Виженера для коммерческого телетайпа. Ключ перфорировался на обычной 5-дорожечной телетайпной ленте, которая могла быть закольцована для формирования цикла. Двойным шифрованием с короткими циклами в 999 и 1000 символов Лиман Ф. Морхаус из группы Вернама получал ключ длиной в 999000 символов, который (что более важно) был «хаотичным». Вернам подал заявку на патент США 13 сентября 1918 г. и получил его в 1919 г. за номером 1310719. Коммерческого успеха предложение Вернама не имело; коды были более востребованы. Но идея была принята в профессиональной дипломатической и военной криптологии и была реализована, среди прочего, в устройстве на двойной перфоленте Сименса со «смесителем» и, позднее, в шифровальной машине SIGTOT армии США.

8.3.3. Удаление переноса. Переход от шага шифрования Виженера в $\mathbb{Z}_2^n$ (производимого сложением с числом $(a_1a_2a_3\ldots)$ в двоичной системе) к n многоалфавитным шагам шифрования Вернама (т. е. шагам шифрования Виженера в $\mathbb{Z}_2$ с последовательными сложениями с a_1, с a_2 и т. д.) равносильно изъятию блока переноса в электронной схеме суммирования.

То же самое справедливо для шага шифрования Виженера в $\mathbb{Z}_{10^n}$, производимого сложением по модулю 10^n чисел из интервала $\{0, \ldots, 10^n - 1\}$. Для механического настольного калькулятора переход к n шагам шифрования Виженера в $\mathbb{Z}_{10}$ равносильно изъятию механического устройства переноса (разд. 5.7).

8.4. Квазинепериодические ключи

8.4.1. Многоалфавитное шифрование считается трудным. Несмотря на свою криптоаналитическую стойкость, периодическое многоалфавитное шифрование с длинными ключами при его использовании должным образом, было менее популярно, чем номенклаторы. Им поначалу пользовались лишь в исключительных случаях: в папской Курии в 1590 г., где оно было взломано Соро, дешифровальщиком Генри IV; кардиналом де Рецем в 1654 г. для связи с принцем Конде (Людовик II из династии Бурбонов) до войны с гугенотами, тогда же оно было взломано Ги Жоли, который угадал ключевое слово, что было в то время основным методом. В 1791 г. Мария-Антуанетта использовала многоалфавитное шифрование в своих любовных и заговорщических письмах (разд. 2.1.1). Ее любовник с 1783 г. шведский граф Аксель Ферсен, придумал набор в духе Порта из 23 взаимообратных алфавитов (рис. 59). Аксель Ферсен был осторожен и не использовал очевидные мнемонические, но *пустые*, слова, наподобие DEPUIS, VOTRE (уже, ваш). Не вина криптоана-

A	(ab)	(cd)	(ef)	(gh)	(ik)	(lm)	(no)	(pq)	(rs)	(tu)	(xy)	(z&)
B	(bk)	(du)	(ei)	(fl)	(gn)	(ho)	(my)	(ps)	(qx)	(rt)	(ac)	(&z)
C	(lr)	(ad)	(bg)	(cz)	(s&)	(ek)	(fm)	(th)	(ix)	(np)	(oq)	(uy)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Рис. 59. Многоалфавитное взаимообратное шифрование Марии-Антуанетты

литиков в том, что бегство Людовика XVI и Марии-Антуанетты закончилось на мосту Вареньи; ни одно из этих сообщений не было расшифровано.

Многоалфавитная подстановка, до тех пор пока она не была механизирована, имела репутацию громоздкой и чреватой ошибками. Уильям Блэйр писал в статье энциклопедии 1819 г.: «многоалфавитная подстановка требует слишком много времени, а минимальная ошибка при записи заводит в тупик...».

Та же самая претензия обнаруживается в книге XVII в. *Traitté l'art dechiffrer*, изданной в Брюсселе: «...требует слишком многого времени для зашифрования, пропуск единственного символа зашифрованного текста искажает сообщение с этого места...».

8.4.2. Многоалфавитное шифрование считается надежным. Однако многоалфавитная подстановка также имеет репутацию невзламываемой. Маттео Ардженти писал: «Шифр на основе ключа — благороднейший и величайший в мире, он самый безопасный и верный, никогда не сыщется человек, который мог бы выведать его».

До XIX в. удавалось взламывать только шифры, которые использовали тривиальные алфавиты со сдвигами, когда можно было угадать слова открытого текста и восстановить короткий ключ (не говоря уже об угадывании ключевого слова, чему следовали Порта и Ардженти в своей работе). Это положение изменилось с середины XIX в.

Чтобы исключить угадывание, желательно (Паркер Хитт) брать ключевую последовательность с длиной периода значительно большей, чем весь открытый текст (квази-непериодический ключ) или использовать непериодический ключ (разд. 8.7).

8.4.3. Последовательное шифрование. При использовании квазинепериодического шифрования желательно, чтобы число применяемых алфавитов значительно превосходило число символов ключа. Эти алфавиты должны выбираться менее регулярно, чем это обеспечивается мнемоническими ключевыми словами (Кан: «нерегулярная последовательность алфавитов»). Более того, если доступно большое число алфавитов, то имеет смысл использовать именно последовательное шифрование.

Последовательным шифрованием называется периодическое многоалфавитное шифрование, в котором никакой алфавит не используют повторно, пока все другие алфавиты не будут использованы. Таким образом, период d последовательного шифрования совпадает с мощностью θ множества шагов шифрования. Квазинепериодическое шифрование появляется, когда сообщение короче, чем θ.

Последовательное шифрование было предложено уже Тритемием в его *tabula recta* (разд. 7.4 и 8.1.2), хотя с 24 сдвинутыми стандартными алфавитами оно не обеспечивало должной безопасности. Последовательное шифрование реализуется цилиндровыми и реечными устройствами, где каждый алфавит доступен лишь в *одной* копии. Последовательное шифрование также широко использовалось в механических и электромеханических шифровальных машинах первой половины XX в.

8.4.4. «Регулярное» перемещение ротора. Хотя ничто не мешало использовать роторы с большим числом контактов (полу-ротор японской машины *angô kikai taipu A* (рис. 66) имел 60 контактов), в случае $\mathbb{Z}_N$ считалось естественным иметь N контактов и, таким образом, только N алфавитов. Чтобы обеспечить длинный период для последовательного шифрования, решение, к которому независимо пришли Хеберн и Шербиус, заключалось в пошаговом (как в счетчике) сцеплении нескольких роторов («регулярное» перемещение ротора). Для четырех роторов, как в ЭНИГМЕ A ($N = 26$), получаем период d, равный или, по крайней мере (при «почти последовательном шифровании»), не намного меньше $\theta = 26^4 = 456976$. Период такой длины позволяет передать сообщение размером в среднюю повесть. С пятью роторами θ приближается к 12 миллионам, что намного больше количества букв, содержащихся во всей Библии.

8.5. Машины, которые генерируют свои собственные ключевые последовательности

Часто шифровальные машины выполняют две функции: они производят многоалфавитные шифрования и они же генерируют свои собственные ключевые последовательности для выбора этих шифрований. В этом случае генерирование ключевой последовательности становится критическим элементом механизации.

8.5.1. Вейтстоун. Тритемий использовал сдвинутые (стандартные) алфавиты, идущие непосредственно один за другим. Такая конструкция была реализована (со сдвинутыми перемешанными алфавитами) в шифраторе *Криптограф* Вейтстоуна 1867 г. (вклейка С). Это означало использование фиксированного ключа. Использование ключей Белазо обеспечивало опытному шифровальщику достаточную степень случайности при выборе алфавитов.

8.5.2. Стремление к нерегулярности. Генераторы ключевых последовательностей с периодом такой длины, что обычно одно сообщение далеко не исчерпывало его полностью, обеспечивали «нерегулярность при выборе алфавитов» путем выбора начальной точки цикла ключевой последовательности. Поэтому *Артур Шербиус* в своей основной патентной заявке от 23 февраля 1918 г. (нем. патент 416-219) использовал регулярное вращательное движение как в счетчиках, только как одну из возможностей.

Арвид Герхард Дамм, один из авторов принципа ротора, сделал в своей шведской патентной заявке от 10 октября 1919 г. довольно слабую попытку обеспечить нерегулярность: четыре шестерни («ключевые колеса»), по одной

для каждого ротора, сдвигают каждый полу-ротор после каждого шага шифрования на меняющееся число позиций. Эта «нерегулярность» была не очень сильна и могла произвести впечатление только на наивного человека: она имела длину периода $d = 17 \cdot 19 \cdot 21 \cdot 23$ сдвигов ротора; $d \approx 150\,000$, что составляет приблизительно одну треть от $\theta = 26^4 = 456\,976$. Получалось почти последовательное шифрование.

В более поздней заявке, зарегистрированной 26 сентября 1920 г. (нем. патент 425 147), Шербиус применил шестерни («ключевые колеса») с нерегулярно расположенными кулачками. Для ЭНИГМЫ 1923 г. с четырьмя роторами (Пауль Бернштейн, заявка от 26 марта 1924 г., нем. патент 429 122), движение ротора было отчасти нерегулярно, поскольку четыре шестерни имели зазоры: одно колесо с 11 позициями имело 5 кулачков и 6 зазоров, второе колесо с 15 позициями имело 9 кулачков и 6 зазоров, третье колесо с 17 позициями имело 11 кулачков и 6 зазоров, последнее колесо с 19 позициями имело 11 кулачков и 8 зазоров. Таким образом, получаемый для сдвигов ротора период $d = 11 \cdot 15 \cdot 17 \cdot 19$, превышающий 50 000, составляет лишь приблизительно девятую часть θ, что, конечно, вновь приводит к почти последовательному шифрованию.

Неравномерное движение посредством шестерен с изменяющимся числом кулачков и зазоров также использовалось в шифровальной машине (вклейка F) Александра фон Края (заявка зарегистрирована 16 января 1925 г., нем. патент 434 642), но с периодом между 260 и 520 криптологически машина была очень слабой.

Борис Хагелин, который приобрел компанию Дамма *Aktiebolaget Cryptograph* в 1935 г., для выполнения шагов шифрования Бофорта заменил полу-роторы «реечным барабаном», также называемым «рейтерной кассетой» (нем. *Stangenkorb*). Он, однако, продолжил использовать неравномерное движение, производимое «ступенчатыми фигурами» ключевых колес. В машинах C-35/C-36 (рис. 60 a, вклейка G) число ключевых колес было увеличено до пяти, что позволило достичь периода $17 \cdot 19 \cdot 21 \cdot 23 \cdot 25 = 3\,900\,225$. В более поздней модели, улучшенной по совету Ива Гильдена, использовалось шесть ключевых колес, обеспечивающих период $17 \cdot 19 \cdot 21 \cdot 23 \cdot 25 \cdot 26$, т. е. более 100 миллионов (вклейка H)[1]. Хагелин получил из Франции заказ на изготовление 5 000 машин, которые были произведены по лицензии фирмой «Ericsson-Colombes». Во время Второй мировой войны в США компанией по изготовлению пишущих машинок L. C. Smith & Corona по лицензии было изготовлено 140 000 машин, называемых M-209 в армии США и CSP 1500 — в ВМФ США. Ненадежная модель C-38m использовалась итальянским ВМФ в Средиземном море. Позже машины Хагелина («хаги») имели период $29 \cdot 31 \cdot 37 \cdot 41 \cdot 43 \cdot 47$, т. е. более 2 миллиардов. Электромеханическая машина BC-543 (рис. 60 b) с клавиатурой и печатающим устройством использовалась в США для связи в армейских подразделениях среднего уровня.

[1] Детали см. в книге Саломаа А. Криптография с открытым ключом. Пер. с англ. — М.: Мир, 1996. — с. 44).

Позже во время войны она была скопирована под именем C-41 немецкой фирмой по изготовлению пишущих машинок «Wanderer Werke». После войны Хагелин продолжал совершенствовать свои машины. В 1952 г. на рынок поступила машина CX-52 фирмы «Хагелин-Крипто», имевшая от шести до двенадцати ключевых колес (H 54 изготавливалась по лицензии фирмой «Dr. Hell Co.» из Киля).

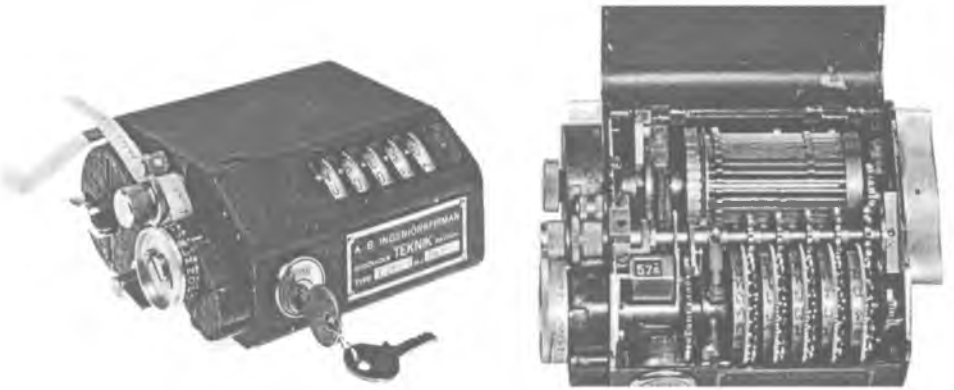


Рис. 60 a. Шифратор C-35, созданный Б. Хагелином (A. B. Cryptoteknik, Стокгольм)

Рис. 60 b. Слева: шифратор BC-543 (Компания «Хагелин Криптограф», США); справа: немецкая копия с шифратора C-41 фирмы «Wanderer Werke»

8.5.3. Вращение колеса с помощью храповиков и пазов. Позднее, введя отражатель и три подвижных ротора, Шербиус отказался от шестерен и заменил их храповиками и пазами на роторах. Период был немного меньше максимального $\theta = 26^3$, а именно $26 \cdot 25 \cdot 26 = 16\,900$, из-за конструкции ку-

лачкового механизма. Соответственно, максимальная длина сообщения была ограничена 180 символами (с 13 января 1940 г. — 250 символами).

В ЭНИГМЕ I и в ЭНИГМЕ *Вермахта* «регулярное» перемещение роторов обеспечивалось благодаря наличию *одного* паза в *алфавитном кольце* каждого ротора. Самый «быстрый» ротор (крайний справа) R_N перемещался на одну позицию на каждом шаге шифрования. Каждый полный оборот его вызывает перемещение на одну позицию «среднего» ротора (центрального) R_M, полный оборот которого, в свою очередь, вызывает перемещение на одну позицию «медленного» (крайнего левого) ротора R_L. Фактически это означало регулярное, как в счетчике, вращение роторов («греческие» роторы β и γ, которые были введены позднее, не имели возможности перемещаться).

Чтобы обеспечить хотя бы некоторую нерегулярность, пазы в роторах ЭНИГМ I–V *Вермахта* размещались в различных позициях алфавитных колец[2]:

Ротор	I	II	III	IV	V
Буква алфавита	Y	M	D	R	H

Но это создавало лишь *иллюзорную сложность*. Хуже того, это была «сложность, которая аннулирует саму себя», как иронически сказал Кан: «Если даже все роторы имеют пазы, вырезанные у одной и той же буквы, криптоаналитики не смогут выяснить, какой из роторов используется в качестве быстрого, определяя (для известных роторов), какая буква вызвала оборот». На флоте (*Kriegsmarine*), по-видимому, были в курсе этого и на новых роторах VI и VII (1938 г.) и VIII (1939 г.) располагали пазы у одинаковых букв. Более того, эти роторы имели два паза: один у буквы H и другой — у буквы U (вклейка K). В отличие от этого в коммерческой ЭНИГМЕ пазы располагались на алфавитном кольце. Таким образом, перемещение роторов зависело от установки колец.

Хотя использование двух пазов делит период пополам и увеличивает опасность наложения (см. разд. 19.1; в качестве контрмеры приходится резко ограничивать допустимую длину каждого единичного сообщения), оно делает криптоанализ более трудным: «Мы имели бы большие затруднения, если бы каждое колесо имело две или три оборотных позиции вместо одной» (Велчман). Три паза, между прочим, не сократили бы период, так как 3 взаимно просто с 26. Было ли это упущением специалистов OKW, неизвестно.

Специальный вариант ЭНИГМЫ для служб контрразведки и разведки (*Абвер*) адмирала Вильгельма Канариса не имел — как ни странно — коммутационной панели, но подобно модели D имел отражающий рефлектор. Другое отличие ЭНИГМЫ *Абвера* от модели D состояло в том, что в ней вместо пазов и храповиков использовались шестеренные муфты (очередной нем. патент 1928 г. Корна), допускающие, в соединении с тахометром, вращение роторов

[2] Оборот совершается, когда (с разностью в 19 символов) в окне появляются буквы R, F, W, K и A. В Блетчли Парк для запоминания этих букв пользовались соответствующим, довольно глупым мнемоническим стихом *Royal Flags Wave Kings Above* (волна королевских флагов выше королей).

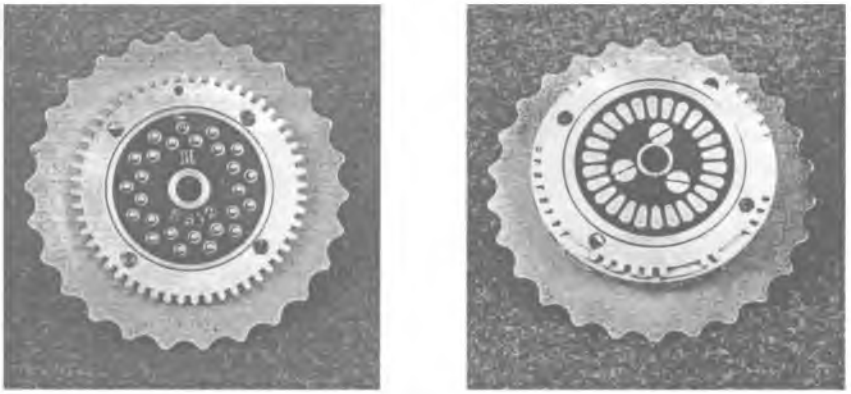


Рис. 61. Лицевая и обратная стороны ротора ЭНИГМЫ Абвера: слева — с 2×26 зубцами; справа — с парой зубцов, используемых как пазы

вперед и назад. Она имела новые роторы; но как и в ЭНИГМЕ *Вермахта*, на алфавитном кольце были установлены остановочные позиции. Три ротора имели 11, 15 и 17 (а не 19, как утверждал Твинн) остановочных позиций. Такая конструкция заставила потрудиться Дилвина Кнокса; но осенью 1941 г. он достиг успеха в криптоанализе этой машины. Кнокс ввел для ЭНИГМЫ *Абвера* специальную терминологию: некоторые совместные вращения роторов R_N, R_M, R_L и отражателя он называл «краб» и «омар».

8.5.4. Typex. На вклейке L показана британская роторная машина TYPEX (Type-X), которая разрабатывалась в течение девяти лет под контролем правительственной комиссии и была готова в 1935 г. Эта машина не предназначалась для коммерческого применения. В некоторых отношениях она была подобна ЭНИГМЕ с 3 роторами; отличие состояло в том, что среди ее пяти роторов два первых входных ротора были неподвижны. Лампочный выход ЭНИГМЫ был заменен ленточным буквопечатающим аппаратом. Криптологически TYPEX была эквивалентом ЭНИГМЫ с 3 роторами с невзаимообратной коммутационной панелью. Существенные различия имелись в перемещении ротора: оно также было регулярно, но порождалось несколькими пазами. Обод с пазами жестко фиксировался на кромке ротора как в коммерческой ЭНИГМЕ, а сердечник ротора был «коммутирующим блоком», расположенным в патроне, несущем обод ротора и алфавитное кольцо. Коммутирующие блоки могли устанавливаться в двух положениях, P или P^{-1}. В типовой версии машины можно было выбрать пять блоков из десяти. Имелись ободы с пятью, семью и девятью пазами; в последнем случае пазы размещались так, чтобы оборот производился, когда в окне появляется один из символов B, G, J, M, O, R, T, V, X. Все роторы при совместном использовании имели идентичное расположение пазов.

В итальянской шифровальной машине OMI Cryptograph-CR (разд. 7.3.4) роторы собирались из патронов, снабженных пазами, и сердечников двух ви-

дов. В машине было 14 роторов, сцепленных парами или 7 роторов, первый из которых можно было выбрать $\binom{14}{2} = 91$ способами, второй — $\binom{12}{2} = 66$ способами и т. д. (рис. 62).

После 1945 г., помимо ЭНИГМ, во многие малые страны поставлялись и TYPEX-ы из запасов, оставшихся после второй мировой войны. Некоторые из них использовались вплоть до 1975 г.

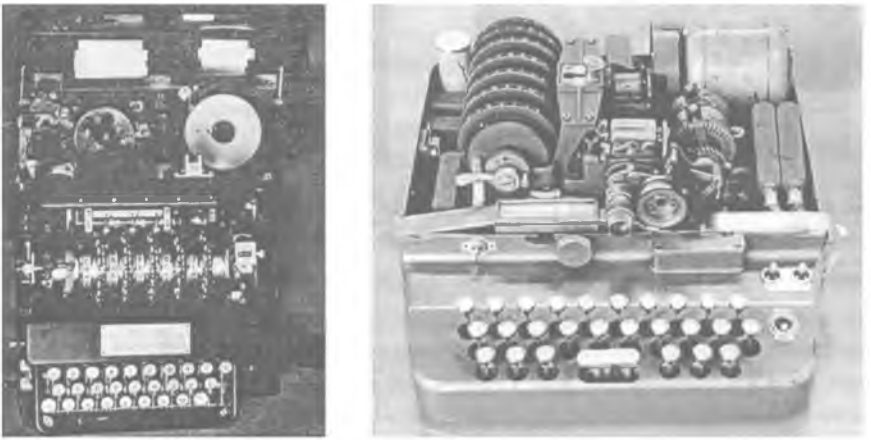


Рис. 62. Слева: TYPEX Марк II; справа: роторная машина фирмы Ottica Meccanica Italiana

С конца 1940-х до начала 1980-х гг. НАТО использовало роторные машины KL-7, разработанные в США для международной связи. (Американская SIGABA считалась правительством США слишком хорошей машиной, чтобы делиться ею). KL-7 (рис. 63) имела семь шифровальных роторов. Используя взаимозаменяемые кодирующие цилиндры и кольца, с точки зрения механики, она была слабым подобием британской TYPEX. Пластмассовые токосъемники, которые управляли перемещением ротора, могли переставляться по кодирующим цилиндрам. Каждый ротор имел 36 контактов, предусмотренных для букв и цифр. KL-7 была одной из последних роторных машин, производимых когда-либо. Уровень безопасности, обеспечиваемый этими машинами, был столь высок, что они поставлялись даже в некоторые не натовские страны. Это показывает, что криптологи в 1960-х гг. полностью принимали принцип Шеннона (разд. 11.2.3), гласящий, что система шифрования должна оставаться безопасной, даже если устройство шифрования имеется в руках противника. Действительно, уже в 1962 г. американский офицер Джозеф Г. Хельмих продал Советскому Союзу техническую информацию о роторах и списки ключей; он был арестован ФБР в 1982 г. Использование KL-7 прекратилось в 1985 г. после шпионского дела Уолкера — к этому времени она уже устарела.

Русский аналог, 10-роторная машина, называлась «Фиалкой».

Рис. 63. Роторная машина KL-7 (кодовое имя ADONIS)

8.5.5. Хеберн. На раннем этапе своей карьеры великий Уильям Фридман изучал роторные машины Хеберна, который сотрудничал с ВМФ США, и в 1925 г. дал им экспертную оценку. Его тестирование машины Хеберна было шедевром (*chef-d'œuvre*). Фридману дали десять сообщений примерно в 300 символов, зашифрованные с одинаковым порядком и начальной установкой роторов. Через две недели работы, он нашел решение, включая реконструкцию коммутации контактов, по крайней мере, некоторых роторов. Итоговый отчет был, в конце концов, рассекречен в 1996 г.; очевидно, что в него попал и его *индекс совпадения* (разд. 16.1).

Криптоаналитики ВМФ во главе с Лоуренсом Ф. Саффордом и армии во главе с Фридманом, вместе с Синковым, Роулеттом и Кульбаком потратили много лет, изучая различные усовершенствования машины Хеберна. В 1932 г. Хеберн, наконец, спроектировал удовлетворительную машину НСМ с пятью роторами. Вращение ее роторов было достаточно нерегулярным, но все еще не удовлетворяло группу Фридмана.

Примерно в 1935 г. Фридман сам разработал на базе ЭНИГМЫ для Корпуса связи армии США машину с 4 роторами, Converter M-325 (которая была даже построена в 1944 г. и дублировала SIGFOY, но из-за некоторых практических недостатков не была принята на вооружение). Затем, вновь под давлением ВМФ, появилась «электрическая шифровальная машина» ECM Mark I, 5-роторная машина с вращением роторов, управляемым цевочными колесами, которая наконец-то удовлетворяла самым высоким требованиям. Но Франк Роулетт продолжал ее совершенствовать, что привело к созданию ECM MarkII, часто называемой в армии США просто ECM, а также M-134-C и SIGABA, а в ВМФ именовавшейся CSP 889 (рис. 64). SIGABA имела 15 роторов: 5 шифровальных, 5 роторов, обеспечивающих неравномерность движения, остальные пять роторов были эквивалентом коммутационной

панели. В 1940-х гг. она обоснованно была самой засекреченной шифровальной машиной в (западном) мире. Она же была и самой дорогой. Система использовалась вплоть до 1959 г. CCM, другая «комбинированная шифровальная машина», также называемая CSP-1700, была гибридной машиной для связи с SIGABA и TYPEX. Из машины Хеберна путем последующих усовершенствований профессиональными криптологами родилась машина ECM Mark III: разработанная после 1934 г. и продолжавшая все еще использоваться во время Второй мировой войны.

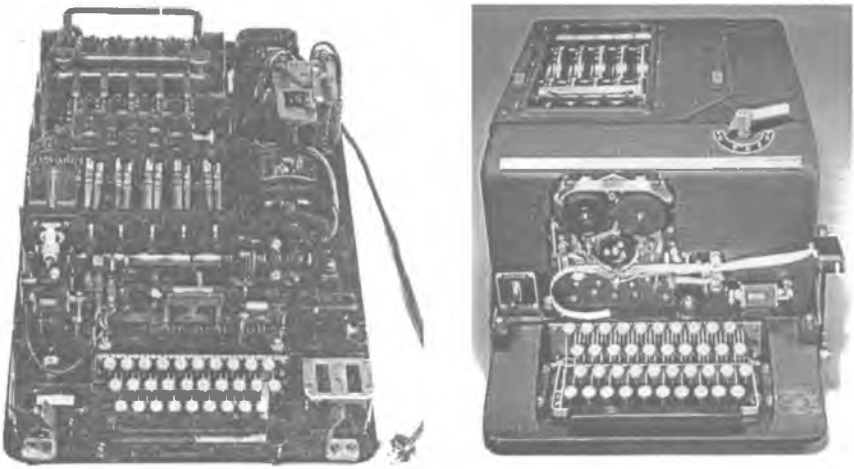


Рис. 64. Роторная машина ECM Mark II (M-134-C, SIGABA, CSP 889)

8.5.6. Ярдли. Япония, становясь великой восточно-азиатской державой, не могла после Первой мировой войны обходиться без дипломатии и, таким образом, нуждалась в криптологии. Дипломаты, как и все остальные, использовали кодовые книги. Правительство привлекло консультантов; польский капитан Ян Ковалевский познакомил их с самыми простыми мерами защиты типа «русского соединения» (разд. 3.4.2). Между 1919 г. и весной 1920 г. японцы ввели одиннадцать кодовых книг, некоторые из которых содержали по 25 000 кодовых групп. Японский радиообмен, естественно, вызвал интерес *«Американского черного кабинета»*, работавшего одновременно на госдепартамент и военное министерство США. Поддержанный отделом МИ-8 военного министерства США, «черный кабинет» был организован после 1918 г. Гербертом Осборном Ярдли (1889–1958 гг.). Официально он был вспомогательным подразделением управления военной разведки. Располагался «черный кабинет» в Нью-Йорке с соблюдением строжайшей маскировки; с 1925 г., после того, как на офис был совершен налет, он прикрывался вывеской «Компании по составлению кодов», которая действительно составила и продавала *«Всеобщий торговый код»*. Ярдли и его люди были трудолюбивы и прилежны; летом 1921 г. они расшифровали телеграмму японского посла в Лондоне

своему министру иностранных дел, содержавшую деликатную информацию о готовящейся международной морской конференции по разоружению и раскрывающую устремления экспансионистской Японии на Дальнем Востоке. К 1929 г. черный кабинет расшифровал около 45 000 телеграмм со всего мира.

4 марта 1929 г. Герберт К. Гувер был провозглашен 31-м президентом Соединенных Штатов Америки, и все внезапно изменилось. Наивность Гувера привела к тому, что он и госсекретарь Генри Л. Стимсон больше не считали возможным пользоваться услугами криптоаналитиков. «Черный кабинет» был без колебаний распущен; окончательно это произошло 31 октября 1929 г. Рабочие материалы перешли к армейскому Управлению связи, руководимому Фридманом. Ярдли вынужден был искать другое место; в разгар Депрессии это было чрезвычайно трудным делом. Он был вынужден зарабатывать деньги и подталкиваемый горечью и нуждой решился написать книгу, потрясающее разоблачение под заголовком «*Американский черный кабинет*» (Индианаполис, 1931 г.). Ярдли был превосходным рассказчиком, и его книгу ожидал немедленный успех. Этим он навлек на себя гнев и презрение правительства. Защищаясь, Ярдли обвинил государственный департамент в глубоком пренебрежении интересами США, выразившемся в использовании «шифров шестнадцатого века», и заявил, что тот не имеет никакого морального права оказывать давление на него. Более серьезными были претензии со стороны его коллег по профессии; они лучше Стимсона знали, что ввиду возможной войны национальные интересы требуют сохранения государственных тайн и преемственности в криптологической работе.

История с Ярдли имела законодательное продолжение. Конгресс США 73-го созыва рассмотрел в 1933 г. спорный законопроект, внесенный администрацией Рузвельта, который предусматривал наказание за издание или передачу без разрешения материалов, полученных в процессе передачи информации между каким-либо иностранным правительством и его дипломатической миссией в Соединенных Штатах. Этим законом были нарушены свобода печати и 37 статья публичного права. *Закон Ярдли* вошел в раздел 952 Титула 18 свода законов Соединенных Штатов, но никакого уголовного преследования Ярдли не последовало.

Ярдли перечислил девятнадцать стран, чьи дипломатические коды были взломаны. Среди этих кодов были коды одиннадцати южноамериканских стран, Либерии и Китая — что не вызвало удивления — но также и коды Англии, Франции, Германии, Испании и Советского Союза, со стороны которых, по крайней мере официально, не могло быть выражения морального осуждения (считается, что в 1920-х годах каждая большая европейская страна обладала одной или несколькими американскими кодовыми книгами) — и Японии.

Книга имела огромный успех, не в последнюю очередь из-за разгоревшегося скандала, всколыхнувшего публику. Было продано 171 931 экземпляров книги в США и позже еще 514 80 в Великобритании, что было неслыханным тиражом для книги по криптологии. Последовали ее переводы на французский, шведский, китайский и японский языки. Сенсацией были проданные

33 931 экземпляров в Японии. Это показало, что Ярдли задел чувствительный нерв японской души.

В Японии член верхней палаты парламента позволил себе довольно-таки невежливые и жесткие выражения; обвиняя свое собственное министерство иностранных дел, он говорил о «подрыве веры, совершенной правительством Соединенных Штатов»; министр иностранных дел и бывший японский посол в Соединенных Штатах назвал это «постыдным делом». Ярдли был опорочен. Тем не менее, Японии он оказал величайшую услугу, стимулировав радикальное улучшение ее криптоаналитической защиты. В результате японцы умножили свои усилия по механизации шифрования.

В 1938 г. Ярдли был нанят Чан Кай Ши и впоследствии вскрывал японские столбцовые перестановки. В 1940 г. он вернулся из Китая; в июне 1941 г. отправился в Канаду, чтобы работать над шпионскими шифрами, но как достоверно утверждает Оливер Стречи, под англо-американским давлением через шесть месяцев был заменен.

Рис. 65. Японская имитация ЭНИГМЫ, машина GREEN

8.5.7. GREEN, RED и PURPLE. Японцы изучали шифровальные машины других стран, в частности, машины, доступные благодаря патентной литературе: ЭНИГМУ, машины Дамма и Хагелина, а также Хеберна. Для машин с латинским алфавитом использовалась общая транслитерация Хепбурна японского алфавита *kana* в латиницу. Японская имитация ЭНИГМЫ D, обозначаемая GREEN американскими криптоаналитиками, была странной конструкцией с четырьмя вертикально установленными роторами (рис. 65) и не имела большого значения. Затем в машине *angō kikai taipu A* (шифровальная машина A), в американском обиходе называемой RED, появились полу-роторы Дамма. В отличие от фиксированной перестановки, задаваемой коммутационной панелью, она имела полу-ротор с 26-ю контактными кольцами (рис. 66).

Электрическая схема переставляла шесть гласных между собой и, соответственно, также 20 согласных и, таким образом, нуждалась (два раза) в 60-и выходных контактах, так как 60 есть НОК 6 и 20. Причина этого криптологически довольно неблагоприятного деления букв лежала, возможно, в тарифных правилах международного телеграфного союза, требующего «произносимости» слов.

Рис. 66. Полу-ротор с 26 контактными кольцами в японской машине *angö kikai taipu A*

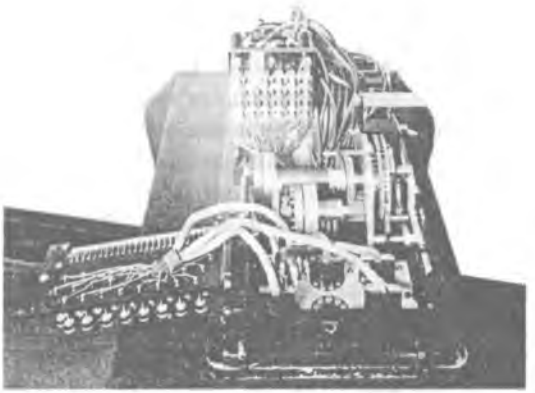


Рис. 67. Американская реконструкция RED машины *angö kikai taipu A* с двумя полу-роторами

Японская машина RED была весьма слабой системой шифрования, не намного лучше машины Края. Вращение ротора производилось шестеренкой с 47 зубьями и с 4, 5 или 6 пазами. Криптологически *angö kikai taipu A* совершает два шифрования АЛЬБЕРТИ: отдельно шифруются группы гласных и согласных букв. Не удивительно, что машина RED была взломана в 1935 г. криптоаналитиками армии США Кульбаком и Роулеттом и к 1936 г. полностью реконструирована Агнессой Дрисколл (реконструкция имела два

полу-ротора, с шестью и с 20-ю контактными кольцами, рис. 67). Весной 1936 г. Кюнце из отдела *Pers Z* немецкого МИДа (*Auswärtiges Amt*) изучил вариант машины, работающей с алфавитом *кана* (у американцев этот вариант носил имя ORANGE). Джек С. Хольтвик из ВМФ США также исследовал эту машину. Оба они добились успеха в ее взломе, Кюнце — в августе 1938 г.

В 1937 г. Япония начала разработку шифровальной машины, которая была бы гораздо более безопасна. Такая машина была введена в строй в 1939 г. и заменила машину RED в дипломатических службах — первые сообщения, зашифрованные с ее помощью, были посланы в марте 1939 г. в Токио из Варшавы. *Angō kikai taipu B* (шифровальная машина B, называемая также «*97-шики обун инжики*» (*97-shiki obun injiki*), буквопечатающее устройство '97), называемая американцами PURPLE, обладала новыми особенностями, впервые использованными японцами. В этой машине применялись шаговые переключатели, используемые на телефонных станциях. Разделение на две группы из шести и 20 символов сохранилось, хотя позже ограничение на группу из шести букв (быть гласными) было снято. В результате число имеющихся алфавитов сократилось до 25, преобразование же было достаточно нерегулярным и определялось внутренней коммутацией. Чтобы вычислить его, потребовалась сосредоточенная месячная работа целой группы людей, не только Уильяма Фридмана и Франка Роулетта, но также и Роберта О. Фернера, Альберта У. Смолла, Самуэля Шнайдера, Женевьевы Фейнштейн и Марии Джо Дьюнинг. Они сначала искали преобразование группы из 6 гласных и получили указания на то, что число алфавитов равно 25. Но что касается группы из 20 согласных, все было покрыто мраком, и никто не мог идентифицировать электромеханический шаг шифрования, который приводил к наблюдаемым результатам. В середине лета 1940 г., когда положение казалось почти безнадежным, недавно принятый новичок из МТИ Лео Розен натолкнулся на идею, что японцы, возможно, использовали шаговые переключатели (рис. 68). Это придало работе дополнительный импульс, и тайна была вскоре раскрыта: в машине было три блока шаговых переключателей, знание этого факта позволило определить коммутацию соединений. Вскоре вслед за важным открытием Женевьевы Грожан, была изготовлена работающая реконструкция машины и в августе 1940 г. после 18 месяцев работы было получено первое полное решение для PURPLE. В январе 1941 г. англичанам в Блетчли Парк была отправлена ее копия. Хотя машина RED положила начало пути к успеху, а низкая дисциплина шифрования японцев давала ключи и подсказки, тем не менее это была победа, добытая криптоаналитическим бюро армии США, «которую не смог повторить кто-либо еще... британская и немецкие криптоаналитические службы были сбиты с толку в своих попытках вскрыть шифр» (Фридман). Речь шла о MAGIC. (Так назывался проект по вскрытию и чтению шифрованных сообщений японских военно-морских сил. — *Прим. перев.*). Вскоре и англичане одержали свою победу — над немецкой ЭНИГМОЙ; все материалы, связанные с ней, они назвали ULTRA. С другой стороны, Дэвид Кан и Отто Лейберих сообщали, что русские и немцы также взломали шифр PURPLE.

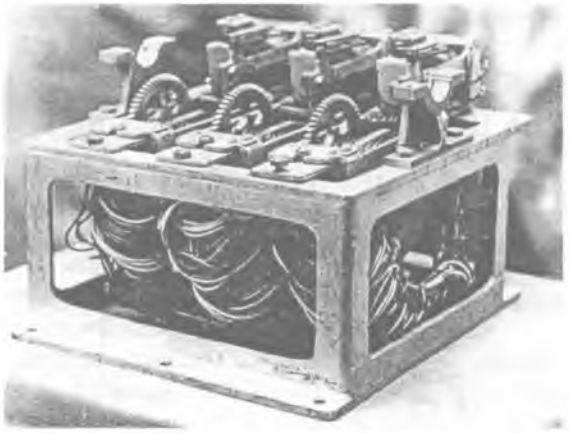


Рис. 68. Устройство шагового переключателя японской машины PURPLE

Как только машина PURPLE была реконструирована, выяснилось, что она обеспечивает лишь посредственный уровень защиты, сопоставимый с уровнем RED. Кажется, японцы недооценили сообразительность американцев; они также уверились, что их язык предохранит их и не будет понят полностью где-либо еще. Последующие машины, которые они также конструировали с использованием шаговых переключателей, были только слегка усложнены: в одной, называемой по американской классификации CORAL, сохранялось разбиение $20 + 6$; окончательно эта машина была взломана группой OP-20-GY в марте 1944 г. Другая машина, называемая JADE, была уникальна тем, что печатала в символах *кана*. Она была немного сложнее других и также была взломана.

У японцев была очень прозрачная система ежедневных установок коммутационной панели и плохая привычка посылать информацию о смене ключей в виде зашифрованных сообщений. Таким образом, взломав эти сообщения, противник получал самые свежие ключи. В 1941 г. Фрэнком Рейвеном был обнаружен даже «ключ на ключи».

8.6. Автономное формирование ключевых последовательностей

8.6.1. Матричные степени. Для шифрований ВИЖЕНЕРа и БОФОРТа требуются «нерегулярные» ключевые последовательности цикловых чисел из $\mathbb{Z}_N$. Чаще других применяется метод, в котором используются последовательные степени *по модулю* N невырожденной $k \times k$ матрицы T, достаточно сильно отличающейся от единичной матрицы. Так как число таких матриц (разд. 5.2.3) не превышает N^{k^2}, то при некотором $r > 0$ степень T^r впервые становится равной единичной матрице. Число $r = r(T, N)$ называется порядком матрицы T в $\mathbb{Z}_N$.

Например, матрица $A=\begin{pmatrix}0&1\\1&1\end{pmatrix}$ с $k=2$ имеет следующие порядки (см. 9.4.2):

$N=$	2	3	4	5	6	7	8	9	10	11	12	13	16	20	23	24	25	26	32	48	64	80	160
$r=$	3	8	6	20	24	16	12	24	60	10	24	28	24	60	48	24	100	84	48	24	96	120	240

Выбрав соответствующий (i, j)-элемент матричных степеней, получим циклическую последовательность чисел, минимальный период которой равен $r(A, N)$. Особенно удобной формой матрицы T является $k\times k$ «сопровождающая матрица» вида

$$A=\begin{pmatrix}0&0&0&\dots&0&\alpha_k\\1&0&0&\dots&0&\alpha_{k-1}\\0&1&0&\dots&0&\alpha_{k-2}\\&&\dots\dots\dots\dots&&&\\0&0&0&\dots&0&\alpha_2\\0&0&0&\dots&1&\alpha_1\end{pmatrix}.$$

В этом случае $(1, k)$-элемент степеней этой матрицы равен последнему элементу итерированного вектора $t_i=t_0A^i=t_{i-1}A$ при начальном векторе $t_0=(100\dots00)$.

Для порождения этих итерированных векторов используется k-разрядный регистр сдвига. Регистры сдвига применительно к сопровождающей матрице также называются линейными регистрами сдвига. Эти регистры легко взламываются при помощи базисного анализа (разд. 20.3). Более стойкими являются нелинейные регистры: они формируют следующий элемент последовательности с помощью некоторой произвольной функции от последних k элементов.

Построение нелинейных регистров с помощью простых преобразований, вроде изменения порядка компонентов векторов после каждого шага может быть достаточно опасным: такая нелинейность может приводить к очень коротким периодам (Сельмер 1993 г., Брюнельссон 1993 г.).

8.6.2. Битовые последовательности. Для двоичного случая $N=2$ шифрования Вернама ключевые последовательности являются (0, 1)-последовательностями, где 0 означает тождественное отображение O и 1 — отражение L. Например, матрица $(k=3)$

$$A=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$$

по модулю 2 порождает последовательность (0 1 0 1 1 1 0 0 1 0 1 1 1 0 0 1 0 1 1 1 0 ...) с периодом $7=2^3-1$. Так как имеется 2^k различных k-битовых векторов, а нулевой вектор инвариантен, то очевидно, что максимальный период достигает величины 2^k-1. Можно показать (Ойштейн, 1948 г.), что если многочлен $x^k-\alpha_1x^{k-1}-\alpha_2x^{k-2}-\dots-\alpha_k$ в поле $\mathbb{Z}_2=\mathbb{F}(2)$ является неприводимым, тогда каждая последовательность векторов, порождаемая с помощью

сопровождающей матрицы A имеет период, который является делителем числа $2^k - 1$.

Для $k = 31$ многочлен $x^{31} + x^{13} + 1$ неприводим, и соответствующий период $2^{31} - 1$ достигает величины, превосходящей 2 миллиарда.

Если $2^k - 1$ простое (число $2^k - 1$ тогда называется простым Мерсенна[3]), то имеются только периоды $2^k - 1$ и 1; период 1 имеет последовательность (0 0 0 0 ...).

Для $N = 2$, т. е. в $\mathbb{Z}_2$, имеются только шаги шифрования ВИЖЕНЕРа и БОФОРТа $+\mathbf{O}$ и $+\mathbf{L}$, т. е. шаги Вернама $O \hat{=} 0$ и $L \hat{=} 1$. Для многоалфавитного *двоичного* шифрования нужны особенно длинный период и хороший механизм порождения нерегулярных (0, 1)-последовательностей.

8.6.3. В принципе, из каждого многоалфавитного множества шагов блочного шифрования χ_i с постоянной шириной блока m, который может быть довольно большим, может быть сформирована конечная последовательность (см. разд. 2.3) $X = (\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_s})$, где X может порождаться из начального ключа $u = (u_1, u_2, \ldots, u_s)$: рекурсивная последовательность

$$u,\ X(u),\ X^2(u),\ X^3(u),\ \ldots$$

является периодической, но обычно с таким большим периодом[4], что может быть пригодной для использования в качестве квазинепериодической ключевой последовательности. Например, в разд. 9.5.2 последовательность X будет определена с помощью h-й степени по *модулю* простого числа p,

$$X(u) = u^k \bmod p, \quad X^s(u) = u^{(h^s)} \bmod p.$$

8.7. Непериодические ключи

Для непериодического шифрования (разд. 2.3.3) требуется $\theta \geqslant 2$ и непериодическая последовательность $(\chi_{i_1}, \chi_{i_2}, \chi_{i_3}, \ldots)$ многоалфавитных шагов шифрования. Эта последовательность характеризуется последовательностью индексов $(i_1, i_2, i_3, \ldots)$ с $0 \leqslant i_\mu < \theta$, или правильной дробью $0.i_1 i_2 i_3 \ldots$ в системе счисления с основанием $\theta > 2$. Таким образом, для каждого иррационального вещественного числа и для каждого $\theta \geqslant 2$ существует непериодическое шифрование.

---

[3] Это справедливо для $k = 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127$. Следующие простые числа $2^{521} - 1$, $2^{607} - 1$, $2^{1279} - 1$, $2^{2203} - 1$, $2^{2281} - 1$ были найдены в 1952 г. Ральфом М. Робинсоном с использованием компьютера SWAC. Было найдено еще пятнадцать простых чисел Мерсенна, за которыми последовали $2^{859433} - 1$ (1994), $2^{1257787} - 1$ (1996), $2^{1398269} - 1$ (1996), $2^{2976221} - 1$ (1997), $2^{3021377} - 1$ (1998) и $2^{6972593} - 1$ (1999). (В 2002 г. было найдено еще одно, самое большое на сегодняшний день, простое число Мерсенна $2^{13466917} - 1$. — *Прим. перев.*)

[4] Следуя Роберту Флойду, со значительным вычислительным затратами, но при минимальной потребности памяти, период последовательности X может быть определен следующим способом.

Пусть $a_0 = u$, $b_0 = u$ и $a_{i+1} = X(a_i)$, $b_{i+1} = X^2(b_i)$. Как только $a_n = b_n$, имеет место равенство $X^n(u) = X^{2n}(u)$, и n есть искомый период.

8.7.1. Заблуждения. При $\theta = 2$ непериодическое шифрование ВЕРНАМа с *бесконечной* индексной последовательностью («бегущий ключ»)

$$(LLOLOOOLOOOOOOOL \ldots),$$

т. е. использование последовательности

$$i_\mu = \begin{cases} L, & \text{если} \quad \mu = 2^k \text{ для целого неотрицательного } k, \\ O & \text{иначе,} \end{cases}$$

не дает никакого преимущества по сравнению с периодическим шифрованием; оно даже хуже. Но даже непериодическое шифрование с ключевой последовательностью (Аксель Туэ, 1904 г.; Марстон Морс, 1921 г.) «Мефисто-полька», использованной Максом Ювом в 1929 г.

$$(10010110011010010110 \ldots),$$

имеет довольно прозрачное правило формирования ключа, допускающее рекурсивные вычисления. И рекурсивная последовательность $(0-1)$ слов

$$\begin{aligned} a_0 &\hat{=} (0) \\ a_1 &\hat{=} (1) \\ a_2 &\hat{=} (01) \\ a_3 &\hat{=} (101) \\ a_4 &\hat{=} (01101) \\ a_5 &\hat{=} (10101101) \\ a_6 &\hat{=} (0110110101101) \\ a_7 &\hat{=} (101011010110110101101) \\ &\vdots \qquad\qquad \vdots \end{aligned}$$

определенная Линденмайером и называемая системой замены (Аристид Линденмайер, 1968 г.)

$$\begin{cases} 0 \to 1 \\ 1 \to 01 \end{cases}$$

также имеет прозрачный закон формирования ключа: для $i \geqslant 2$ верно соотношение $a_i = a_{i-2} \circ a_{i-1}$.

Очевидно, что непериодические последовательности могут быть довольно «регулярными». Насколько легко можно получить непериодическую индексную последовательность, которая «нерегулярна» и в то же время известна и отправителю и получателю?

Идея взятия текста из широко распространенной книги регулярно изобретается главным образом любителями. Согласно принципу Шеннона «противник знает используемую систему» это ведет к фиксированному ключу, со

всеми опасностями, которые уже упоминались в разд. 2.6.1. Для осмысленных ключевых текстов существует систематический метод взлома шенноновских систем шифрования с известными алфавитами (разд. 2.6.4) — метод «зигзага» (разд. 14.4).

8.7.2. Автоключи. Не удивительно, что перспективы для получения непериодического ключа из открытого текста обсуждались очень давно. Решительный шаг был сделан Джеронимо (Джироламо) Кардано (1501–1576 г.). Позже Белазо в своей книге *De Subtilitate* в 1550 г. ввел многоалфавитную подстановку с ключами, вслед за Кардано используя открытый текст, начиная ключ с начала в каждом новом слове открытого текста:

s i c	e r g o	e l e m e n t i s
S I C	S I C E	S I C E R G O E L
N T F	Z C L T	Z V H R Y V I P E

Здесь используется алфавит $Z_{20} \cup \{x, y\}$, линейное многоалфавитное шифрование со следующей нумерацией букв:

a	b	c	d	e	f	g	h	i	l	m	n	o	p	q	r	s	t	v	x	y	z
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

Идея автоключа (фр. *autoclave*, *autochiffrant*) задумывалась с наилучшими намерениями, даже завораживала; но Кардано, возможно, никогда не испытывал ее. Шифрование полифонное: буквы s и S также как f и F переходят в N; i и I также как x и X переходят в T; c и C также как p и P переходят в F и т. д. По сравнению с получателем взломщику потребуется не больше усилий для нахождения правильной комбинации из 2^k возможных (если первое слово имеет k букв). Белазо пробовал исправить дефект, шифруя первое слово согласно Тритемию и затем для каждого следующего слова в качестве ключа используя первую букву предыдущего слова и следующие после нее буквы алфавита:

s i c	e r g o	e l e m e n t i s
A B C	S T V X	E F G H I L M N O
T M E	Z N D M	L R N V P Z G Y H

Однако это был все еще слишком прямолинейный метод. Затем Блез де Виженер высказал блестящую идею введения короткого произвольно выбранного начального ключа: он выбирал первый символ ключа и брал следующие ключевые символы либо из открытого текста, либо из шифротекста («автоключ»):

a	u	n	o	m	d	e	l	e	t	e	r	n	e	l
D	A	U	N	O	M	D	E	L	E	T	E	R	N	E
X	I	A	H	G	U	P	T	M	L	S	H	I	X	T

a	u	n	o	m	d	e	l	e	t	e	r	n	e	l
D	X	H	E	E	C	O	U	M	X	G	N	A	B	Q
X	H	E	E	C	O	U	M	X	G	N	A	B	Q	O

В этом случае многоалфавитное шифрование над $\mathbb{Z}_{20}$ (рис. 69) становилось взаимообратным шифрованием Порта, а не шифрованием ВИЖЕНЕРа.

Второй случай, однако, является бесполезным: ключ полностью открыт, и все сообщение (кроме, возможно, первой буквы) может быть расшифровано сразу (Шеннон, 1949 г.).

Лучше обстоит дело с секретностью в первом случае: здесь имеет место рекуррентный метод, помогает только знание первой буквы ключа. Число различных вариантов ключа не превосходит пары дюжин. Выход состоит в использовании начального ключа из d букв вместо одной буквы. Комбинаторная сложность получается, тем не менее, такая же, как для периодического шифрования с ключом длины d. При достаточно большом d проверка выполняется не дольше, но если один и тот же начальный ключ используется

A	*B*	↕	a	b	c	d	e	f	g	h	i	l
			m	n	o	p	q	r	s	t	u	x
C	*D*	↕	a	b	c	d	e	f	g	h	i	l
			x	m	n	o	p	q	r	s	t	u
E	*F*	↕	a	b	c	d	e	f	g	h	i	l
			u	x	m	n	o	p	q	r	s	t
G	*H*	↕	a	b	c	d	e	f	g	h	i	l
			t	u	x	m	n	o	p	q	r	s
I	*L*	↕	a	b	c	d	e	f	g	h	i	l
			s	t	u	x	m	n	o	p	q	r
M	*N*	↕	a	b	c	d	e	f	g	h	i	l
			r	s	t	u	x	m	n	o	p	q
O	*P*	↕	a	b	c	d	e	f	g	h	i	l
			q	r	s	t	u	x	m	n	o	p
Q	*R*	↕	a	b	c	d	e	f	g	h	i	l
			p	q	r	s	t	u	x	m	n	o
S	*T*	↕	a	b	c	d	e	f	g	h	i	l
			o	p	q	r	s	t	u	x	m	n
U	*X*	↕	a	b	c	d	e	f	g	h	i	l
			n	o	p	q	r	s	t	u	x	m

Рис. 69. Шифрование Порта для $\mathbb{Z}_{20}$ Дж. Б. и М. Ардженти

неоднократно для различных сообщений, наложение (разд. 19.1) может помочь взломать его. Таким образом, начальный ключ должен быть сопоставим по длине с сообщением — но тогда продолжение автоключа дальше не имеет смысла.

Следующий недостаток — распространение ошибок шифрования — является общей слабостью всех автоключевых методов.

Бэббидж повторно изобрел автоключ — в этот раз даже со смешанным алфавитом — и, хотя он поначалу думал, что тот невскрываем, позже нашел решения для частных случаев. Намного позже, в 1949 г. Шеннон отметил,

что рекуррентное шифрование ВИЖЕНЕРа эквивалентно шифрованию ВИЖЕНЕРа с периодом 2. Если открытый текст разделен на группы $a_1a_2a_3\ldots$ длины d и если D есть начальный ключ, то следующие тождества (по модулю N) имеют место для шифротекста $C_1C_2C_3\ldots$:

$$C_1 = a_1 + D, \quad C_i = a_i + a_{i-1} \qquad (i = 2, 3, \ldots)$$

и, таким образом, справедливы рекуррентные тождества

$$\begin{aligned} C_1 &= a_1 + D \\ C_2 - C_1 &= a_2 - D \\ C_3 - C_2 + C_1 &= a_3 + D \\ C_4 - C_3 + C_2 - C_1 &= a_4 - D \quad \text{и т. д.} \end{aligned}$$

Следовательно, последовательность

$$C_1,\ C_2 - C_1,\ C_3 - C_2 + C_1,\ C_4 - C_3 + C_2 - C_1,\ \ldots$$

может рассматриваться как многосимвольное шифрование Виженера с периодом 2, т. е. как дважды чередующееся многосимвольное сложение ЦЕЗАРЯ. Не помогает даже использование смешанного алфавита. Аналогичный результат справедлив и для рекуррентного шифрования БОФОРТа.

8.7.3. Функция открытого текста. Идея воздействия на процедуру порождения ключа шифровальной машины каким-либо скрытным способом с помощью открытого текста вновь появляется в патентной литературе примерно в 1920 г. («символ влияния», в приложении к патенту от 10 октября 1919 г. Арвида Герхарда Дамма, шведский патент 52279, патент США 1 502 376). Таким образом, начиная с телетайпных шифровальных машин T52d и T52e Сименса и SZ 42 Лоренца, (нерегулярное) перемещение шифровальных элементов могло далее запутываться (*«mit Klartextfunktion»*) и шифрование становилось практически непериодическим. Однако в случае передающих каналов с помехами это часто вело к проблеме «сдвига по фазе» в шифровании; функция открытого текста, поэтому, весьма облегчала жизнь британским дешифровальщикам и использовалась только в течение нескольких месяцев в конце 1944 г.

8.7.4. Потоковый шифр. Рекуррентное шифрование вида $c_i = f(p_i, p_{i-1})$, применявшееся Кардано и Виженером, является частным случаем современного потокового шифра (нем. *Stromchiffrierung*) $c_i = \mathbf{X}(p_i, k_i)$ непериодического шифрования, где бесконечный ключ k_i порождается конечным автоматом G, являющимся генератором ключей $k_i = G(k_{i-1}, p_{i-1})$ с начальным ключом k_1. В этом случае сложность ключей сосредоточена в G.

8.8. Индивидуальные, одноразовые ключи

8.8.1. Вернам. Тот факт, что рекуррентное шифрование не намного лучше квазинепериодического шифрования, все еще не исключает возможность того,

что в безопасной системе шифрования отправителю и получателю обеспечено теоретически неограниченное снабжение секретными ключами, каждый из которых является подлинно нерегулярным и не несущим никакого смысла, ключами случайными и используемыми только один раз, т. е. одноразовыми ключами (нем. жаргон *i-Wurm*). Вернам, по-видимому, выдвинул в 1918 г. эту идею случайно, но она быстро распространилась в период между двумя мировыми войнами; первые ее следы могут быть обнаружены в США, в Советском Союзе и в Германии.

8.8.2. Бесконечный и бессмысленный. Майор Джозеф Майборн, позднее генерал-майор и начальник войск связи армии США (в период 1937–1941 гг.) в 1918 г. принял во внимание предостережение 1914 г. Паркера Хитта — «никакое сообщение в зашифрованном виде не находится в безопасности, если ключевая последовательность не сопоставима по длине с самим сообщением» — и вместе с шагом шифрования Вернама (разд. 8.3) ввел понятие одноразового ключа (*одноразовая лента, одноразовый блокнот* (ОРБ)), таким образом, связывая эпитет «*бесконечный*» (*endless*) с термином Мохауса «*бессмысленный*» (см. разд. 8.3.2).

В Германии Кюнц, Шауфлер и Ланглотц в 1921 г. предложили блокнот из 50 листов, каждый из которых содержал 240 цифр (48 групп по пять цифр) для перешифрования числовых кодов. С 1926 г. немецкая шифровальная система *Auswartiges Amt* представляла собой перешифрование (см. разд. 9.2.1) с помощью одноразовых блокнотов.

Советский Союз также перешел на использование индивидуальных ключей в 1926 г. к большому несчастью Фетерлейна, специалиста по Советскому Союзу в британской службе M.I.1 (b). Советский Союз продолжал использование индивидуальных ключей; на вклейке О показан лист из шифроблокнота размером со спичечный коробок, найденный у русского шпиона.

Природа одноразовых ключей такова, что они должны уничтожаться немедленно после использования. В шифровальных машинах Вернамовского типа измельчение бумажной ленты (*шреддинг*) можно легко механизировать. Большую трудность на практике представляет обеспечение достаточным количеством ключей при интенсивном обмене сообщениями, в частности в критических ситуациях на поле боя. С этими трудностями способны справиться военные штабы, дипломатическая почта или секретные агенты, ведущие двустороннюю переписку, и в таких ситуациях одноразовые ключи действительно часто используются, при условии постоянного снабжения ключами.

8.8.3. Плохое исполнение. Последовательности символов или цифр индивидуального ключа должны демонстрировать отсутствие всякой регулярности, должны быть случайными. Хорошие стохастические источники дороги. Кан сделал следующее замечание по русским индивидуальным ключам:

«Интересно, что некоторые блокноты, по-видимому, произведены машинистками, а не машинным способом. В них видны перебои и подчистки — невозможная вещь для блокнотов, выполненных машинами. Более существенен статистический анализ цифр. Один такой блокнот, например, имел число групп, в которых цифры от 1 до 5 чередуются с цифрами от 6 до 0, вроде

18293, не менее чем в семь раз превышающее число групп со случайным набором цифр. Это наводит на мысль, что машинистка нажимала поочередно клавиши левой рукой (которой нажималась бы группа цифр от 1 до 5 на европейской машинке) и правой рукой (которой нажималась бы группа цифр от 6 до 0). Помимо этого, групп, начинающихся с малых цифр (от 1 до 5), вместо 1/2, которая ожидалась бы при случайном наборе, наблюдалось 3/4 от их общего количества, возможно, потому, что машинистка печатала пробелы правой рукой, когда начинала новую группу левой рукой. Меньше встречалось удвоенных и утроенных цифр, чем ожидается при случайном распределении цифр. Возможно, девушки, которым было приказано печатать наугад, считали что, некоторое количество дуплетов и триплетов должно попадаться в случайном тексте, но в целях конспирации, сокращали их число. Однако, несмотря на эти аномалии, имеется не слишком много цифрового материала, чтобы сделать криптоанализ возможным».

8.8.4. Стойкое шифрование. Если индивидуальный ключ поступает из стохастического источника, выдающего все символы независимо и с равной вероятностью, то здравый смысл говорит, что открытый текст, зашифрованный такой ключевой последовательностью, порождает «невзламываемый» шифротекст, т. е. шифрование является стойким. (Выражение использовалось Плинием Эрлом Чейсом уже в 1859 г.) Что это интуитивно означает, на первый взгляд, кажется, ясно; стоит также заметить, что в этой книге все криптоаналитические методы основаны на предпосылках, которые отсутствуют при стойком шифровании. Но это ничего не доказывает; фактически проблема состоит в том, чтобы дать точную рецептуру «стойкости», и непременно стохастической природы. До сих пор наиболее вразумительный рецепт, основанный на работе А. Н. Колмогорова, был дан в 1974 г. Г. И. Чайтиным. Следуя ему и Шнорру (1970 г.), мы будем говорить, что непериодическое шифрование, порожденное бесконечной последовательностью индексов, будет стойким, если имеет место следующее:

Для каждой конечной последовательности не существует алгоритмического описания короче перечисления элементов последовательности, т. е. никакая последовательность не может быть сжата в более короткое алгоритмическое описание.

Как следствие этого, никакая последовательность, сгенерированная машиной, т. е. фиксированным алгоритмом, не является стойкой. Алгоритмы в данном контексте должно понимать в универсальном смысле тезиса Черча. Таким образом, не годятся цифровые последовательности, которые характеризуют рациональные или вычислимые иррациональные числа (разд. 8.7). Числа, подобные $\sqrt{2}$, $\sqrt{5}$, $\sqrt{17}$, никак не подходят, поскольку они слишком легко могут быть угаданы. Множество невычислимых вещественных чисел, тем не менее, очень велико.

Не известно, определяет ли каждое невычислимое вещественное число стойкое шифрование.

8.8.5. Формирование стойких ключевых последовательностей. Физические процессы, используемые сегодня для порождения «истинно» случайных ключей, основаны на наложении несоизмеримых колебаний или на хаотических нелинейных системах. По-видимому, они более надежны, чем использовавшиеся примерно до 1950 г. шумовые эффекты вакуумных ламп и стабилитронов или показания счетчика Гейгера. Шум электронной лампы использовался в 1943 г. для формирования индивидуальных ключей в британской системе вернамовского шифрования ROCKEX, которая обслуживала радиообмен Англии с США на высшем правительственном уровне — приблизительно один миллион слов в день, или, в более современных терминах, объемом в четыре 3.5-дюймовых дискеты в день.

8.8.6. Практическая реализация. В 1944 г. государственный департамент США начал использовать SIGTOT, армейскую вернамовскую шифросистему с одноразовыми ключами для своих наиболее секретных дипломатических депеш. Армия также использовала пятироторную машину M-134-A (SIGMYC), чьи роторы перемещались одноразовой 5-дорожечной лентой. Система Вернама в январе 1943 г. была заменена на такую же роторную систему M-228 (SIGCUM), разработанную Фридманом. После нескольких дней эксплуатации Роулетт обнаружил слабость системы, и поэтому она была временно снята и заменена в апреле 1943 г. на улучшенную версию.

Только малая часть послевоенных проектов была системами с истинными одноразовыми ключами, среди них можно упомянуть машину Mi-544 компании Standard Elektrik Lorenz (Германия), и машины Хагелина T-52 и T-55 компании Crypto AG (Швейцария). У русских имелась своя машина с одноразовой перфолентой «АГАТ».

8.8.7. Ошибки в использовании. Эксплуатация одноразовых ключей порождает свои собственные философские проблемы. Эрих Хюттенхейн сообщал, что в МИДе Германии по правилам безопасности каждый лист одноразовых ключей из блокнота в сотню страниц должен был существовать только в одном оригинале и одной копии. Реально же делалось девять копий, которые распределялись в меняющемся порядке по пяти дипломатическим миссиям.

«Веноновское» вскрытие («Venona breaks») Ричардом Халлоком, Сесилом Джеймсом Филлипсом (1925–1998 гг.), Женевьевой Фейнштейн и Люсиль Кэмпбелл самых стойких советских шифросистем (начиная с ноября 1944 г.) также произошло благодаря случайному использованию одних и тех же одноразовых шифроблокнотов. Летом 1944 г. Филлипс выяснил, что первая пятизначная шифрогруппа является указателем ключа. Этот взлом шифра позднее стоил жизни советским шпионам Джулиусу и Этель Розенбергам и окончательно разоблачил Гарольда «Кима» Филби, Ги Берджесса и Дональда Дюарта МакЛейна, Клауса Фукса, Гарри Голда, Девида Гринглассa, Гарри К. Уайта и Пьера Кота как шпионов. С другой стороны, Советский Союз был предупрежден в 1946 г. Уильямом Вейсбэндом и в августе 1949 г. Филби, которые, возможно, заставили прекратить использование дубликатов ОРБ после 1949 г.

Явным нарушением идеи надежного шифрования является формирование ключевых последовательностей машиной. Если затем происходит компрометация шифротекст-шифротекст между такой системой и, скажем, системой, использующей периодическую ключевую последовательность, и если последняя система в должное время взломана, то одноразовый блокнот с предполагаемой стохастической ключевой последовательностью становится открытым. При наличии достаточного количества материала, машина, которая генерировала ключевую последовательность, может быть реконструирована. Подобное действительно произошло с немецкой дипломатической системой FLORADORA, реконструированной англичанами (см. разд. 9.2.1): в шифре *i-Wurm*, который использовался немецким МИДом, обнаружили регулярность, и в Блетчли Парк могли даже определять, какая машина была использована — согласно Р. У. Филби, который принимал участие во вскрытии. Найджел де Грей сообщал, что м-р Лоренц предлагал такую машину в 1932 г. английскому МИДу. Эрих Ланглотц с немецкой стороны еще не знал о доктрине Чайтина. А тесты могли лишь опровергать, но не доказывать случайность.

8.9. Обмен и управление ключами

8.9.1. Одиночные символы ключа служат для формирования или выбора (см. разд. 2.6) шагов шифрования в системе шифрования. Такая система может быть одноалфавитной или многоалфавитной: в любом случае шаг шифрования никогда не должен использоваться второй раз, если требуется высокая безопасность шифра.

В одноалфавитном случае шифрование, удовлетворяющее этому требованию, должно быть полиграммным с шириной блока, которая может покрыть все сообщение. Это создает большое практическое неудобство. Таким образом, приходится рассматривать многоалфавитное шифрование с меньшей шириной блока, в частности, односимвольное. Кроме того, строгое требование никогда не использовать шаг шифрования повторно можно ослабить до требования индивидуального одноразового ключа (разд. 8.8) — т. е. ключа, вообще никогда вновь не используемого — что показывает полное отсутствие любой закономерности в последовательности шагов шифрования и после этого уже гарантирует в смысле Чайтина и Колмогорова, что шифрование не будет взломано. Хотя перед 1930 г. в США, Германии, Советском Союзе и в других странах индивидуальные одноразовые ключи уже высоко оценивались применительно к специальным задачам, их практические недостатки вели к широко распространенной тенденции мириться с более слабой, лишь относительной безопасностью шифрования.

8.9.2. Не нужно специально подчеркивать (см. разд. 2.6.1), что согласование ключей между пользователями является специфической слабостью любой криптологической системы. На большом расстоянии управление безопасностью часто зависит (см. ниже) от надежности посыльных, которую трудно гарантировать, и от доступности, которую не менее трудно обеспечить.

Поэтому в истории криптологии предпринималось много попыток защитить само согласование ключей криптологическими средствами; возможно даже с помощью стеганографических мер.

Выполнение согласования ключей для некоторой криптологической системы в рамках самой этой системы может выглядеть привлекательным — тем более, если имеется полная уверенность в невскрываемости такой системы. Однако этого нужно избегать любой ценой, так как вскрытие информации, обслуживающей согласование ключей, может скомпрометировать всю систему. По крайней мере необходимо, как это делал немецкий ВМФ в конце войны с помощью таблицы биграмм, подвергать согласование ключей некоторому дополнительному шифрованию в системе другого вида.

Идея шифрования согласования ключей с помощью ключа сообщения путем указания начальной позиции некоторого механического ключевого генератора некоторое время не была широко известна и не распространялась, например, на коммерческую ЭНИГМУ 1923 г., но в то же время была принята на вооружение в трехроторных ЭНИГМАХ Рейхсвера и Вермахта. Согласование ключей было отправным пунктом для взлома молодыми польскими математиками зашифрованного немецкой ЭНИГМОЙ трафика, произошедшего в 1932 г. Немецкая сторона — исключая ВМФ — сильно недооценила возможности взломщиков и не посчитала необходимым сделать процедуру согласования ключей более сложной, неизменно ссылаясь на необходимость не перегружать каналы связи и возможности шифровальщиков.

Имеются экстравагантные методы обхода таких уязвимых мест согласования ключей, например, при помощи коммутирующих шифрований $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$:

$$\chi_i^{(1)}\chi_i^{(2)}x = \chi_i^{(2)}\chi_i^{(1)}x,$$

Например, при помощи шифрований ВИЖЕНЕРа или ВЕРНАМа. В этом случае отправитель зашифровывает свое открытое сообщение шифрованием $\mathbf{X}^{(1)}$ с ключом $k^{(1)}$, выбранным им наугад; получатель применяет $\mathbf{X}^{(2)}$ с ключом $k^{(2)}$, выбранным наугад, и *посылает эту новую криптограмму назад отправителю*. В силу коммутативности $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$ тот интерпретирует его как сообщение, им зашифрованное, и которое он может расшифровать с помощью своего ключа $k^{(1)}$. *Частично расшифрованное сообщение он посылает теперь получателю*, который, в свою очередь, интерпретирует его как сообщение, им зашифрованное, и которое он может расшифровать с помощью своего ключа $k^{(2)}$. Таким образом, он получает первоначальный открытый текст. Недостатком этого метода является необходимость в тройной передаче. Если сообщение короткое, это допустимо. Поэтому такой метод был бы хорош для передачи важной информации, вроде паролей или ключа, который нужно использовать впоследствии с некоторым другим методом шифрования. Так как в итоге ни сообщение, ни один из ключей не передаются в открытом виде, метод кажется безопасным. Однако дьявол уже тут как тут, как показывает следующий простой пример с двумя шифрованиями ВИЖЕНЕРа на $\mathbb{Z}_{26}$.

Отправитель А выбирает ключ $AQSID$, который не известен получателю.

Получатель B выбирает ключ $PZHAF$, который не известен отправителю. Открытый текст /image/ шифруется отправителем на ключе $AQSID$

$$\begin{array}{r} \text{i m a g e} \\ +\,A\;Q\;S\;I\;D \\ \hline \text{I C S O H} \end{array}$$

I C S O H отправляется получателю, тот шифрует его на ключе $PZHAF$:

$$\begin{array}{r} \text{I C S O H} \\ +\,P\;Z\;H\;A\;F \\ \hline \text{X B Z O M} \end{array}$$

X B Z O M отправляется назад отправителю, тот расшифровывает его с помощью ключа $AQSID$

$$\begin{array}{r} \text{X B Z O M} \\ -\,A\;Q\;S\;I\;D \\ \hline \text{X L H G J} \end{array}$$

X L H G J, наконец, отправляется назад получателю, тот расшифровывает его с помощью ключа $PZHAF$ и, таким образом, получает сообщение /image/.

$$\begin{array}{r} \text{X L H G J} \\ -\,P\;Z\;H\;A\;F \\ \hline \text{i m a g e} \end{array}$$

Сообщения X B Z O M и I C S O H посылаются по открытой линии связи, а их разность раскрывает ключ B:

$$\begin{array}{r} \text{X B Z O M} \\ -\,\text{I C S O H} \\ \hline P\;Z\;H\;A\;F \end{array}$$

(аналогично, X B Z O M и X L H G J раскрывают ключ A). Это означает, что расшифрование X L H G J с помощью этого ключа $PZHAF$ приведет к компрометации открытого текста /image/:

$$\begin{array}{r} \text{X L H G J} \\ -\,P\;Z\;H\;A\;F \\ \hline \text{i m a g e} \end{array}$$

Причина такой возможности взлома кроется в том, что ключи образуют группу относительно суперпозиции шифрований (см. разд. 9.1) и, более того, что является типичным для метода шифрования, основынным на циклической группе порядка 26. В этом случае шаги шифрования будут известны.

Защита против взлома имеется только в том случае, когда хотя бы один из двух процессов расшифрования сделан настолько трудным, что он является практически труднореализуемым. Это означает использование метода

шифрования, где знание ключа шифрования не достаточно для эффективного вычисления ключа расшифрования. Такая мысль была высказана в 1970 г. Джеймсом Х. Эллисом (ум. 1997 г.) и обнародована в 1998 г. британской Группой по безопасности электронных коммуникаций (*Communication-Electronics Security Group*). Но в таком случае получатель В мог бы публично объявить ключ для шифрования предназначенных ему сообщений, которые он и только он был бы в состоянии расшифровывать. Более того, первый и второй шаги метода могут быть опущены. Это порождает идею метода асимметричного шифрования, *опубликованного* в этой форме впервые в 1976 г. Уитфилдом Диффи и Мартином Э. Хеллманом (далее см. разд. 10.1.2).

8.9.3. Как только система связи включает большое количество узлов и соединений, «обслуживание ключа» должно быть расширено до «управления ключами». Безопасное распределение ключей становится наиболее трудной задачей в схеме управления ключами. Ключи при передаче должны быть защищены от перехвата. Ключи могут распределяться по безопасным путям курьерами (в основном дипломатами и военными) или заказной корреспонденцией (прежде всего коммерческими пользователями), в то время как телеграф, телефон, телефакс и Интернет в этой роли сомнительны. Более старые каналы, как правило, нельзя использовать непосредственно для передачи секретного сообщения, потому что они слишком медленны и, в большинстве случаев, слишком дороги. Часто их надежность ограничена их малой емкостью.

Кроме того, в хорошей системе управления ключами присутствует безопасный ввод ключей в шифросистему, с носителем ключей, устойчивым к внешним воздействиям и с устройством «аварийного стирания», также как и сертификация качества ключей.

Схемы управления ключами, которые включают регистрацию и распределение ключей, создают риск, который может быть уменьшен стеганографически аббревиатурной системой обозначений.

Следуя этой линии рассуждений и руководствуясь требованиями практики, необходимо рассмотреть иерархии ключей с различными уровнями защиты (системы ключей доступа, использование главных, вторичных и третичных ключей). Например, главный ключ может быть сгенерирован машиной, но так как машина сама может попасть в руки противника, возможно использование вторичного ключа, пригодного только для сравнительно короткого сообщения, скажем, не более 250 символов, и слегка защищенного системой, отличной от основной системы, но которая также не является невзламываемой. Поэтому используется третичный ключ, передаваемый безопасным способом, которым можно пользоваться в течение более длинного периода — скажем один день[5)].

[5)] В качестве примера для ЭНИГМЫ Вермахта согласно процедуре, которая держалась с 8 июля 1937 г. до 15 сентября 1938 г.: первичный ключ генерировался машиной; вторичный ключ сообщения (индикатор) определял начальное положение роторов для каждого сообщения; третичный *Tagesschlüssel* (рис. 51 а) включает порядок дисков, кольцевую установку, исходную установку параметров перекрестного подключения дисков («*Grundstellung*») и начальную топологию соединений (steckering). Однако первичная и вторичная системы шифрования были идентичны, а третичный ключ передавался курьером.

Например, алгоритм обмена ключей (KEA), созданный АНБ, использующий ключи длиной в 1024 бита, и рассекреченный в июне 1998 г. министерством обороны США, защищен трудностью вычисления дискретного логарифма, см. разд. 10.2.4.2.

Такие иерархические системы делают задачу управления ключами слишком сложной. Кроме того, они подвергаются риску постепенной атаки: компрометируется генератор ключей, компрометируется вторичный ключ, компрометируется система в целом. Это особенно опасно, если согласование ключей для вторичного ключа выполняется в пределах системы первого контура: единичный взлом может слишком легко привести к постоянному взлому.

Все правила управления ключами справедливы также для индивидуальных, одноразовых ключей. Для них естественным образом справедливо предостережение Хитта о том, что длина ключевой последовательности должна быть не меньше длины открытого текста. Но это исключает использование действительно невзламываемых систем во многих практических случаях. Поэтому они все более и более заменяются квазинепериодическими ключами, в действительности периодическими ключами с гарантируемыми периодами экстремальной длины и сертифицированными, как прошедшие строгие стохастические тесты (псевдослучайные ключи). Часто для сертификации этих ключей привлекаются сложные теоретико-числовые исследования. Это справедливо не только для открытых систем, но и для коммерческих систем шифрования, которые также следуют этой тенденции. Однако прогресс в технологии хранения данных может прекратить подобную практику. Легкие диски памяти (CD-ROM) с высокой плотностью записи (гигабайты на декаграмм) дают индивидуальным, одноразовым, случайным ключам новый шанс быть использованными в дипломатической, стратегической военной и коммерческой связи, где имеется реальная потребность в абсолютной невскрываемости.

Глава 9

Композиция различных методов

Напомним, что любой шифр $\mathbf{X}$: $V^* \to W^*$ обычно конечно порождается какой-либо криптосистемой M. Пусть M^* обозначает множество всех шифров, порожденных таким образом с помощью M. Метод шифрования S — это некоторое подмножество множества M^*. Через M^d обозначается подмножество периодических шифров с периодом d. Через M^∞ обозначается множество шифров с невычислимыми ключами.

Композиция двух шифров путем упорядоченного соединения их схем шифрования требует, чтобы пространство криптотекста первого метода шифрования совпадало с пространством открытого текста второго метода.

Любители склонны верить, что композиция двух методов более стойка для незаконного дешифрования, чем каждый из них. Но это не обязательно так. Второй метод может даже частично или полностью нейтрализовать действие первого. Покажем это на примере. Пусть S — простая подстановка, порожденная, как обычно, некоторым паролем, скажем, таким (которым вполне мог бы воспользоваться Базерье):

BASEDOWS DISEASE IS CURABLE (Базедова болезнь излечима, *англ.*)

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
B	A	S	E	D	O	W	I	C	U	R	L	F	G	H	J	K	M	N	P	Q	T	V	X	Y	Z

Она имеет четыре 1-цикла, два 2-цикла и один 18-цикл. Примененная дважды, она дает результат:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
A	B	N	D	E	H	V	C	S	Q	M	L	O	W	I	U	R	F	G	J	K	P	T	X	Y	Z

в котором восемь букв, включая частые гласные e и a, инвариантны.

9.1. Групповое свойство

Криптосистемы M, для которых $V \doteqdot W$, обладают тем свойством, что композиция двух схем шифрования не выводит за пределы M. В этом случае будем говорить, что такая криптосистема образует группу. Примерами являются группа $\mathcal{P}_{26}$ всех простых подстановочных схем шифрования над Z_{26} и группа $\mathcal{P}_{24}$ всех перестановок ширины 24.

Для других криптосистем с условием $V \doteqdot W$ это необязательно: например, множество одноцикловых схем подстановок не образует группу, поскольку групповая единица не может быть одноцикловой подстановкой. Примеры из разд. 7.2.4 показывают, что множество схем шифрования АЛЬБЕРТИ и множество схем шифрования РОТОР для некоторых первичных алфавитов не образуют группу. Композиция таких схем шифрования увеличивает комбинаторную сложность. Это оправдывает применение трех и четырех роторов в машине ENIGMA. В этом смысле групповое свойство может приносить ущерб.

9.1.1. Ключевые группы. Однако, если криптосистема M образует группу, то композиция двух элементов χ_s, $\chi_t \in M$ тоже является некоторым $\chi_u \in M$, где индекс u однозначно определяется индексами s и t: $u = s \bullet t$. Таким образом, $\chi_s(\chi_t(p)) = \chi_{s \bullet t}(p)$, где $\bullet$. $\bullet$ — групповая композиция ключевых символов, которые образуют ключевую группу.

Криптосистема с ключевой группой была названа Шенноном «чистой», а Саломаа назвал ее «замкнутой относительно композиции».

9.1.2. Композиция методов. Метод шифрования тоже может быть группой относительно композиции входящих в него шифров, как например группа всех линейных подстановок данной ширины n, группа всех многоалфавитных (односимвольных) подстановок данного периода d или группа всех блочных перестановок данной ширины n.

Композиция двух методов шифрования (произведение шифров), вообще говоря, приводит, к некоторому новому методу шифрования, хотя часто родственному исходным: например, композиция двух общих или линейных многоалфавитных методов шифрования с периодами d_1 и d_2 является общим или линейным многоалфавитным методом шифрования с периодом НОК(d_1, d_2); то же самое имеет место и для блочных перестановок ширины n_1 и n_2. Для подстановок это было подмечено еще Бэббиджем в 1854 г. В указанных случаях комбинаторная сложность тоже возрастает.

Иногда композиция двух методов шифрования коммутативна, подобно композиции группы всех простых подстановок с группой всех блочных перестановок данной ширины n. Если два метода шифрования, каждый из которых является группой, коммутируют между собой, то произведение шифров тоже образует группу (Шеннон: «Произведение двух чистых шифров, которые коммутируют, является чистым»).

9.1.3. Т 52. Схемы шифрования телетайпных шифромашин, созданных Сименсом, работали над $\mathbb{Z}_2^5$ и использовали композицию пентаграфических подстановок (схемы шифрования ВЕРНАМ, действующие на 5-битных кодовых группах), а также перестановок пяти бит (перестановки их позиций) —

в теоретико-групповых терминах — это некоторое подмножество гипероктаэдральной группы порядка $2^5 \cdot 5! = 3840$. Эти шифромашины были основаны на патенте, полученном 18 июля 1930 г. Джиппом и Россбергом.

Модели T 52a и T 52b использовались немецким флотом с 1931 г.; модель T 52c сначала использовалась Люфтваффе, а с середины 1941 г. стала использоваться всем Вермахтом (немцы называли ее *Geheimschreiben* (тайнописец, *нем.*, а англичане использовали кодовое наименование *Sturgeon* (осетр, *англ.*)). Выпускалось около 1000 таких машин в год.

Шифрование и расшифрование (т. е. «законное» дешифрование) производились десятью шифродисками w_s, каждый из которых оперировал бинарными выключателями i_s, со значениями $i_s = 0$ или $i_s = 1$, $s = 1, 2, \ldots, 10$. Пять дисков $w_1 \ldots w_5$ производили на множестве $\mathbb{Z}_2^5$ 32 подстановки BERNAM, остальные пять $w_6 \ldots w_{10}$ производили перестановки, порождаемые 2-циклами.

В T 52 это было множество $\{(12)^{i_6}(23)^{i_7}(34)^{i_8}(45)^{i_9}(51)^{i_10}\}$. Поскольку $(12)(23)(34)(45) = (23)(34)(45)(51) = (54321)$, а $(23)(34)(45) = (12)(23)(34)(45)(51) = (5432)$, то число различных среди этих перестановок равно 30. В общем, 10 дисков порождали 960 алфавитов. В модели T 52c, усовершенствованной под руководством Вюстенея (1899–1988 гг.), ключ сообщения легко можно было изменить, и в силу новой схемы там имелось лишь 15 разных подстановок и 16 различных перестановок, так что число используемых для сообщения алфавитов снижается до 240.

Движение шифродисков управлялось их зубцами (диски имели соответственно 47, 53, 59, 61, 64, 65, 67, 69, 71 и 73 зубцов). На каждом шаге все диски поворачивались на один зубец, что приводило к регулярному движению дисков с периодом $47 \cdot 53 \cdot 59 \cdot 61 \cdot 64 \cdot 65 \cdot 69 \cdot 71 \cdot 73 \approx 10^{18}$.

Модели T 52d и T 52e (введенные 1943 г. и 1944 г.) были соответственно вариантами T 52a и T 52b, отличаясь более нерегулярным, прерывистым движением дисков, которое поддерживалось *Klartextfunktion* (функцией открытого текста, *нем.*). T 52b отличалась от T 52a лишь улучшенным подавлением помех.

К концу войны Сименсом в нескольких копиях была построена телетайпная шифромашина T 43. Она использовала индивидуальный одноразовый ключ и вероятно была идентична машине, названной англичанами *thransher* (морская лисица, *англ.*).

9.1.4. SZ. Криптологически более простой была телетайпная шифромашина SZ 40, SZ 42, SZ 42a (английское кодовое название *tunny* (тунец, *англ.*)). Она строила лишь BERNAM подстановки, и соответственно шифр был взаимообратным.

В SZ 42 (Вклейка N) первая группа из пяти шифродисков с 41, 31, 29, 26 и 23 зубцами (эти диски англичане называли χ-дисками) оперировала схемами шифрования BERNAM на 5-битовых кодовых группах; на каждом шаге диски поворачивались на один зубец. Вторая группа из пяти дисков с 43, 47, 51, 53 и 59 зубцами (эти диски назывались ψ-дисками) оперировала аналогично схемам шифрования BERNAM на 5-битовых кодовых группах, диски рабо-

тали последовательно. Еще два диска (называемых движущими), служили лишь для организации нерегулярного движения. Один, с 61 зубцом, двигаясь вместе с χ-дисками, управлял движением другого, с 37 зубцами, который, в свою очередь, управлял одновременным (слабость!) движением ψ-дисков. Период превосходил 10^{19}. Все диски были снабжены случайным образом расположенными выступами, управляющими ВЕРНАМ выключателями и могли быть быстро приведены в произвольные начальные позиции.

9.1.5. Оливетти. Гораздо меньше известно о практическом применении телетайпных шифромашин, построенных Оливетти (итальянский патент 387 482 от 30 января 1941 г.), которые имели лишь пять шифродисков и два движущих диска, порождающих слабую нерегулярность.

9.2. Перешифрование

9.2.1. Перешифрование. Перешифрованием называется общий случай произведения шифров: т. е. когда буквенный или цифровой код шифруется снова. Для цифрового кода используется шифрование ВИЖЕНЕР над $\mathbb{Z}_{10}$, т. е. с $N = 10$; и соответственно применяется сложение по $\bmod 10$, т. е. «без переноса», которое можно было выполнять на сумматоре без переноса (разд. 8.3.3).

В 1780-х гг. Арнольд, британский шпион в Новых английских штатах, применял для перешифрования прибавление 7 по модулю 10 к кодовым группам, т. е. обычное сложение ЦЕЗАРЯ. Если вместо этого для групп ширины m, где m — криптоширина кода, некоторое число прибавляется по $\bmod 10^m$ (многосимвольное сложение ЦЕЗАРЯ), то говорят о «добавке». Использование добавок в связи с кодами стало широко практиковаться в XIX в., когда начали появляться коммерческие кодовые книги. Особый род двойного перешифрования возник следующим образом. Кодовые книги, которые содержат как числовые, так и буквенные кодовые группы (см., например, рис. 39), использовались для обратного перевода числового кода, полученного прибавлением добавки в буквенный код. Это особенно интересно, если добавка для легкости числового вычисления имеет весьма специальный вид, например, 02000. В 1876 г. Патрик был обвинен Конгрессом США в коррупции за использование такой системы. Американский ВМФ использовал эту систему в Испано-американской войне 1898 г., и она рассматривалась как наиболее надежная и совершенная кодовая система того времени при условии, что добавки меняются через достаточно короткие интервалы.

МИД германского Рейха с 1919 г. применял двойное перешифрование 5-значного числового кода («Deutsche Satzbuch»), в котором использовалось две добавки, каждая из которых покрывала шесть 5-значных групп. В распоряжении имелось 5000 таких добавок. Англичане в течение долгого времени безуспешно пытались взломать эту систему, получившую кодовое название GEC или FLORADORA. Однако в мае 1940 г. в немецком консульстве в Рейкьявике (Исландия) им удалось захватить немецкие шифродокументы, включающие две кодовые книги № 22 и № 46 и десять добавок. Сначала

это обеспечило незначительный прогресс при тестировании стереотипных выражений. Затем в 1942 г. британский консул в столице португальского Мозамбика случайно получил добавки для двух месяцев. И в 1944 г. Деннистон, «Билл» Филби и Феттерляйн, возобновивший деятельность в Блетчли Парк, совместно с Соломоном Кульбаком из Вашингтона добились полного взлома этой системы.

Для перешифрования может также использоваться перестановка: Ситтлер, один из наиболее удачливых создателей кодов рекомендовал четырехсимвольные перестановки в своих кодовых группах. Если эту перестановку фиксировать, то, тем не менее, эффект перешифрования получается такой же, как и при использовании нового кода (разумеется, не более надежного, чем старый). При таком использовании ширина применяемой перестановки и длина кода должны быть взаимно просты.

9.2.2. Потребность в перешифровании. В особо важных случаях — даты, хронометраж времени, координаты, имена и т. д., прямо рекомендуется композиция кода с каким-либо предпочтительно независимым методом перешифрования. Одним из примеров является перешифрование кода двухдольными биграммными подстановками — оно использовалось для 3-значного кода первой строки (*Schlüsselheft* — ключевая тетрадь, *нем.*) немецкой армией во время Первой мировой войны после марта 1918 г. Перешифрование с помощью машины ENIGMA применялось для решеток, связанных с картами (разд. 2.5.3), немецким ВМФ во время Второй мировой войны.

9.2.3. Штепсельный коммутатор. Перешифрование шифра машины ENIGMA было усовершенствовано введением штепсельного коммутатора (разд. 7.3.3). Подстановки были взаимообратными, но это было необязательно, так как, хотя произвольная подстановка и должна была сохранить взаимообратный характер ENIGMA, но это отменялось, однако, диагональным коммутатором Уэлчмена. Другие же криптоаналитические методы польских (простыни Зыгальского) или английских («бомбы» Тьюринга) ученых были нечувствительны к «штекерингу» и могли работать при произвольных подстановках штепсельных коммутаторов.

9.2.4. ADFGVX. Ранним примером произведения шифров с помощью перемешивания была система ADFGVX немецкой армии, изобретенная в Первую мировую войну лейтенантом Небелем (во время Второй мировой войны он был офицером связи в Люфтваффе) и введенная генералом Людендорфом на Западном фронте в 1918 г., с двудольной 6×6 подстановкой Полибия (разд. 3.3.1) в алфавите {A, D, F, G, V, X}, ясно отличимой от сигналов Морзе (разд. 2.5.2) и от перестановки ширины 20. Ключ менялся ежедневно, и это стоило французскому криптологу Пэнвину, по крайней мере, целого дня работы для дешифрования сообщений — если он вообще мог их дешифровать.

9.2.5. Перешифрование с помощью ENIGMA. В шифросети ENIGMA немецкого военно-морского флота сообщения особой важности проходили второе перешифрование с помощью ENIGMA; эти сообщения отделялись от основного текста словом offizier. Причина была криптологическая, и кроме того это делало информацию недоступной для рядового состава. Поздно вече-

ром 20 июля 1944 г. такой сигнал циркулировал по всем немецким кораблям и был дешифрован в Блетчли Парк:

OKMMM ANANA LLEXX EINSA TZJWA LKUER EJNUR DURCH OFFIZ IERZU ENTZI
FFERN OFFIZ IERJD ORAJD ERFUE HRERJ ADOLF HITLE RJIST TOTXD ERNEU
EFUEH RERIS TFELD MARSC HALLJ VONWI TZLEB ENJ...

(Ко всем! Конец! Обращение. Только для офицеров. Зашифровано. Фюрер Адольф Гитлер умер. Новый фюрер — фельдмаршал фон Вицлебен, *нем.*)

Вальтер Эйтан [Ettinghausen], дежурный вахты Z в корпусе 4 Блетчли Парка, не знал, насколько мрачны были неверные новости; во всяком случае он сохранил этот секрет от вездесущих юных леди, военнослужащих женской вспомогательной военно-морской службы.

9.3. Подобие методов шифрования

Шеннон назвал два класса $\mathcal{S}$ и $\mathcal{T}$ методов шифрования подобными, если существует (независимое от ключей) взаимно однозначное отображение A множества криптотекстовых слов из $\mathcal{T}$ в множество криптотекстовых слов из $\mathcal{S}$ такое, что

$$\text{для любого } T \in \mathcal{T} \text{ найдется } S \in \mathcal{S}, \text{ такое что } S = AT,$$
$$\text{т. е. } S(x) = AT(x) \text{ для всех } x.$$

Методы шифрования из подобных классов криптоаналитически эквивалентны: можно допустить, что отображение A известно (ведь предупреждение Керкхоффса в констатации Шеннона гласит: «Врагу известна используемая система»). Любой возможный путь для взлома $\mathcal{T}$ годится и для взлома $\mathcal{S}$.

Классами подобных методов шифрования являются, например, методы ЦЕЗАРЯ и обращенные методы ЦЕЗАРЯ (отображением A является «инвертирующая» подстановка (разд. 3.2.1) криптотекста), методы ВИЖЕНЕРа и методы БОФОРТа (A — инвертирующая подстановка криптотекста), методы простой колонной перестановки и методы блочной перестановки (A является матричной перестановкой криптотекста).

9.4. Шенноновское «перемешивание теста»

Композиция многодольной многосимвольной подстановки и перестановки некоммутативна. Композиция собственно многосимвольной подстановки ширины k и перестановки ширины k тоже некоммутативна. Некоммутативна и композиция простой подстановки и метода ВИЖЕНЕРа.

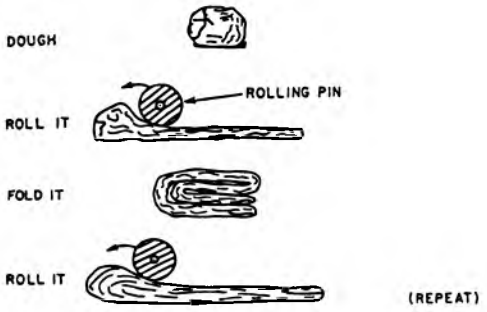


Рис. 70. Процесс перемешивания теста

Шеннон в 1945 г. подчеркивал, что композиция некоммутирующих методов шифрования работает подобно «перемешиванию теста»[1]. Подобные явления в компактных пространствах были изучены Хопфом (рис. 70)[2].

9.4.1. Перемешивание и рассеивание. Можно полагать, что композиция будет эффективной, если методы, участвующие в композиции, не только не коммутируют, но даже скорее не зависят друг от друга, подобно перестановке, рассеивающей шифруемый текст и линейной многосимвольной подстановке, выполняющей его перемешивание. Если произведение шифров не является

Рис. 71. Модулярное преобразование

группой, можно взять его итерацию, и тогда комбинаторная сложность увеличивается. Однако в дискретных пространствах шифров любая итерация любого фиксированного преобразования в конце концов становится периодической, в результате чего перемешивание Хопфа становится иллюзией. Это эффектно показано в следующем примере итерации двумерного преобразования фотографии, которая на первых шагах очень убедительно демонстрирует свой перемешивающий характер.

[1] N. J. A. Sloane. Encrypting by Random Rotations. Lecture Notes in Computer Science, Springer, 1990.

[2] Eberhard Hopf. On Causality, Statistics and Probability // Journal of Mathematics and Physics. — 1934. — v. 13, pp. 51–102.

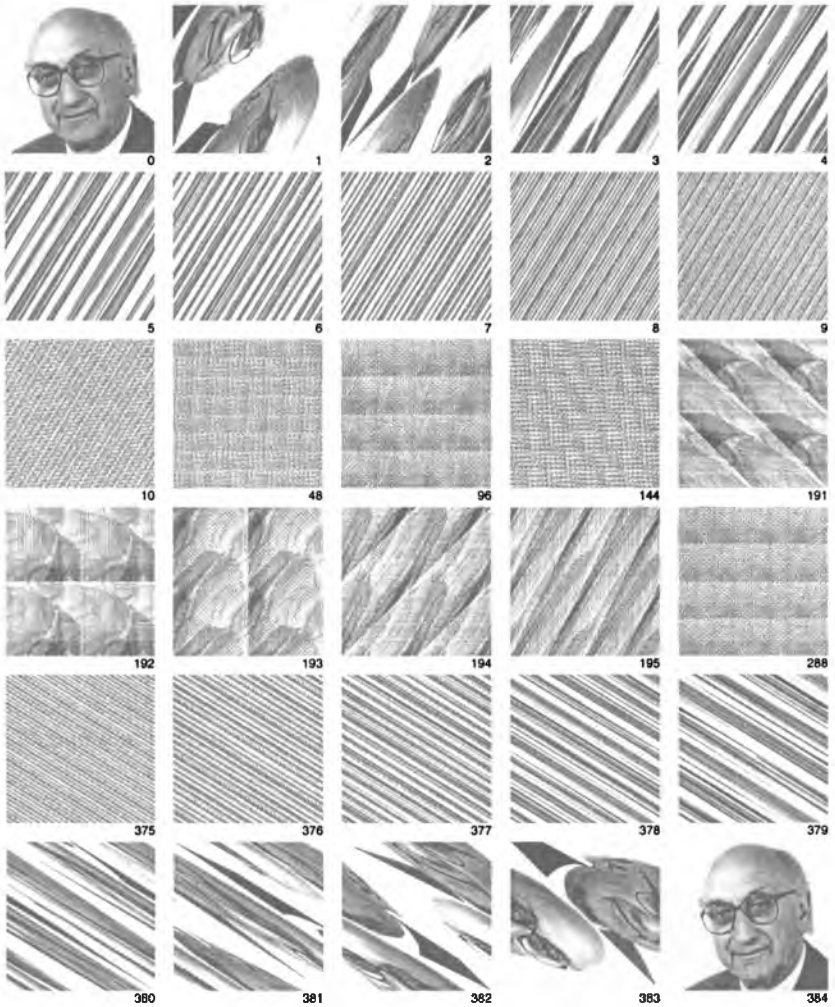


Рис. 72. Воскрешение FLB

Шаг рассматриваемого преобразования состоит из отражения с аффинным искажением, производимым приведением к основному формату путем обрезания выступающих углов и нового их приклеивания (рис. 71). Результат последовательных шагов преобразования приведен на рис. 72. Сначала это выглядит как будто портрет FLB[3)] собираются полностью перемешать, но после 48 шагов структура рисунка начинает меняться в обратную сторону и после 192 шагов получается четырехкратное подобие некоего призрака, а после 384 шагов восстанавливается исходная картина.

[3)] FLB = Friedrich L. Bauer, автор книги. — *Прим. перев.*

Рис. 73. Модулярное преобразование T

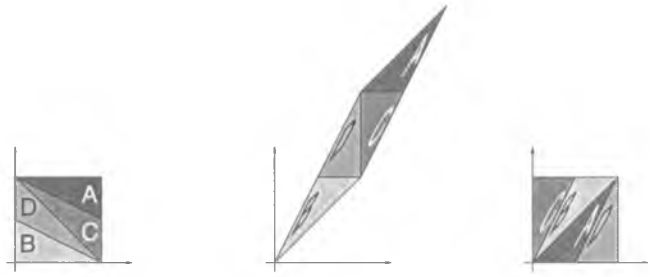


Рис. 74. Модулярное преобразование T^2

9.4.2. Эврика! Этот феномен можно объяснить. Рассмотрим квадрат Q: $0 \leqslant x < 1$, $0 \leqslant y < 1$ с тороидальной связью и на нем модулярное преобразование (рис. 73):

$$T: \begin{cases} x' & = y \\ y' & = \begin{cases} x+y-1, & \text{если} \quad x+y \geqslant 1, \\ x+y, & \text{если} \quad 0 \leqslant x+y < 1. \end{cases} \end{cases}$$

Локальное аффинное искажение, включающее отражение, задается матрицей

$$T = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

которая уже встречалась нам в разд. 8.6.1. Заметим, что

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix},$$

где F_n есть n-е число Фибоначчи[4)].

[4)] См. F. L. Bauer Efficient Solution of a Non-Monotonic Inverse Problem. (В книге W. H. J. Feijen Beauty is our Business. — Springer, 1990. — pp. 19–26.)

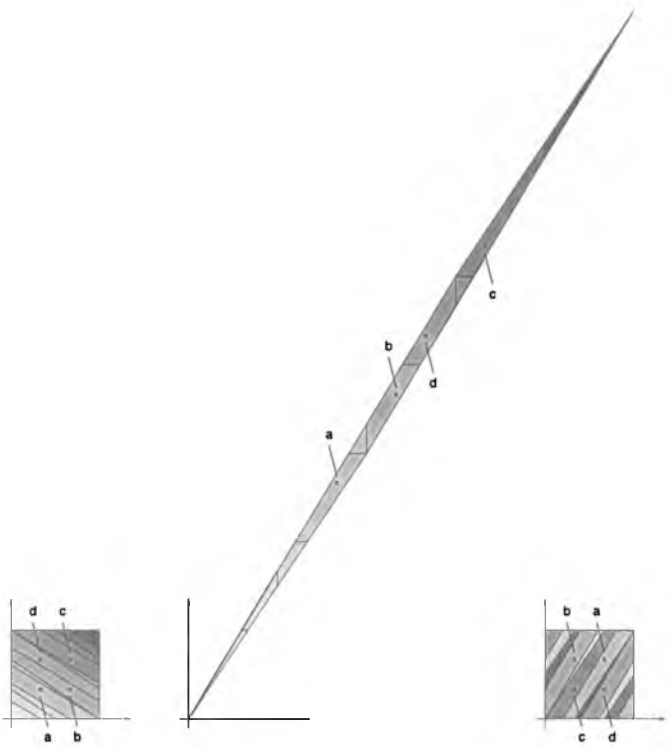


Рис. 75. Модулярное преобразование T^4

Рис. 74 показывает действие преобразования T^2 с локальным аффинным искажением, вызываемым матрицей

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

В заключение приводим рис. 75, показывающий действие преобразования T^4, где

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^4 = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} + 3 \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Здесь уже можно видеть, что шаблон из четырех точек

$$a = \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}, \quad b = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}, \quad c = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}, \quad d = \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix}$$

поворачивается на 180°. Следовательно, эти четыре точки будут неподвижными точками для преобразования T^8 с матрицей

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^8 = \begin{pmatrix} 13 & 21 \\ 21 & 34 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 3 \cdot \begin{pmatrix} 4 & 7 \\ 7 & 11 \end{pmatrix}.$$

Для T^{16} с матрицей

$$\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}^{16} = \begin{pmatrix}610 & 987\\ 987 & 1597\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + 21 \cdot \begin{pmatrix}29 & 47\\ 47 & 76\end{pmatrix}$$

появляются новые неподвижные точки; фактически,

$$\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}^{16} \begin{pmatrix}i/21\\ k/21\end{pmatrix} = \begin{pmatrix}i/21\\ k/21\end{pmatrix} + \begin{pmatrix}29i+47k\\ 47i+76k\end{pmatrix}.$$

Таким образом, все 400 точек с координатами $(i/21, k/21)$, $0<i<21$, $0<k<21$, являются неподвижными точками преобразования T^{16}. Преобразование T^{48} имеет матрицу

$$\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}^{48} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + 46\ 368 \cdot \begin{pmatrix}64\ 079 & 103\ 682\\ 103\ 682 & 167\ 761\end{pmatrix},$$

так что

$$\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}^{48} \begin{pmatrix}i/46368\\ k/46368\end{pmatrix} = \begin{pmatrix}i/46368\\ k/46368\end{pmatrix} + \begin{pmatrix}64\ 079i+103\ 682k\\ 103\ 682i+167\ 761k\end{pmatrix}.$$

В результате это дает $46367^2 = 199^2 \cdot 233^2 \approx 2.15 \cdot 10^9$ неподвижных точек. Вне этих точек происходит полное перемешивание. Однако, если перемешивание ограничено точками решетки, то понятно воскрешение портрета, когда множество неподвижных точек совпадает с множеством всех точек мозаики. Заметим, что $46368 = 2^5 \cdot 3^2 \cdot 7 \cdot 23$. Таким образом,

$$\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}^{48} (\mathrm{mod}\, 2^5) = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

Это значит, что в примере рис. 72 процесс перемешивания 32×32 мозаичных точек должен привести к их воскрешению через 48 шагов. В действительности перемешивание производилось с 256×256 мозаичными точками. Теперь получаем

$$\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}^{48} (\mathrm{mod}\, 2^8) = \begin{pmatrix}2\ 971\ 215\ 073 & 4\ 807\ 526\ 976\\ 4\ 807\ 526\ 976 & 7\ 778\ 742\ 049\end{pmatrix} \bmod 2^8 = \begin{pmatrix}225 & 64\\ 64 & 33\end{pmatrix},$$

$$\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}^{96} (\mathrm{mod}\, 2^8) = \begin{pmatrix}225 & 64\\ 64 & 33\end{pmatrix}^2 \bmod 2^8 = \begin{pmatrix}193 & 128\\ 128 & 65\end{pmatrix},$$

$$\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}^{192} (\mathrm{mod}\, 2^8) = \begin{pmatrix}193 & 128\\ 128 & 65\end{pmatrix}^2 \bmod 2^8 = \begin{pmatrix}129 & 0\\ 0 & 129\end{pmatrix},$$

$$\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}^{384} (\mathrm{mod}\, 2^8) = \begin{pmatrix}129 & 0\\ 0 & 129\end{pmatrix}^2 \bmod 2^8 = \begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}.$$

Данный результат дополняет таблицу, приведенную в разд. 8.6.1, а именно, $r(T,\ 256) = 384$.

Приведенный пример двумерного шифрования мог бы, конечно, быть выполнен и для шифрования текста при помощи перестановки.

9.4.3. Шеннон. Он рекомендовал совершенно общие композиции вида $\mathcal{SFT}$, где $\mathcal{S}$ и $\mathcal{T}$ — классы сравнительно простых методов, а F — некоторая (фиксированная) перестановка («барьер»), выполняющая тщательное перемешивание. В примере томографических методов (разд. 4.2) это должна быть перестановка, помещенная посередине, в примере колонной перестановки, перемешивающей строки (разд. 6.2.3) — матричная перестановка. В современных приложениях это может быть какой-нибудь чип, определяющий какое-либо семейство из 64 полиграфических подстановок шириной 64 бит.

Представляется уместным сделать следующее предупреждение: слепая вера в эффективность таких барьеров неоправдана, так как всегда существует опасность иллюзорной сложности. Кроме того, чем лучше перемешивание, тем сильнее локальная ошибка шифрования «размножается» на весь криптотекст. А самое худшее, что любая локальная ошибка в передаваемом криптотексте будет распространяться на весь дешифрованный открытый текст («лавинный» эффект), делая его полностью нечитабельным; со всеми скверными последствиями в случае повторения (гл. 11). В этом смысле хорошее перемешивание опасно.

9.4.4. Томографические методы. Они достигают простого, практичного и достаточно эффективного перемешивания с помощью применения многодольной подстановки, где барьером является некоторая специальная перестановка, а именно разрезание текста на куски и новая сборка кусков, т. е. перегруппировка промежуточного криптотекста, и последующего использования многосимвольной подстановки. Кажется, автором этой идеи является Рикети, граф де Мирабо (1748–1791 гг.), французский публицист и политик эпохи перед Французской революцией.

После двухдольной взаимно однозначной подстановки Полибия (см. разд. 3.3.1) $Z_{25} \to Z_5 \times Z_5$ он группировал вместе сначала все первые цифры, а затем все вторые цифры. Наконец, он применял обратную подстановку Полибия $Z_5 \times Z_5 \to Z_{25}$. (Предположительно это получается с помощью стеганографического метода — Базерье утверждал, что он расшифровал подлинные письма маркизы Софи де Монье ее известному любовнику графу де Мирабо, и, ссылаясь на их *teneur pornographique* (порнографическое содержание, *фр.*), не указал дальнейших подробностей, а привел лишь замечание в связи с Ботцелем и О'Кинаном, разд. 1.2.) Прямой и обратной подстановкой Полибия $Z_{25} \leftrightarrow Z_5 \times Z_5$ интересовался также юный Льюис Кэрролл (дневниковая запись от 26 февраля 1858 г.).

9.4.5. Полибий. Вместо фиксированной перестановки в качестве барьера можно также брать какое-либо семейство линейных преобразований, например, ВИЖЕНЕР или БОФОРТ над Z_5. Это (без обратной подстановки) является основной идеей «Шифра нигилиста», который может работать с периодическим или с бегущим ключом.

Кроме того, прямая и обратная подстановка Полибия использовалась в ранних шифромашинах B-21 и B-211 Хагелина. Для барьера Хагелин использовал два полуротора Дамма с десятью позициями, чтобы получить для каждого из двух Z_5 десять разных алфавитов (два сдвинутых квинтуплета);

в результате получается 100 разных алфавитов. Движение роторов выполнялось двумя парами шестеренок с 17, 19, 21 и 23 зубцами. В B-211 кроме того имелся штепсельный коммутатор.

9.4.6. Коэл. Томографический метод, предложенный Коэлом (разд. 4.2.3), согласно Гильдену, был основан на отображении $Z_{25} \to Z_{10} \times Z_{10}$. Деластель[5)] также рассматривал томографический метод более общий, чем его локальный метод, упомянутый в разд. 4.2.3: метод включал перегруппировку больших кусков, как это делал Мирабо. Для перегруппировки он предложил использовать параметр de sériation (длину серии, *фр.*). Например, равный 7 в следующем примере подстановки Полибия, порожденной паролем BORDEAUX:

e n v o y e z	u n b a t a i	l l o n i n f	a n t e r i e
1 4 5 1 5 1 5	2 4 1 2 5 2 3	4 4 1 4 3 4 2	2 4 5 1 1 3 1
5 3 3 2 4 5 5	2 3 1 1 2 1 3	1 1 2 3 3 3 5	1 3 2 5 3 3 5
14515155332455	24125232311213	44143421123335	24511311325335
D S S Z I C Z	C O T H G O R	P D J A O I K	C S R B H V K

Деластелль был достаточно практичен, чтобы выбрать длину серии, которая была бы не слишком большой: в противном случае ошибка шифрования распространилась бы на все сообщение. Однако опасность, исходящая от незаконного дешифровальщика, выше.

9.4.7. Другие методы. Деластелль рассматривал также томографические методы на основе трехдольной подстановки (разд. 4.1.3). Другие томографические методы используют тернарный код Морзе, например, метод, называемый POLLUX, а также обращение метода *Kulissenverfahren* (тайное действие, *нем.*) Охавера, которое в следующем примере использует шифр ширины 7:

	s	e	n	d	s	u	p
Символы Морзе	···	·	–·	–··	···	··–	·––·
Кодовая длина	3	1	2	3	3	3	4
Обращение	4	3	3	3	2	1	3
Перегруппированные	···	–·–	···	···	·–	·	––·
	H	K	S	S	A	E	G

9.5. Перемешивание и рассеивание арифметическими операциями

Тщательное перемешивание, в частности, производится арифметическими операциями. Один метод, любимый математиками, переоткрытый недавно, основан на произвольном одноалфавитном блочном шифровании сообщения как последовательности чисел, получаемых шифрованием каждого из этих чисел с помощью арифметических операций приведения по модулю некоторого подходящего числа q, возможно с последующим перешифрованием в буквенную форму («символьное сложение», «символьное умножение»).

[5)] Felix Marie Delastelle, 1840–1902. Verfasser des Traité Élementaire de Criptographie, Gauthier–Villar, Paris, 1902.

Сложение по mod q так же, как и умножение на множитель h по mod q, были рассмотрены выше в связи с линейными подстановками (разд. 5.7). Аналогично может быть образована r-я степень по mod q. Эти операции являются все более перемешивающими: умножение как итерационное сложение и возведение в степень как итерационное умножение. Мы увидим, что если существуют обратные операции, то они реализуют расшифрование (т. е. «законное» дешифрование) примерно с той же затратой усилий, какие были нужны для шифрования.

Для заданного числа q блок открытого текста может быть зашифрован выражением, чей числовой эквивалент x удовлетворяет условию $0 \leqslant x < q$. Шифр как число может быть даже представлен обычными числовыми кодами; это приводит к сжатию, которое может быть значительным для часто используемых текстов: стандартные коммерческие коды включают в среднем 8.5 букв открытого текста на одну 5-значную цифровую группу.

Можно использовать числовое представление в системе с любым основанием B, где $B \geqslant |V|$, которая удобна и в которой возможно быстрое выполнение арифметических операций. При $V = \mathbb{Z}_{26}$ часто использовалось основание $B = 100$, т. е., по существу, десятичная арифметика $(\mathbb{Z}_{10}^2)$ с цифровыми парами, или основание $B = 32$, т. е., по существу, бинарная арифметика $(\mathbb{Z}_2^5)$ с 5-битовыми группами. Сейчас обычно используются байты $(B = 256)$, 16-битовые группы $(B = 2^{16})$, 32-битовые группы $(B = 2^{32})$ и 64-битовые группы $(B = 2^{64})$.

9.5.1. Арифметика вычетов по модулю q. Некоторые предварительные сведения об умножении числа x на фиксированный множитель h по *модулю* q:

$$M_h(x) = x \cdot h \pmod{q}$$

были сделаны в разд. 5.7.

Для простого числа $q = p$ умножение *по модулю* q образует группу. Для каждого $h \neq 0 \pmod{q}$ существует обратное h' такое, что $h \cdot h' \pmod{p} = 1$, так что для умножения в поле Галуа $\mathbb{F}(p)$ (т. е. в конечном поле порядка p):

$$M_{h'}(M_h(x)) = x.$$

Для составного (т. е. непростого) q число h имеет обратное (и тогда h называется регулярным относительно q) тогда и только тогда, когда оно взаимно просто с q (разд. 5.6).

По техническим причинам q часто берется вида $q = 2^k$ или $q = 2^k - 1$, если вычисление производится в двоичной системе, и вида $q = 10^k$ или $q = 10^k - 1$, если в десятичной. В случае $q = 2^k$ лишь нечетные числа имеют обратные; в случае $q = 2^k - 1$ составных q следует избегать.

Для больших значений q определение обратного h' для числа h выглядит нетривиальным лишь на первый взгляд: существует простой алгоритм, являющийся аналогом алгоритма быстрого деления, который мы обычно применяем в позиционной системе для $\mathbb{Z}$, и который был приведен в разд. 5.7.1

для $\mathbb{Z}_q = \{0, 1, 2, 3, \ldots, q-1\} \subset \mathbb{Z}$. Рассмотрим этот алгоритм на следующем примере[6)].

$$17 \cdot h' \equiv 1 \pmod{1000}$$

$$\begin{array}{l} -1 \\ \left.\begin{array}{l} 16 \\ 33 \\ \mathbf{50} \end{array}\right\} 3 \\ \left.\begin{array}{l} 220 \\ 390 \\ 560 \\ 730 \\ 900 \end{array}\right\} 5 \\ \left.\begin{array}{l} 2600 \\ 4300 \\ 6099 \end{array}\right\} 3 \end{array}$$

$$17 \cdot 353 = 6001 \equiv 1 \pmod{1000}.$$

Можно легко запрограммировать микропроцессор, чтобы сделать это эффективно. Как только h' определено, для дешифрования требуется то же усилие, что и для зашифрования.

Если q является произведением двух различных простых чисел: $q = p' \cdot p''$, и если $h \cdot h'_1 \equiv 1 \pmod{p'}$ и $h \cdot h'_2 \equiv 1 \pmod{p''}$, то $h \cdot h' \equiv 1 \pmod{q}$, где $h' \equiv h'_1 \pmod{p'}$ и $h' \equiv h'_2 \pmod{p''}$.

Арифметика вычетов значительно уменьшает усилия, необходимые для определения обратного h' для h.

9.5.2. Возведение в степень. Для возведения числа x в фиксированную степень h *по модулю* q,

$$P_h(x) = x^h \pmod{q} \quad (x^0 = 1)$$

можно установить следующее.

Для простого $q = p$, если для некоторого h' имеем $h \cdot h' \equiv 1 \pmod{(p-1)}$ (тогда h взаимно просто с $p-1$), то $P_{h'}(x)$ является обратным для $P_h(x)$. Таким образом, для возведения в степень в поле Галуа $F(p)$

$$P_{h'} \cdot (P_h(x)) = x.$$

[6)] Для обращения числа по модулю другого числа в общем случае можно использовать алгоритм Евклида. — *Прим. ред.*

Доказательство. При подходящем α имеем:

$$\begin{aligned} P_{h'}(P_h(x)) &= x^{h\cdot h'} \bmod p = x^{h\cdot(h' \bmod p-1+\alpha\cdot(p-1))} \bmod p = \\ &= x^{h\cdot h' \bmod p-1} \cdot x^{h\alpha\cdot(p-1)} \bmod p = \\ &= x^1 \cdot (x^{p-1} \bmod p)^{h\alpha}. \end{aligned}$$

Но поскольку, в силу малой теоремы Ферма, $x^{p-1} (\mathrm{mod}\ p) = 1$, то окончательно

$$P_{h'}(P_h(x)) = x.$$

Примеры. Некоторые взаимно обратные пары можно найти в табл. 1 (разд. 5.5), например,

для $p = 11$: (3, 7) и (9, 9), $(N = 10)$,
для $p = 23$: (3, 15), (5, 9), (7, 19), (13, 17) и (21, 21), $(N = 22)$,
для $p = 31$: (7, 13), (17, 23), (11, 11), (19, 19) и (29, 29), $(N = 30)$.

Случай $p = 11$:

x^3 (mod 11) имеет цикловое представление (0) (1) (2867) (3549) (10),

x^9 (mod 11) имеет цикловое представление (0) (1) (26) (87) (34) (59) (10),

x, x^3 (mod 11), x^9(mod 11) и x^7(mod 11) образуют циклическую группу C_4 порядка 4.

Случай $p = 31$:

x^7 (mod 31) имеет цикловое представление A порядка 4:

(0)(1)(5)(25)(9 10 20 18)(17 12 24 3)(2 4 16 8)
(6)(26)(14 19 7 28)(22 21 11 13)(15 23 29 27)(30),

x^{11} (mod 31) имеет цикловое представление B порядка 2:

(0)(1)(5 25)(9 14)(10 19)(17 22)(12 21)(2)(16)(4)(8)
(6 26)(20 7)(18 28)(24 11)(3 13)(15)(29)(23)(27)(30),

x^{17} (mod 31) имеет цикловое представление AB порядка 4:

(0)(1)(5 25)(14 10 7 18)(22 12 11 3)(2 4 16 8)
(6 26)(9 19 20 28)(17 21 24 13)(15 23 29 27)(30),

x^{19} (mod 31) имеет цикловое представление A^2 порядка 2:

(0)(1)(5)(25)(9 20)(10 18)(17 4)(12 3)(2 16)(4 8)
(6)(26)(14 7)(19 28)(22 11)(21 13)(15 29)(23 27)(30),

$x^{29} \pmod{31}$ имеет цикловое представление A^2B порядка 2:

$$(0)(1)(5\ 25)(9\ 7)(10\ 28)(17\ 11)(12\ 13)(2\ 16)(4\ 8)$$
$$(6\ 26)(14\ 20)(19\ 18)(22\ 24)(21\ 3)(15\ 29)(23\ 27)(30).$$

$x^7 \pmod{31}$ и $x^{11} \pmod{31}$ порождают группу $C_4 \times C_2$ порядка 8.

Случай $p = 23$:

$x^7 \pmod{23}$ имеет цикловое представление

$$(0)(1)(2\ 13\ 9\ 4\ 8\ 12\ 16\ 18\ 6\ 3)(5\ 17\ 20\ 21\ 10\ 14\ 19\ 15\ 11\ 7)(22)$$

и порождает циклическую группу C_{10} порядка 10.

Вообще, для данного нечетного простого числа p множество $\{P_h\}$, где h регулярно относительно $(p-1)$, образует абелеву (т. е. коммутативную) группу M_{p-1}, зависящую от p. Возможны многоалфавитные шифры с такой ключевой группой. Эта группа имеет порядок $(p-1)/2-1$, если число $(p-1)/2$ простое. Простое число p, для которого $p' = (p-1)/2$ тоже является простым, называется надежным (safe) или сильным (strong) простым числом (Блэкли, 1978 г.). Надежными простыми числами являются 5, 7, 11, 23, 47, 59, 83, 107, 167, 179, 227, 263, 347, 359, 383, 467, 479, 503, 563, 587, 719, 839 863, 887, 983, 1019, 1187, 1283, 1307, 1319, 1367, 1439, 1487, ...; но имеются также и большие, например, $45 \cdot 2^{37} - 1$ или $10^{100} - 166517$. Кроме 5 и 7 все надежные простые числа имеют вид $12a - 1$.

$P_h(x)$ имеет тривиальную неподвижную точку $x = 0$ и две нормальные неподвижные точки $x = 1$ и $x = p - 1$, возможно есть и другие. Степени $P_h(x)$ и $P'_h(x)$ могут быть получены как произведения повторяющихся квадратов; двоичное представление h и h' показывает, как это должно быть сделано.

Для $p = 11$, поскольку $3_{10} = 11_2$ и $7_{10} = 111_2$; $9_{10} = 1001_2$:

$$P_3(x) = x \cdot x^2, \quad P_7(x) = x \cdot x^2 \cdot (x^2)^2, \quad P_9(x) = x \cdot ((x^2)^2)^2,$$

где $\cdot\cdot$ означает умножение, а $.^2$ возведение в квадрат, причем всякий раз *по модулю* 11.

Фактически, для n-битных чисел, при $2^n < p < 2^{n+1}$, возведение в степень *по модулю* p требует, грубо говоря, такого же усилия, как и выполнение n умножений. Учитывая тенденцию настоящего времени к использованию в схемах шифрования микропроцессорных чипов, следует заключить, что арифметические методы будут становиться все более важными в будущем.

Особенно это относится к простым числам вида $p = 2^{2^k} + 1$ (простые числа Ферма). Сейчас очень важной является проблема отыскания взаимно обратных пар *по модулю* 2^{2^k}, для чего существуют специальные решения.

Для составных чисел q ситуация более сложна. Частный случай, когда q является произведением двух (различных) простых чисел, $q = p' \cdot p''$, будет рассмотрен в разд. 10.3.

9.5.3. Двусторонняя связь. Так как h и h' в разд. 9.5.1 и разд. 9.5.2 взаимозаменяемы, то во взаимной связи двух партнеров A и B можно применять h как для шифрования, так и для расшифрования (дешифрования), и другое h', аналогично, как для шифрования, так и для расшифрования.

9.5.4. Чейз. Предвестником арифметических методов был Чейз; в 1859 г. он описал во вновь созданном журнале *Mathematical Monthly* следующий метод. После двухдольной инъективной подстановки $V \to W^2$, где $W = Z_{10}$, образуется число x, как это делал Мирабо (разд. 9.4.4), из первых знаков, и другое число y из вторых знаков. Затем выполним простые арифметические операции, скажем, умножим x на 7 и y на 9 и снова переведем результат в V. Эта простая система является более безопасной, чем многие привычные схемы, хотя и не находит практического применения.

9.6. DES и IDEA®

Алгоритм DES (Data Encryption Standard) был обнародован в 1977 г. Национальным бюро стандартов (National Bureau Standarts, NBS) США для использования с «несекретными компьютерными данными»[7]. Метод DES является блочным шифрованием для октограмм байтов. Последовательность фиксированных перестановок и зависящие от ключа многодольные нелинейные подстановки производят тщательное перемешивание. DES является томографическим методом; это можно лучше всего увидеть на примере оригинального метода шифрования LUCIFER Фейстеля, служащего компании IBM (рис. 76). Сразу возникает впечатление, что крестным отцом разработки является Шеннон (разд. 9.4.3). На короткой эффективной ключевой длине настояло АНБ (National Security Agency)

9.6.1. Алгоритм DES. Мы дадим лишь схему метода; с деталями можно ознакомиться в официальном источнике[8].

9.6.1.1. Шифрование. Принципиальное строение схемы шифрования DES показано на рис. 77. Восьмибайтный блок открытого текста сначала подвергается действию (независимой от ключей) исходной перестановки T, а потом расщепляется в два четырехбайтовых блока L_0 и R_0. Последующие 16 циклов ($i = 1, 2, \ldots, 16$) определяются рис. 77, где

$$L_i = R_{i-1} \quad \text{и} \quad R_i = L_{i-1} \oplus f(R_{i-1}, K_i).$$

Здесь символ $\oplus$ означает сложение *по модулю* 2, а K_i — 48-битный ключ, порожденный заданным ключом с помощью выбранной функции f. Заключительная перестановка T^{-1}, обратная T, завершает схему шифрования DES.

---

[7] Federal Information Processing Standards Publication 46, National Technical Information Service, Springfield, VA, April 1977. Federal Register, March 17, 1975 and August 1, 1975. For the presentation of background information (from the point of view of N.B.S.). См. Smid M. E., Branstad D. K.: *The Data Encryption Standard: Past and Future*, Proceedings of the IEEE, Vol. 76, № 5, May 1988.

[8] Efficient DES Key Search / CRYPTO '93, Santa Barbara, CA, Aug. 22–26, 1993.

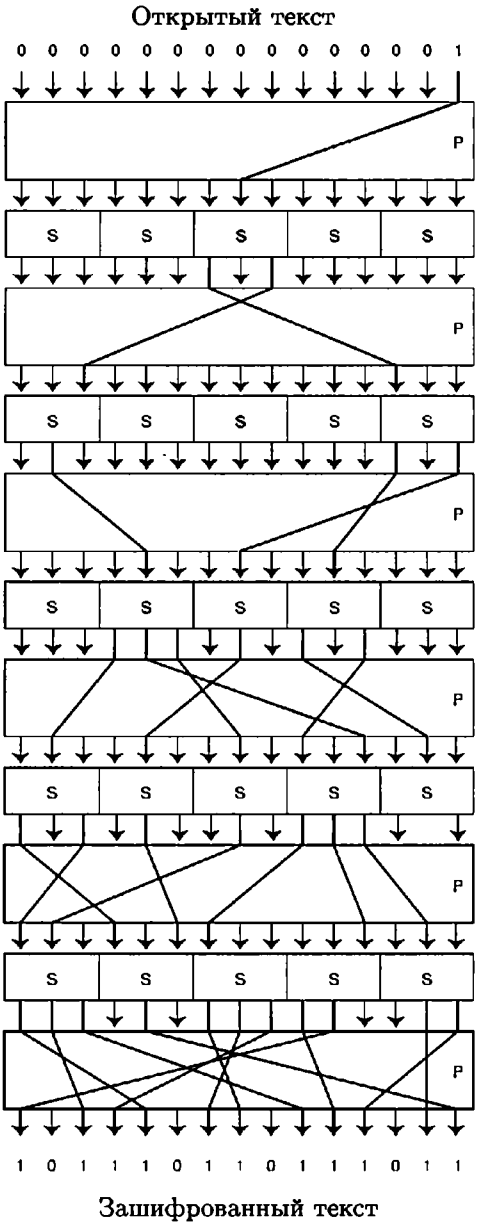


Рис. 76. Шифрование LUCIFER (Фейстель, 1973)

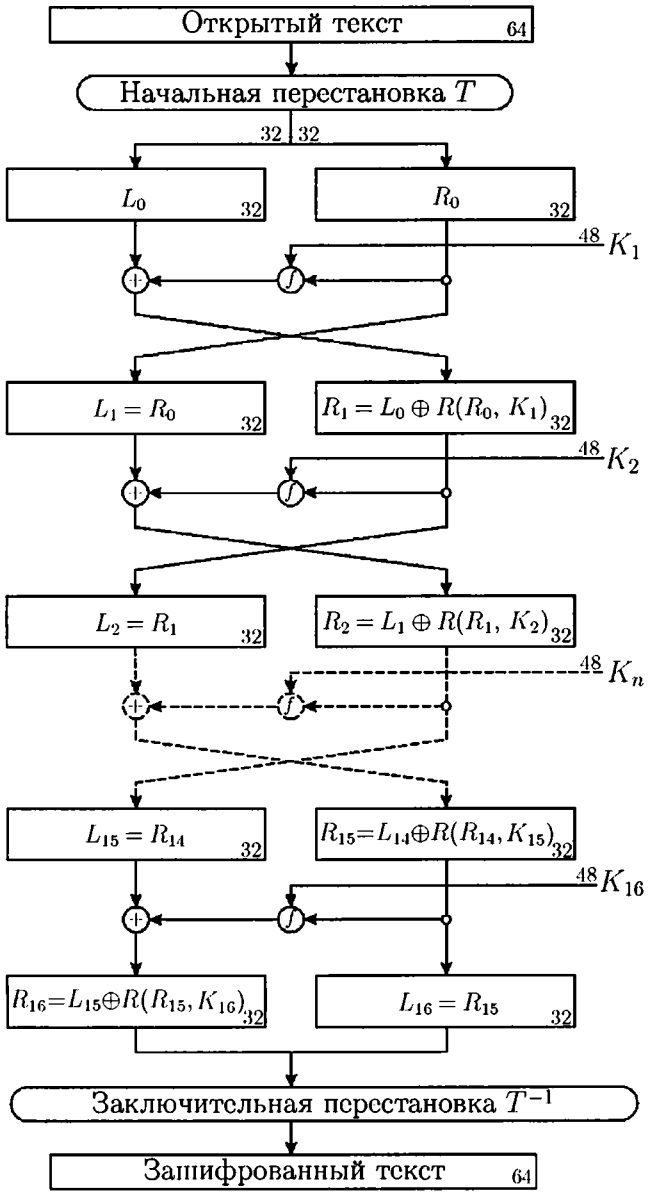


Рис. 77. Схема шифрования DES

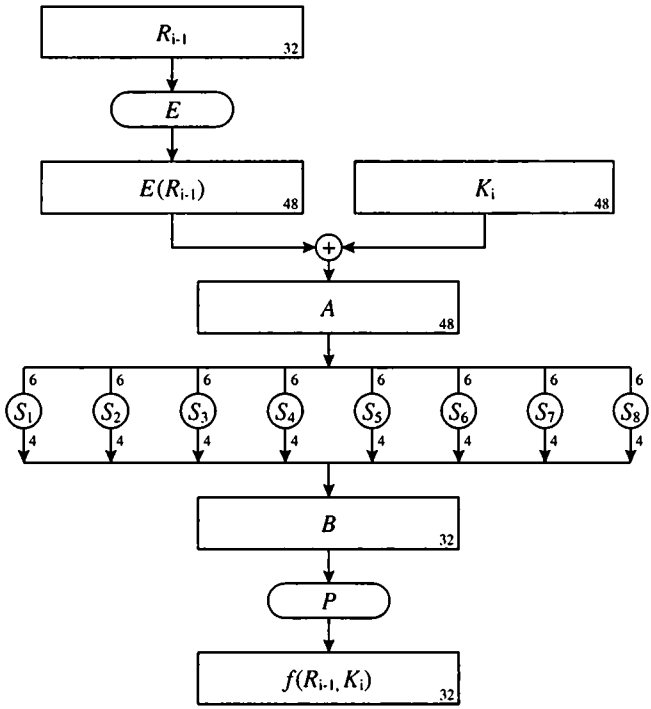


Рис. 78. Функция f

Функция f является центральной частью этого алгоритма (рис. 78). Блок R_{i-1} из 32 битов расширяется в 48-битный блок $E(R_{i-1})$ путем дублирования определенных битовых позиций и прибавления *по модулю* 2 к K_i. Полученный в результате 48-битный блок расщепляется на восемь 6-битных групп, служащих в качестве входов для каждого из восьми подстановочных модулей S_1, S_2, ..., S_8 («S-блоков»). Каждый из этих модулей осуществляет четыре различных нелинейных подстановки. В следующей таблице показаны эти подстановки для S_1.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
S_1: 1	0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
2	4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
3	15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13

Первый и шестой биты 6-битной группы интерпретируются как двоичное число, определяющее строку (а тем самым и подстановку), биты со 2-го по 5-й определяют столбец. Для входа 110010 (2-я строка, 9-й столбец в таблице) модуля S_1 выходом битовой группы является 12, т. е. 1100. Восемь 4-битных

выходных блоков подстановочных модулей S_1, S_2, ..., S_8 связаны воедино и подвергаются фиксированной заключительной перестановке P («P-блок»).

Остается вопрос о порождении подключей. Сначала в исходном ключе пользователем выделяются контрольные биты, которые удаляются (должно остаться 56 бит). Потом оставшиеся биты перемещаются согласно фиксированному предписанию и расщепляются на два 28-битных блока. Эти блоки циклически сдвигаются влево в каждом цикле на одну или больше позиций — в зависимости от номера цикла. Два из них (в соответствии с некоторым точно установленным правилом) порождают 48-битный подключ (K_i).

9.6.1.2. Расшифрование. Для расшифрования используется тот же алгоритм, но теперь подключи (K_i) применяются в обратном порядке. Этот алгоритм по-существу ключесимметричен: один и тот же ключ используется для шифрования и расшифрования. Циклы шифрования можно описать взаимно обратными отображениями

$$\begin{aligned} &h_i\colon (R,\, L) \mapsto (R,\, L + f(R,\, K_i)) && \text{(обработка)},\\ &g\colon (R,\, L) \mapsto (L,\, R) && \text{(обмен)}. \end{aligned}$$

Очевидно, отображение g инволютивно (т. е. взаимно обратно), что для отображения h_i вытекает из тождества

$$L \oplus f(R,\, K_i) + f(R,\, K_i) = L.$$

В то время как полное шифрование описывается равенством

$$\mathrm{DES} \equiv T^{-1} \circ h_{16} \circ g \circ h_{15} \circ g \circ \ldots \circ h_2 \circ g \circ h_1 \circ T$$

(в последнем цикле нет обмена), порядок подключей обращается при расшифровании:

$$\mathrm{DES}^{-1} \equiv T^{-1} \circ h_1 \circ g \circ h_2 \circ g \circ \ldots \circ h_{15} \circ g \circ h_{16} \circ T.$$

Поскольку все отображения взаимно обратны, композиция DES и DES^{-1} является тождественным отображением.

9.6.2. Лавинный эффект. Оказывается, что после нескольких циклов каждый бит промежуточного результата зависит от каждого бита открытого текста и от ключа. Минимальные изменения в открытом тексте или в ключе приводят к изменению около 50% бит («лавинный эффект»).

9.6.3. Способы применения DES. Для открытого текста, состоящего более чем из восьми байтов, блочное шифрование подразумевает разбиение открытого текста на 8-байтные блоки. В некоторых приложениях, например, если информация с каждым шагом становится все более полезной, и должна быть передана без задержки, открытый текст может быть и короче восьми байтов. Для обоих рассматриваемых случаев можно сравнить различные способы применения DES с учетом различий в скорости шифрования и размножении ошибок, и в результате решить, какой из способов является более предпочтительным.

Национальный институт стандартов и технологий (N.I.S.T., первоначально N.B.S.) стандартизировал четыре разных режима использования DES в США — по два для каждой из упомянутых выше ситуаций[9].

Режим ECB (*Electronic Code Book* — электронная кодовая книга, *англ.*) трактует все 8-байтные блоки как независимые. Совпадающие блоки открытого текста приводят к совпадающим блокам криптотекста. Этого режима с его строго одноалфавитным использованием алгоритма DES следует избегать насколько возможно.

Режим CBC (*Cipher Block Chaining*, криптоблоковые цепи, *англ.*) зависит от истории шифрования. Отправной точкой является начальный блок, который должен быть согласован между партнерами (ключ согласования c_0).

Шифрование блоков открытого текста m_1, m_2, m_3, ... приводит к следующим блокам криптотекста c_1, c_2, c_3, ..., где

$$c_1 = \mathrm{DES}(m_1 \oplus c_0), \quad c_2 = \mathrm{DES}(m_2 \oplus c_1), \quad c_3 = \mathrm{DES}(m_3 \oplus c_2).$$

Расшифрование производится следующим образом:

$$m_1 = \mathrm{DES}^{-1}(c_1) \oplus c_0, \quad m_2 = \mathrm{DES}^{-1}(c_2) \oplus c_1, \quad m_3 = \mathrm{DES}^{-1}(c_3) \oplus c_2.$$

Теперь метод шифрования многоалфавитный, но с довольно регулярным строением алфавитов, фактически это некоторый метод автоключа с повторяющимся ключом, защищаемым лишь нелинейностью барьера DES.

Кроме двух описанных режимов шифрований, существуют режимы, использующие алгоритм DES для порождения псевдослучайного ключа.

Режим CFB (*Cipher Feedback*, шифр с обратной связью, *англ.*) предлагает выбор 1-битного, 8-битного, 16-битного, 32-битного или 64-битного выходного сигнала для последующего использования с другим методом шифрования.

Режим OFB (*Output Feedback*, выход с обратной связью, *англ.*) имеет внутреннюю обратную связь, причем механизм обратной связи не зависит от потоков как открытого, так и шифрованного текста. Обычно этот режим применяется для производства 8-битного (64-битного) выхода. Он находит применение также при аутентификации.

9.6.4. Надежность DES. Первые описания DES, появившиеся в печати, вызвали дискуссию и критику, которые не утихали и после принятия этого стандарта. В частности, указывалось, что 16 внутренних циклов не могут обеспечить должный уровень безопасности. Другие важные недостатки DES состояли в следующем.

- Не были раскрыты критерии проектирования S-блоков — сначала не было никакой информации, позднее появились отрывочные сведения; правда, в конце концов полная информация была опубликована Копперсмитом в 1990-х гг. Но эти барьеры, существенные для надежности,

[9] DES Modes of Operation, National Bureau of Standards (US), Federal Information Processing Standards Publication 81, National Technical Information Service, Springfeld, VA, December 1980.

могли содержать «ловушки», делающие незаконное дешифрование легким (или по крайней мере облегчать его).

- Длина ключа относительно мала. Существует всего $Z = 2^{56} \approx 72 \cdot 10^{15}$ различных возможных ключей (для оригинального проекта LUCIFER — и это сравнение безусловно соответствует действительности — число ключей было намного больше, а именно порядка $2^{128} \approx 34 \cdot 10^{37}$).
- В режиме ECB ключ довольно долго сохраняется фиксированным; одноалфавитный характер шифра допускает классические атаки («построение глубины», разд. 19.1).

С другой стороны, DES является довольно быстрым алгоритмом шифрования. Увеличение ключевой длины и количества циклов, а также другие подобные вещи уменьшают скорость шифрования. Всемирное признание DES как стандарта *de facto* до некоторой степени оправдывает этот проект.

Однако в значительной мере возникшая дискуссия объяснялась глубоким недоверием общественности: American National Standards Institute (Американский национальный институт стандартов, *англ.*) подозревался (и даже обвинялся) в пособничестве АНБ, которое предполагалось заинтересованным во взломе шифров DES. Официальные сообщения не способствовали уменьшению этих подозрений.

Даже сегодня ни о каких «ловушках» публично неизвестно. С другой стороны, нет и никаких доказательств их отсутствия. Конечно, удивительные свойства алгоритма DES должны быть изучены. Например, симметрия относительно дополнительности: если открытый и ключевой тексты являются дополнительными, то и результирующий криптотекст тоже является дополнительным. Могут существовать и другие симметрии, не раскрытые до сих пор. Осадок недоверия остается. Фактически, приходится полагать, что при помощи все более и более быстрых машин государственная власть сможет взломать DES, если это окажется необходимым для национальной безопасности США. Частная инициатива не должна и, вероятнее всего, не способна это сделать.

Если предположить, что DES недостаточно надежен, например в режиме ECB, который широко используется в коммерческой сфере, то лекарством могло бы быть многократное шифрование с независимыми ключами. Однако при этом остается опасность иллюзорной сложности.

Верхний предел усилий по взлому метода шифрования определяется атакой грубой силы (т. е. перебором). Следует учитывать, что DES допускает неограниченное число опробываний, и таким образом может быть подвергнут этой атаке. В 1993 г. Бихам и Шамир нашли криптоаналитические контрмеры против методов перемешивания, используя для этого малые вариации открытого текста («разностный криптоанализ»). Оказалось, что для DES атака грубой силы в действительности может быть сокращена от 2^{56} тестов при полном переборе до 2^{47}, что, хотя и представляет лишь теоретический интерес, — но выглядит заманчиво[10)].

[10)] Подробности см. в работе Susan Landau, Notices AMS, Vol. 47, p. 341, p. 450.

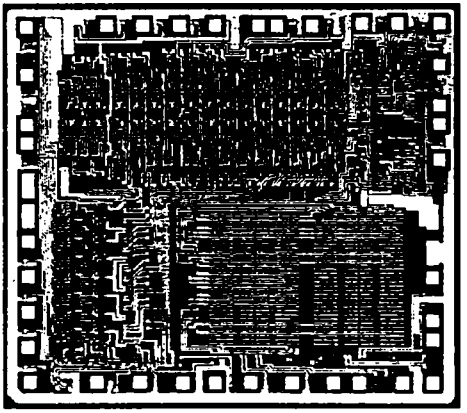


Рис. 79. DES-чип 1979 г.

Между тем, Копперсмит обнаружил, что не позднее 1974 г. проектировщики DES попытались сделать все возможное, чтобы предотвратить такую атаку.

Примерно с 1980 г. специальные чипы для выполнения DES попали на рынок. В первой половине 1990-х гг. подобные устройства производили шифрование и расшифрование со скоростью 10–20 Мбит/сек. На рис. 79 показан чип более раннего времени (1979 г.).

Метод DES стал всемирным лидером на рынке (безотносительно к американским экспортным ограничениям). Это должно было удовлетворить АНБ. Будут ли возможные преемники DES столь же успешными? Это предстоит увидеть.

9.6.5. Преемники для DES. Экспорт 40-битных алгоритмов RC2 и RC4 производства RSA Data Security, Inc., поступивших на рынок в 1993 г., был свободным. В связи с этим Пур заметил: «Если вы получили разрешение от США [экспортировать какой-либо алгоритм шифрования], это, скорее всего, означает, что его слишком легко дешифровать».

Широко распространено мнение (например, Хорак, 1996 г.: «DES близок к исчерпанию кредита доверия»), что 56-битный DES, якобы предназначенный лишь на десятилетие, скоро должен быть заменен. В конце концов, можно считать, что за двадцать лет с 1977 г. по 1997 г. максимальная достижимая скорость вычислений была увеличена примерно в 2^{10} раз, и в 1998 г. это было подтверждено успешным взломом грубой силы (перебором). Как следствие, N. I. S. T. (Национальный институт стандартов и технологий) пришел к заключению, что он больше не может поддерживать использование DES для многих приложений и рекомендовал временно использовать утроенный DES еще несколько лет, пока не будет закончен новый AES (Advanced Encryption Standard, улучшенный стандарт шифрования, *англ.*).

По сравнению с DES значительно надежнее выглядит алгоритм SKIPJACK, имеющий 80-битный ключ, работающий в 32 циклах. Этот алгоритм также симметричен относительно ключа.

Заметим, что в 1993 г. специальный компьютер (стоимостью в $100 000) осуществил полный перебор ключевого множества DES за 3.5 часа[11]. Подобного результата для алгоритма SKIPJACK следует ожидать только к 2029 г., а в 1993 г. на это потребовалось бы $26.2 \cdot 2^{20}$ лет, в 2005 г. — $6.55 \cdot 2^{10}$ лет, в 2017 г. — 1.64 г. Приведенные цифры справедливы для атаки простого перебора («грубая сила»). Любой же профессиональный дешифровальщик должен найти способы и средства сделать это быстрее.

Увы! Согласно американским законам, алгоритм SKIPJACK, доступный в виде специального устройства (MYK-78 в системе Clipper, программирование Mikotronx, Torrence, CA, USA, с производительностью до 20 Мбит/сек и стоимостью около $10) был до июня 1998 г. засекречен и недоступен в качестве средства программирования. Это, конечно, означает, что возможность применения этого алгоритма в компьютерных сетях и интернете была очень ограниченной. Та роль стандарта de facto, которую приобрел алгоритм DES, едва ли будет достигнута алгоритмом SKIPJACK даже после рассекречивания.

Европа, вообще более либеральная, чем США, не связывает себя действиями американского правительства. Среди различных независимых попыток создания преемника для DES, может быть, наиболее обещающей является IDEA[12] (*International Data Encryption Algorithm*), разрабатывавшийся Мэсси и другими с 1990 г. и запатентованный и зарегистрированный швейцарской компанией Ascom Tech AG, Solothurn. IDEA поступил в продажу с 1993 г. в виде VLSI чипа и как программное обеспечение без коммерческих ограничений. С ключом из 128 бит, IDEA в следующем веке[13] выдержит атаку грубой силы, хотя для других криптоаналитических методов, особенно тех, которые используют ошибки шифрования противника, этот алгоритм так же уязвим, как SKIPJACK или DES. Все эти алгоритмы шифрования работают с 8-байтными блоками открытого текста.

Иногда, однако, не все доступные биты используются криптологически. Так, в начале 1998 г. оказалось, что в 64-битном ключе, используемом повсеместно для защиты доступа в мобильные телефоны (D1, D2, E-plus) системы GSM последние 10 бит были постоянно установлены на 0. Соответственно трудоемкость атаки грубой силы сокращается в 1000 раз и время ее выполнения составляет от нескольких дней до нескольких часов.

9.6.6. Криптосистемы и крипточипы. Много волнений доставил один алгоритм, который в виде программы распространялся в июне и июле 1991 г. по международным компьютерным сетям и был частью системы PGP («Pretty Good Privacy», довольно хорошая секретность, *англ.*, в шутку называемой

[11] Винер (M. J. Wiener), Efficient DES Key Search, CRYPTO'93, Santa Barbara, CA, Aug. 22–26, 1993.

[12] IDEA — это зарегистрированная торговая марка.

[13] XXI веке.

также «Pretty Good Piracy», довольно хорошее пиратство). Так же, как и другие алгоритмы шифрования и расшифрования, PGP содержит средства для распространения секретных ключей и для аутентификации (разд. 10.5, 10.6). Алгоритм PGP за удивительно короткое время вырос в de facto стандарт электронной почты в интернете. PGP ускользнул от своего создателя Циммермана, а также проскользнул сквозь сито американского закона, — в очень большой степени из страха перед АНБ, но также в значительной мере из злорадного озорства международного анархического движения криптопанков. В январе 1996 г. американское правительство прекратило преследование Циммермана. Между тем, программное обеспечение PGP продавалось в США примерно за $ 100.

Как бы то ни было, национальные законы не встречают понимания в мире. Например, фирме Netscape, выпустившей в марте 1998 г. очередной браузер, для того, чтобы удовлетворить американским законам (ITAR, *International Traffic in Arms Regulations*), пришлось отказаться от использования SSL (*Secure sockets layer*) для безопасного обмена данными. После этого в течение какой-то пары дней в Австралии выпустили «cryptocilla», — версию браузера Netscape, снабженную 128-битными алгоритмами шифрования.

Как бы то ни было, основной слабостью всех существующих криптосистем является распределение ключей. Это урок, который польские криптоаналитики преподали немцам. Как только взломано распределение ключей, весь алгоритм шифрования становится бесполезным.

Тем временем микропроцессорные чипы становились все более и более мощными. Недавно (1996 г.) созданный 64-битный процессорный чип общего назначения Alpha-AXP (21164), произведенный DEC (Digital Equipment Corporation), работает с частотой 300 Мгц, содержит 9.3 миллиона транзисторов и выполняет $1.2 \cdot 10^9$ команд в секунду. Первоначально он был изготовлен по 0.5μ-технологии (1μ (микрон)=10^{-6}м), т. е. его проводники имели ширину порядка 0.0005 мм. К 1999 г. частота возросла; частота его преемника-процессора Alpha 21264 равна примерно 600 Мгц, он изготовлен по 0.35μ-технологии. DEC объявил о переходе на 0.25μ- и 0.18μ-технологии, которые позволяют достичь частоты 1 Ггц к 2000 г.

Ныне микропроцессоры широко используются в серверах, настольных компьютерах и ноутбуках и часто связаны в сети, и таким образом все более и более ощущается нужда в криптографической защите их данных. Швейцарская фирма Crypto AG, Zug предложила в 1996 г. криптоплату для автономных или сетевых настольных персональных компьютеров и ноутбуков, которая предусматривает идентификацию пользователя и контроль доступа к ЭВМ, шифрование содержимого жестких дисков, дискет, директорий и файлов со скоростью не менее 38 Мбит/сек. Она имеет защиту от воровства и неумелого пользования, хранит пароли и работает с индивидуально порожденными псевдослучайными ключами, используя симметричный блочный алгоритм шифрования. Управление ключами использует многоуровневую иерархию. Главные ключи, ключи зашифрованных данных, ключи файлов и ключи дис-

ков образуют числовое многообразие мощности $2^{124} = 2 \cdot 10^{37}$. Показанная на вклейке P, эта криптоплата размером 85 мм × 54 мм и всего 3.3 мм толщиной очень мала. Тем не менее, если она используется должным образом, то можно ожидать, что она может противостоять мощнейшему суперкомпьютеру в течение значительного времени.

Глава 10

Системы шифрования с открытыми ключами

В рассмотренных до сих пор многоалфавитных методах шифрования используются ключи для шифрования и расшифрования. В криптосистемах с взаимообратными схемами шифрования используется для этого, разумеется, один и тот же ключ. В общем случае имеется две возможности.

1) Применяется только один ключ. Один и тот же ключевой символ имеет одинаковое значение для шифрования и для расшифрования. Этот случай имеет место для алгоритма DES (разд. 9.6.1). При использовании криптомашин необходим переключатель, позволяющий сделать выбор между режимом шифрования и режимом расшифрования.

2) Применяется два ключа, — один для шифрования, другой для расшифрования. В этом случае криптомашина работает в одном режиме, однако определение ключа расшифрования по ключу шифрования требует применения сверхусилий.

В качестве примера рассмотрим классический метод шифрования ВИЖЕНЕР. Можно считать, что в этом методе, с одним и тем же ключом k, используются шифрование E и расшифрование D, заданные равенствами

$$E = \{\chi_{k_j}(x)\}: \quad \chi_{k_j}(x) = x + k_j \pmod{N^n},$$
$$D = \{\chi_{k_j}^{-1}(x)\}: \quad \chi_{k_j}^{-1}(x) = x - k_j \pmod{N^n},$$

т. е. имеет место случай (1).

С другой стороны, полагая

$$D = \{\chi_{k_j}^{-1}(x)\}: \quad \chi_{k_j}^{-1}(x) = x + (-k_j) \pmod{N^n},$$

т. е. $\chi_{k_j}^{-1} = \chi_{(k_j)^{-1}}$, где $(k_j)^{-1} = -k_j$, видим, что имеет место случай (2).

Получение $(k_j)^{-1}$ из (k_j) в этом примере довольно простое, и поскольку ключ расшифрования $(k_j)^{-1}$ должен оставаться секретным, то ключ шифрования (k_j) тоже должен быть секретным. Но если получение $(k_j)^{-1}$ из (k_j) так же сложно, как и взлом криптотекста любыми другими средствами, то ключ (k_j) шифрования может быть сделан открытым (публичным). В таком случае мы получим систему с открытым ключом шифрования. Удивительно, но такие криптосистемы существуют.

Возникает такой вопрос: будет ли система с открытым ключом шифрования (система с публичным ключом) давать какие-либо преимущества? И если да, то почему такая простая идея появилась так поздно — в середине 1970-х гг. — в длинной истории криптологии? Ответ состоит в том, что использование криптографических методов в коммерческих сетях имеет особенности, которые отсутствовали в классической ситуации двух партнеров. В самом деле, открытый ключ дает существенные преимущества в случае, когда большое число лиц участвует в обмене секретной информацией, а также в случае, когда аутентификация имеет такое же или даже большее значение, чем секретность, — это имеет место в различных ситуациях, возникающих при современных глобальных финансовых сделках.

10.1. Симметричные и асимметричные методы шифрования

10.1.1. Симметричные методы (методы частных ключей). Ключ, согласованный двумя партнерами, определяет в классических криптосистемах как схему шифрования, так и схему расшифрования достаточно простым образом, который является симметричным в том смысле, что для выполнения этих действий требуется выполнить одинаковый объем работ. Более того, обычно два партнера в разное время выступают в разных качествах — отправителя и получателя, в случаях, где шифрование и расшифрование коммутируют (разд. 2.6.2, 9.5.3), каждому требуется только один частный ключ партнера.

Симметричные методы частных ключей не прекратили своего существования и с приходом, примерно в 1950 г., электронного века. И метод DES, рассмотренный в разд. 9.6, который является самым известным современным методом блочного шифрования, принадлежит к этому классу; в нем имеется только один ключ (случай (1)), и шифрование отличается от расшифрования лишь порядком, в котором выполняются циклы, порождаемые этим ключом. Шифрование и расшифрование быстрые, их скорость в 1995 г. составляла около 20 Мбит/сек.

Криптоаналитическая надежность зависит от степени секретности используемого ключа. Более того, если пользователь считает, что возможный незаконный дешифровальщик, даже знающий класс применяемых методов, никогда не найдет этот ключ, — это почти ничего не говорит о том, что в реальных условиях никто не смог бы так зашифровать некое фальшивое сообщение, чтобы получатель мог расшифровать его, ничего не заподозрив.

Аутентификация не является проблемой, но лишь в том случае, пока гарантируется надежность (см. разд. 10.5). Однако существуют определенные недостатки:

(1) Отправитель сообщения не может доказать своему партнеру или третьему лицу, что это именно он послал данное сообщение. Этот недостаток является препятствием для передачи приказов и для финансовых сделок.

(2) Ключи должны сообщаться или передаваться по каналу, чья криптоаналитическая надежность намного выше, чем надежность канала, используемого для нормальных передач. Передача ключей по обычным каналам может оказаться невозможной.

(3) При большом числе партнеров, которым требуется надежная связь, число двусторонних каналов (а значит, и число ключей) становится очень большим. Например, для сети с n партнерами, каждый из которых должен обмениваться сообщениями с каждым, требуется $\binom{n}{2} = n \cdot (n-1)/2$ взаимообратных ключа или $n \cdot (n-1)$ симметричных ключей. При $n = 1000$ эти числа равны 499500 и 999000 соответственно.

10.1.2. Асимметричные методы (методы с открытыми ключами). Ключ расшифрования несимметричной криптосистемы, конечно, должен быть надежно защищен; но асимметрия может зайти так далеко, что ключ шифрования будет не только незащищенным, но даже открытым (публичный ключ). В двустороннем канале с партнерами $\mathcal{A}$ и $\mathcal{B}$ имеется теперь четыре ключа: открытый ключ шифрования для $\mathcal{B}$ и соответствующий частный ключ расшифрования для $\mathcal{A}$, открытый ключ шифрования для $\mathcal{A}$ и соответствующий частный ключ расшифрования для $\mathcal{B}$. Это вдвое больше прежнего числа. Но $\mathcal{A}$ может теперь получать сообщения от многих партнеров, каждый из которых знает и использует для своего шифрования публичный ключ получателя $\mathcal{A}$ и верит, что лишь законный получатель $\mathcal{A}$ может расшифровать сообщение. Таким образом, каждый партнер имеет открытый (публичный) ключ и частный ключ, полное число ключей для сети из 1000 партнеров уменьшается от почти миллиона до двух тысяч.

Это устраняет недостаток (3), отмеченный выше. Что касается (1), то решение этой проблемы будет дано ниже в разд. 10.5. Недостаток (2) так же исчезает, так как публичный ключ не должен передаваться; если же партнер $\mathcal{A}$ решит открыть канал связи с $\mathcal{B}$, он может это сделать, обратившись к справочнику с ключами всех участников.

Понятие системы шифрования с открытыми ключами было введено в 1976 г. Диффи и Хеллманом[1)].

10.1.3. Шифрование и цифровая подпись. Пусть KP_i обозначает публичный ключ i-го партнера, а KC_i — его частный ключ. Ключ KP_i определяет шифрование E_i, а KC_i определяет расшифрование D_i. Как E_i, так и D_i эффективно выполняются, но если собрание ключей $\{KP_i\}$ представляет собой некий публичный справочник, тогда как ключ KC_i известен лишь i-му парт-

[1)] В статье New Directions in Cryhtography (Новые направления в криптографии, англ.), IEEE Transactions on Information Theory, IT-22, v. 6, pp. 644–654, 1976.

неру. Ни один из партнеров (кроме i-го) не может получить KC_i из KP_i (точнее, для этого потребуется слишком большие усилия).

Если E_i и D_i обладают свойством

$$D_i(E_i(x)) = x, \qquad (*)$$

мы говорим об (асимметричном) методе шифрования, обеспечивающем секретность.

Если, кроме того, E_i и D_i обладают свойством

$$E_i(D_i(x)) = x, \qquad (**)$$

мы говорим об (асимметричном методе) цифровой подписи, обеспечивающей аутентификацию.

Асимметричный метод шифрования работает следующим образом: если партнер $\mathcal{A}$ хочет послать некоторое зашифрованное сообщение m партнеру $\mathcal{B}$, он выбирает из справочника под рубрикой $\mathcal{B}$ ключ $KP_{\mathcal{B}}$ (это определяет шифр $E_{\mathcal{B}}$:

$$c = E_{\mathcal{B}}(m) \qquad (A)$$

и посылает криптотекст c по открытому каналу партнеру $\mathcal{B}$.

Партнер $\mathcal{B}$ применяет свой частный ключ $KC_{\mathcal{B}}$ (который определяет $D_{\mathcal{B}}$), чтобы расшифровать сообщение m:

$$D_{\mathcal{B}}(c) = D_{\mathcal{B}}(E_{\mathcal{B}}(m)) = m \quad \text{(в силу } (*)\text{)}. \qquad (B)$$

Асимметричное шифрование и цифровая подпись работают следующим образом: если партнер $\mathcal{A}$ хочет послать зашифрованное сообщение m, подписанное его подписью «A», партнеру $\mathcal{B}$, он сперва шифрует m *своим собственным ключом* $KC_{\mathcal{A}}$ (это определяет $D_{\mathcal{A}}$):

$$d = D_{\mathcal{A}}(m), \qquad (A1)$$

и присоединяет к d свою подпись «A». Затем он берет из справочника под рубрикой $\mathcal{B}$ ключ $KP_{\mathcal{B}}$ (это определяет $E_{\mathcal{B}}$) и шифрует пару («A», d):

$$e = E_{\mathcal{B}}(\text{«}A\text{»}, d) = E_{\mathcal{B}}(\text{«}A\text{»}, D_{\mathcal{A}}(m)). \qquad (A2)$$

$\mathcal{A}$ посылает криптотекст e по открытому каналу своему партнеру $\mathcal{B}$. Получив его, $\mathcal{B}$ применяет свой частный ключ $KC_{\mathcal{B}}$ (который определяет $D_{\mathcal{B}}$), чтобы расшифровать полученную пару:

$$D_{\mathcal{B}}(e) = D_{\mathcal{B}}(E_{\mathcal{B}}(\text{«}A\text{»}, d)) = (\text{«}A\text{»}, d) \quad \text{(в силу } (*)\text{)}. \qquad (B1)$$

$\mathcal{B}$ узнает из части «A», что отправителем является $\mathcal{A}$, и тогда применяет его публичный ключ $KP_{\mathcal{A}}$ (который определяет $E_{\mathcal{A}}$ для того, чтобы получить из d сообщение m:

$$E_{\mathcal{A}}(d) = E_{\mathcal{A}}(D_{\mathcal{A}}(m)) = m \quad \text{(в силу } (**)\text{)}. \qquad (B2)$$

Получив содержательный текст, $\mathcal{B}$ убеждается, что он получил сообщение от $\mathcal{A}$, так как ни один другой партнер не мог зашифровать это ключом $D_{\mathcal{A}}$.

10.2. Однонаправленные функции

Весь успех асимметричных систем шифрования с открытым ключом основан на вопросе: как можно добиться того, чтобы D_i, т. е. E_i^{-1} нельзя было легко получить из E_i, и, следовательно, чтобы взлом этого шифра был практически очень труден?

10.2.1. Строго однонаправленные функции. Инъективная функция $f\colon X \to Y$ называется строго однонаправленной, если выполняется следующее.

Существует эффективный[2)] метод вычисления $f(x)$ для всех $x \in X$, но не существует эффективного метода для вычисления x из соотношения $y = f(x)$ для всех $y \in f[X]$.

Саломаа дал замечательный пример однонаправленной функции. Шифрование с омофонными схемами шифрования $Z_{26} \to Z_{10}^7$ определяется так: для некоторой буквы X в телефонном справочнике крупного города Z ищется какое-нибудь имя, начинающееся с буквы X, и в качестве криптотекста берется 7-значный телефонный номер, выписанный напротив этого имени. Например, для шифрования слова /kindergarten/ (детский сад, *нем.*) делаются следующие шаги:

k ↦ Koch	↦ 8202310	i ↦ Ivanisevic	↦ 8119896
n ↦ Nadler	↦ 6926286	d ↦ Dicklberger	↦ 5702035
e ↦ Esau	↦ 8348578	r ↦ Remy	↦ 7256575
g ↦ Geith	↦ 2730661	a ↦ Aranyi-Gabor	↦ 2603760
r ↦ Rexroth	↦ 5328563	t ↦ Tecins	↦ 6703008
e ↦ Eisenhauer	↦ 7913174	n ↦ Neunzig	↦ 3002123

В результате слово /kindergarten/ шифруется последовательностью из 12 семизначных кодовых групп

8202310 8119896 6926286 5702035 8348578 7256575 2730661 2603760
5328563 6703008 7913174 3002123

оператором-человеком менее чем за минуту. Дешифрование однозначно, но потребуются многие часы, если делать это с помощью телефонного справочника с 2000 страниц. Однонаправленная функция, использованная для шифрования, ставит для законного дешифровальщика непреодолимые трудности. Поэтому строго однонаправленные функции нельзя использовать разумным образом для шифрования сообщений, так как шифрование должно сопровождаться расшифрованием. Однако строго однонаправленные функции без омофонов можно использовать для идентификации и аутентификации: пароль шифруется строго однонаправленной функцией и хранится в таком виде. И всякий раз, когда требуется произвести машинный поиск, представленный пароль шифруется (этой однонаправленной функцией) и оба криптотекста

[2)] Эффективным считается вычислительное усилие, которое имеет полиномиальную сложность относительно $\log |X|$. Если же выполняется условие $P = NP$, то ни одна строго однонаправленная функция вообще не может существовать.

сравниваются. Эта схема употребляется в широко используемой операционной системе UNIX[3)], но основывается она всего на одном варианте алгоритма DES, который не квалифицируется как строго однонаправленная функция.

Однако существует возможность, когда легальный пользователь этой системы обладает обратным телефонным справочником — либо полученным нелегально из почтового офиса, либо приобретенным за деньги. Такой справочник делает процесс расшифрования столь же простым, как и шифрование. Это является своего рода скрытой приостановкой одностороннего направления функции подобно следующему секрету: человек не может пройти назад через захлопнувшуюся дверь, но посвященный знает, где найти спрятанную кнопку, чтобы ее открыть.

10.2.2. Однонаправленные функции с секретом. Для обеспечения безопасности данных необходима однонаправленная функция с секретом, которая позволяла бы законному пользователю иметь доступ к данным путем расшифрования[4)].

Инъективная функция $f : X \to Y$ называется однонаправленной функцией с секретом, если:

- существует эффективный метод вычисления $f(x)$ для всех $x \in X$;
- существует эффективный метод вычисления $f^{-1}(y)$ для всех $y \in f[X]$, но он не может быть получен эффективно из соотношения $y = f(x)$ для всех $y \in f[X]$: необходима дополнительная секретная информация, «секрет».

Секретом в примере Саломаа является обратный справочник. Он может быть создан пользователем, который может найти для этого время, если испытывает в нем частую необходимость (или если он может купить его), и если сложность хранения не препятствует этому. Такая предварительная обработка информации является одной из лучших стратегий взлома систем асимметрического шифрования, если только эти системы функционируют достаточно долгое время.

10.2.3. Граница эффективности. Трудности с обращением однонаправленных функций возникают только из-за недостатка времени и памяти, необходимых для выполнения вычислений. Но технологический прогресс постоянно сдвигает пограничную линию между «неподатливостью» и «эффективностью»; в настоящее время, грубо говоря, каждые два года скорость самого быстрого компьютера удваивается, и в среднем за 15 месяцев стоимость компьютера сокращается вдвое[5)]. Эта последняя тенденция позволяет привлекать параллелизм. Криптолог пытается противодействовать техническому прогрессу подходящим увеличением некоторых параметров шифрования, и таким образом не дает криптоаналитикам преодолеть этот барьер. Приведем пример. Для некоторых методов обращение однонаправленной функции равносильно разложению некоторого числа n на простые сомножители, что

[3)] UNIX — зарегистрированная торговая марка.

[4)] Так же используются термины «ключевые однонаправленные функции», «однонаправленные функции с ловушкой». — *Прим. ред.*

[5)] Скорее следует говорить о снижении удельной стоимости вычислений. — *Прим. ред.*

требует огромной работы по сравнению с перемножением этих сомножителей. Один из быстрейших известных алгоритмов, «Квадратичное решето» («Quadratic Sieve», *англ.*)[6] имеет «субэкспоненциальную сложность», т. е. для его выполнения требуется порядка

$$e^{\sqrt{\ln n \cdot \ln(\ln n)}} = n^{\sqrt{\ln(\ln n)/\ln n}}$$

операций. Для $n = 10^{70}$ это число равно $2.69 \cdot 10^{12}$. В 1984 г. факторизация числа $(10^{71} - 1)/9$ потребовала 9.5 часов работы суперкомпьютера CRAY X-MP. Грубо говоря, работа равносильна $80 \cdot 10^6$ «макро-компьютерных шагам» в секунду. Экстраполируя этот факт, получим следующую таблицу (в предположении, что каждые два года быстродействие отдельного компьютера удваивается):

n	биты	$e^{\sqrt{\ln n \cdot \ln(\ln n)}}$	время 1984 г.	время 1994 г.	время 2004 г.
10^{50}	166	$1.42 \cdot 10^{10}$	181 с.	5.66 с.	181 мс.
10^{70}	232	$2.69 \cdot 10^{12}$	9.5 ч.	0.297 ч.	34.2 с.
10^{100}	332	$2.34 \cdot 10^{15}$	344 л.	10.75 дн.	8.26 ч.
10^{120}	399	$1.31 \cdot 10^{17}$	52.57 л.	600 дн.	19.3 дн.
10^{140}	465	$5.49 \cdot 10^{18}$	$2.2 \cdot 10^3$ л.	68.75 л.	803 дн.
10^{154}	512	$6.69 \cdot 10^{19}$	$2.7 \cdot 10^4$ л.	837 л.	26.77 л.
10^{200}	664	$1.20 \cdot 10^{23}$	$4.8 \cdot 10^7$ л.	$1.5 \cdot 10^6$ л.	$4.8 \cdot 10^4$ л.

Граница эффективности определяется числом операций, выполненных за год работы. Она сдвигается от $n = 10^{100} \approx 2^{332}$ (в 1984 г.) до $n = 10^{120} \approx 2^{399}$ (в 1994 г.) и ожидается, что она достигнет величины $n = 10^{140} \approx 2^{405}$ в 2004 г., (оставаясь все еще ниже 2^{512}).

Сенсация произошла в 1994 г., когда десятичное 129-значное число (429 бит) было разложено на два простых сомножителя каждое с 65 десятичными знаками. В соответствии с указанной выше экстраполяцией, отдельному суперкомпьютеру понадобилось бы для этого 3330 дней. В действительности же вся работа была распределена между 1600 (менее сильными) компьютерами, связанными по интернету, и была закончена за 8 месяцев. В 1999 г. было факторизовано число из 465 бит. Начиная где-то с 2004 г., числа с 512 бинарными разрядами уже не будут доставлять непреодолимых препятствий при их факторизации.

Все более и более специализированные компьютеры, с высокой степенью параллелизации, вступают в жизнь. Однако, какие бы усилия ни делались

[6] Основан на ранней работе Крайтчика, улучшенной Померанцем (1985 г.), Монтгомери (1987 г.) и Сильверменом (1987 г.). Быстрейшая версия этого алгоритма называется «ppmpqs», Double Large Prime Variation of the Multiple Polynomial Quadratic Sieve (двойные большие простые числа: изменение кратного полиномиального квадратичного решета, *англ.*). Однако ни этот метод, ни даже еще лучший метод Полларда (1988 г.) «Number Field Sieve», требующий числа шагов порядка $e^{((\ln n)^{1/3} \cdot (\ln(\ln n))^{2/3})} = n^{(\ln(\ln n)/\ln n)^{2/3}}$, не являются эффективными в смысле разд. 10.2.1

в этом направлении, продолжают сохраняться границы, обусловленные ограниченностью памяти и времени, которые нельзя преодолеть по физическим причинам. Например, в соответствии с нашими современными знаниями, для построения 10^{60}-битного блока памяти потребуется масса всей нашей солнечной системы, точно так же для выполнения 10^{70} операций потребуется больше времени, чем прошло от рождения Вселенной при «большом взрыве» примерно 10^{18} секунд тому назад — даже если каждая операция продлится не более 10^{-43} с., т. е. будет выполняться за планковское время — кратчайший временной интервал, который допускает современная физика.

Правда, эти большие числа могут ввести в заблуждение. Не существует доказательства отсутствия алгоритма более быстрого, чем квадратичное решето или сравнимых с ним. Не исключено, что разложение числа n на простые сомножители может быть осуществлено за время, растущее в зависимости от логарифма n не выше, чем полиномиально. Но это предположение неправдоподобно, так как оно противоречит двухтысячелетней истории изучения факторизации целых чисел. С другой стороны, несуществование секретов, отличных от уже известных, трудно доказать. И теория сложности в ее нынешнем состоянии мало помогает, так как она дает лишь верхние оценки для требуемых усилий. Саломаа (1990 г.) отметил, что «не существует доказуемых нижних оценок количества работы криптоаналитика, анализирующего криптосистему с открытым ключом». Какой-либо вновь найденный секрет может поставить под угрозу надежность всей криптосистемы точно так же, как прямая атака дешифрования, обходящаяся без обращения функции. В этом заключается основной риск использования всех асимметричных методов и их большой недостаток, делающий весьма сомнительным их использование при передаче и хранении особо важной информации.

10.2.4. Известные примеры однонаправленных функций. В настоящее время нет доказательств существования строго однонаправленных функций. Это связано с отсутствием методов получения хороших нижних оценок сложности. Правда, существуют хорошие кандидаты, основанные на операциях умножения и возведения в степень над полем Галуа $\mathbb{F}(p)$ для простого числа p.

10.2.4.1. Однонаправленная функция без секрета: умножение простых чисел. Как заметил Тьюринг в 1937 г. (разд. 5.7), можно сравнительно просто перемножить два числа, состоящие из десяти тысяч десятичных знаков, в том числе и два простых числа такой величины; на настольном компьютере это занимает секунды. Но сегодня не существует (разд. 10.2.3) ни одного (широко) известного эффективного метода, пригодного для разложения 200-значного десятичного числа на его простые сомножители (кроме отдельных частных случаев).

Пусть $X = \{(x_1, x_2) \mid x_1, x_2 \text{ — простые числа}, K \leqslant x_1 \leqslant x_2\}$ для достаточно большого K. Инъективная функция $f\colon X \to \mathbb{N}$, определенная условием

$$f(x_1, x_2) = x_1 \cdot x_2,$$

является однонаправленной функцией. В настоящее время нет данных, позволяющих утверждать обратное.

10.2.4.2. Однонаправленная функция без секрета: возведение в степень в поле Галуа $\mathbb{F}(p)$. Пусть p — простое число. Для числа $a \neq 0$ из поля Галуа $\mathbb{F}(p)$ a-экспоненциальная функция (т. е. степень с основанием a) F_a: $\mathbb{Z}_{p-1} \to \mathbb{Z}_p \setminus \{0\}$, определяемая условием

$$F_a(n) = a^n \pmod{p},$$

является для достаточно больших p и a однонаправленной функцией (о кольце $\mathbb{Z}_p$ см. гл. 5).

Пример: $p = 7$, $\mathbb{Z}_p \setminus \{0\} = \{1, 2, 3, 4, 5, 6\}$, $a = 2$:

n	0	1	2	3	4	5	$6 \equiv 0$
2^n	1	2	4	8	16	32	64
$2^n \bmod 7$	1	2	4	1	2	4	1

Трудоемкость нахождения F_a находится в терпимых пределах даже для значений p и a, превышающих 10^{200}. Основная идея — повторное возведение в квадрат и умножение — та же, что и в разд. 9.5.2, ее можно продемонстрировать следующим примером, где всякий раз возведение в квадрат и умножение берутся *по модулю* p:

$$a^{25} = \left(\left((a^2 \cdot a)^2\right)^2\right)^2 \cdot a, \quad \text{так как} \quad 25_{10} = 11001_2.$$

Пример с $a = 2$, $p = 7$ показывает, что a-экспоненциальная функция не обязательно инъективна. Если же функция F_a инъективна над $\mathbb{Z}_p \setminus \{0\}$ (и таким образом является групповым изоморфизмом групп $\mathbb{Z}_{p-1}$ и $\mathbb{Z}_p \setminus \{0\}$), то число a называется первообразным корнем в кольце $\mathbb{Z}_p$. Например, значения a, равные 3, 11, 12 и 22 для $p = 31$ являются первообразными корнями кольца $\mathbb{Z}_{31}$. Функция $F_3 \equiv 3^n \pmod{31}$ определяет перестановку с цикловым разложением $(13 + 9 + 8)$:

(1 3 27 23 11 13 24 2 9 29 21 15 30)(6 16 28 7 17 22 14 10 25)(4 19 12 8 20 5 26 18);

$F_{11} \equiv 11^n \pmod{31}$ — перестановку с цикловым разложением $(26 + 3 + 1)$:

(1 11 24 8 19 22 18 2 28 10 5 6 4 9 23 12 16 20 25 26 7 13 21 27 15 30)(3 29 17)(14);

$F_{12} \equiv 12^n \pmod{31}$ — перестановку с цикловым разложением $(10+7+7+5+1)$:

(1 12 4 28 14 18 8 9 15 30)(12 20 5 26 10 25 6)(3 23 22 7 24 16 19)(13 17 11 21 29)(27);

$F_{22} \equiv 22^n \pmod{31}$ — перестановку с цикловым разложением $(24 + 5 + 1)$:

(1 22 10 5 6 8 28 18 16 9 27 29 24 4 20 25 14 7 21 23 3 15 30)(2 19 11 17 12)(13).

Можно показать, что для каждого простого числа p существует по крайней мере один первообразный корень в кольце $\mathbb{Z}_p$. На самом деле существует $\varphi(p)$ первообразных корней для каждого p, где φ — функция Эйлера. Для $p = 31$,

кроме указанных выше, первообразными корнями являются еще 21, 17, 13 и 24. Для простых чисел специального вида могут существовать некоторые особенности. Например, для $p = 17$, 257, 65537 и всех больших простых чисел Ферма, т. е. чисел вида $p = 2^{2^k} + 1$, числа 3 и 7 всегда являются первообразными корнями (Бейлер, Лейтбехер).

Если число a является первообразным корнем в кольце $\mathbb{Z}_p$, то поскольку функция F_a инъективна, она имеет обратную функцию F_a^{-1}, называемую (дискретной) a-логарифмической функцией или индексом в группе $\mathbb{Z}_p \setminus \{0\}$. И если возведение в степень в кольце $\mathbb{Z}_p$ — эффективная операция, то получить эффективные алгоритмы для вычисления дискретного логарифма весьма затруднительно.

Среди известных алгоритмов дискретного логарифмирования в такой мультипликативной группе, как $\mathbb{Z}_p \setminus \{0\}$, даже хорошие алгоритмы, подобные алгоритму Шэнкса (1971 г.)[7] имеют трудоемкость, пропорциональную

$$\sqrt{|\mathbb{Z}_p|} = \sqrt{p} = e^{1/2 \ln p},$$

и, таким образом, неэффективны.

Лучший метод подсчета индекса требует нахождения подходящего базиса мультипликативной группы кольца $\mathbb{Z}_p$, обычно первых t простых чисел, и таким образом требует предварительного построения огромной базы данных. Поэтому он работает лишь в специальных случаях, имея при этом субэкспоненциальную сложность. Точнее, этот метод требует числа операций того же порядка $e^{\sqrt{\ln p \cdot \ln(\ln p)}}$, что и простое разложение с помощью квадратичного решета.

Кольцо $(\mathbb{Z}_p, +, \times)$ является полем, а именно простым полем Галуа $\mathbb{F}(p)$ характеристики p. Кроме того, рассмотрим общее поле Галуа $\mathbb{F}(p^k)$ характеристики p, являющееся расширением поля $\mathbb{F}(p)$, а в нем мультипликативную группу $\mathbb{F}(p^k) \setminus \{0\}$, которая, как известно, является циклической. Эта группа порождается некоторым элементом x, который является нетривиальным корнем уравнения $x^{p^k} - x = 0$. Элементами поля Галуа $\mathbb{F}(p^k)$ являются p^k многочленов степени не выше $k-1$ над полем $\mathbb{F}(p)$; поле $\mathbb{F}(p^k)$ можно рассматривать также как k-мерное векторное пространство над полем $\mathbb{F}(p)$.

Пример. $p = 2$, $k = 3$, $\mathbb{F}(2^3) = \{0, 1, x, x+1, x^2, x^2+1, x^2+x, x^2+x+1\}$. Многочлен $x^8 - x$ имеет неприводимый множитель третьей степени $x^3 + x + 1$, соответствующий нулевому элементу поля $\mathbb{F}(8)$, поэтому можно приводить степени с помощью соответствия $x^3 \mapsto x + 1$.

Для нас важна лишь мультипликативная группа $\mathbb{F}(p^k) \setminus \{0\}$. Заметим, что при $k > 1$ умножение в ней отличается от модулярной арифметики. Так мы окончательно получаем групповой изоморфизм

$$F_a \colon \mathbb{Z}_{p^k-1} \to \mathbb{F}(p^k) \setminus \{0\},$$

[7] Программу см. в работе: Otto Forster, *Algorithmische Zahlentheorie*, Vieweg, Braunschweig, 1966.

определяемый условием

$$F_a(n) = a^n \quad \text{в} \quad \mathbb{F}(p^k).$$

В примере $\mathbb{F}(2^3)$ с $a = x$ и с $a = x + 1$ имеем:

n	0	1	2	3	4	5	6	$7 \equiv 0$
x^n	1	x	x^2	x^3	x^4	x^5	x^6	x^7
x^n приведенное	1	x	x^2	$x+1$	x^2+x	x^2+x+1	x^2+1	1
n	0	1	2	3	4	5	6	$7 \equiv 0$
$(x+1)^n$	1	$x+1$	$(x+1)^2$	$(x+1)^3$	$(x+1)^4$	$(x+1)^5$	$(x+1)^6$	$(x+1)^7$
$(x+1)^n$ приведенное	1	$x+1$	x^2+1	x^2	x^2+x+1	x	x^2+x	1

Случай $p = 2$ при большом k имеет практический интерес. В 1988 г. Поллард предложил вариант *Number Field Sieve* (NFS) — метода расчета индексов, в котором простые числа заменяются неприводимыми многочленами — его сложность определяется числом

$$e^{((\ln p)^{1/3} \cdot (\ln \ln p)^{2/3}}.$$

Параллельные вычисления использовались в подготовительной работе для таких амбициозных задач, как вычисление индексов в $\mathbb{F}(2^{503})$ (Копперсмит (1986 г.), Маккэли (1990 г.), Гордон и Маккэли (1993 г.)).

Следующий естественный шаг в применении групповых изоморфизмов был сделан с помощью метода эллиптических кривых (ECM), развитого Коблицем (1985 г.), Миллером и Ленстрой (младшим) (1986 г.). Этот метод использует известную теорию алгебраических кривых третьего порядка (эллиптических кривых) в проективной плоскости над конечным полем K, в частности, над $\mathbb{F}(p^k)$. При этом некоторые алгоритмы, (например, алгоритм Шэнкса) могут быть перенесены на группу точек эллиптической кривой. Однако представляется, что для метода подсчета индексов нахождение подходящего базиса в случае эллиптических кривых является делом безнадежным. Таким образом, метод эллиптических кривых, возможно, предлагает более высокую надежность для идентификации и аутентификации. Особый интерес представляют эллиптические кривые над полем Галуа $\mathbb{F}(2^k)$, так как арифметические операции в этом поле легко выполняются для больших n.

При составном числе q функция $F_a(n) = a^n$ имеет один секрет, если q является произведением двух различных простых чисел: разложение q на простые сомножители позволяет создать таблицу, которая вместе с китайской теоремой об остатках (разд. 10.4.3) делает вычисление дискретного логарифма более легким, чем в общем случае.

10.2.4.3. Однонаправленная функция с секретом: возведение в степень по модулю q. В разд. 9.5.2 возведение в фиксированную степень рассматривалось в поле Галуа $\mathbb{F}(p)$. Теперь модуль q может быть составным; рассмотрим случай, когда

$$P_h(x) = x^h \pmod{q}.$$

Как и раньше, существуют подходящие пары (h, h') фиксированных чисел из $\mathbb{F}_q \setminus \{0\}$ таких, что

1) существует эффективный метод вычисления $P_h(x)$ для всех $x \in \mathbb{F}_q$;

2) существует эффективный метод вычисления $P_{h'}(x)$ для всех $x \in \mathbb{F}_q$, и при этом выполняются равенства

$$P_{h'}(P_h(x)) = x \quad \text{и} \quad P_h(P_{h'}(x)) = x.$$

Но если известны лишь h и q и при этом q довольно велико, скажем, $q > 10^{200}$, то не существует эффективного (публично) известного метода для вычисления h'.

Здесь имеется секрет. Нахождение h' становится намного легче, если число q составное и известна факторизация q на два сомножителя (оба довольно больших). Подробнее это будет рассмотрено в разд. 10.3.

10.2.4.4. Однонаправленная функция с секретом: возведение в квадрат по модулю $q = p' \cdot p''$. Эта функция является важным частным случаем из разд. 10.2.4.3 с $h = 2$, который использует так называемые «квадратичные вычеты», т. е. квадратные корни *по модулю* q; для этих корней существует теория, восходящая к Лежандру и Гауссу. Одно ее приложение для систем шифрования с открытым ключом было изучено в 1985 г. Вильямсом.

Рассмотрим сначала случай $q = p$, где p — простое число. Таблица 1 из разд. 5.5 показывает, что для $2 < p < 37$ и $N = p - 1$ отсутствуют взаимно обратные пары по модулю N чисел $\{h, h'\}$, содержащие $h = 2$. Функция $P_2(x) = x^2 (\mathrm{mod} p)$ не является ни инъективной, ни сюръективной. Для двузначной обратной функции мы будем использовать обозначение $\sqrt{\ }$. Для $p = 17$, обращая функцию $P_2(x)$, получим:

$$\sqrt{1} = \pm 1, \quad \sqrt{2} = \pm 6, \quad \sqrt{4} = \pm 2, \quad \sqrt{8} = \pm 5,$$
$$\sqrt{9} = \pm 3, \quad \sqrt{13} = \pm 8, \quad \sqrt{15} = \pm 7, \quad \sqrt{16} = \pm 4.$$

Для простого числа p существуют эффективные методы вычисления квадратного корня *по модулю* p, основанные на золотой теореме Гаусса, — квадратичном законе взаимности.

Иная ситуация имеет место для составного q, скажем, для $q = p' \cdot p''$. Если разложение q на простые сомножители известно, то для каждого простого делителя p числа q квадратные корни $\sqrt{a} = \pm u$ *по модулю* p можно вычислить эффективно, и тогда $\sqrt{a}$ *по модулю* q легко может быть вычислен. Но для того, кто не знает разложения q на простые сомножители, вычисление $\sqrt{a}$ *по модулю* q, как показано Рабином (1979 г.), столь же трудно, как и факторизация числа q.

10.3. Метод RSA

Метод RSA является самым известным среди методов шифрования с открытым ключом[8]. Он назван в честь Ривеста, Шамира и Адлемана (1978 г.), чей американский патент действовал до 20 сентября 2000 г. Метод RSA основан на широко признанном предположении, что при определенных условиях возведение в фиксированную степень *по модулю* q является однонаправленной функцией с секретом (разд. 10.2.4.3).

10.3.1. Пусть для i-го партнера в асимметричной шифровальной сети

(1) $q_i = p_i' \cdot p_i''$, где p_i' и p_i'' — нечетные простые числа, $p_i' \neq p_i''$.

(2) $e_i, d_i \in \{1, 2, \ldots, q_i\} \subset \mathbb{Z}_{q_i} \setminus \{0\}$, где

(2a) $\text{НОД}(e_i, \psi(q_i)) = 1$,

(2b) $\text{НОД}(d_i, \psi(q_i)) = 1$,

(2c) $e_i \cdot d_i \pmod{\psi(q_i)} = 1$,

здесь ψ обозначает функцию Кармайкла[9][10]

$$\psi(p_i' \cdot p_i'') = \text{НОК}\ (p_i' - 1, p_i'' - 1) = 2 \cdot \text{НОК}\left(\frac{p_i' - 1}{2}, \frac{p_i'' - 1}{2}\right).$$

Метод RSA является многосимвольным, одноалфавитным блочным шифрованием с символами $p_j \in \mathbb{Z}_{q_i}$ открытого текста и символами $c_j \in \mathbb{Z}_{q_i}$ криптотекста, $q = q_i$.

Для i-го партнера используются следующие ключи:

публичные e_i (для шифрования)[11], q_i,
частные d_i для расшифрования.

Схема шифрования определяется с помощью однонаправленной функции $E_i\colon \mathbb{Z}_{q_i} \to \mathbb{Z}_{q_i}$,

$$E_i(m_j) = m_j^{e_i} \bmod q_i = c_j.$$

Схема расшифрования определяется с помощью однонаправленной функции $D_i\colon \mathbb{Z}_{q_i} \to \mathbb{Z}_{q_i}$,

$$D_i(c_j) = c_j^{d_i} \bmod q_i = m_j.$$

---

[8] Американский патент № 4 405 829, 20 сентября 1983.

[9] В исходной публикации вместо функции Кармайкла используется функция Эйлера φ, $\varphi(p_i' \cdot p_i'') = (p_i' - 1) \cdot (p_i'' - 1)$. Поставленные условия гарантируют выполнение оригинального метода.

[10] $\psi(2 \cdot p_i' \cdot p_i'') = \text{НОК}\,(1, p_i' - 1, p_i'' - 1) = \text{НОК}\,(p_i' - 1, p_i'' - 1)$.

Для общего определения функции Кармайкла $\psi(n)$ см. Arnold Scholz и Bruno Schoentberg, Einführung in die Zahlentheorie, 5th ed., de Gruyter, Berlin 1973. Функция $\psi(n)$ является делителем $\varphi(n)$. Теорема Кармайкла является «очень полезным, но забытым обобщением теоремы Эйлера» (H. Riesel, Prime Numbers and Computer Methods for Factorization. Birkhäuser, Basel, 1985.) В действительности Ривест, Шамир и Адлеман не используют эту теорему и Саломаа не указывает ее в своей книге 1990 г. Шольц и Шенеберг не ссылаются на Кармайкла (1879–1967 гг.).

[11] Поскольку $\psi(p_i' \cdot p_i'')$ четно, то $e = 2$ исключается.

Эти схемы определяют цифровую подпись, так как

$$D_i(E_i(x)) = E_i(D_i(x)) = x \quad \text{для всех} \quad x \in \mathbb{Z}_{q_i}.$$

Доказательство вытекает из доказательства аналогичного равенства в разд. 9.5.2. Это можно обосновать следующим следствием из теоремы Кармайкла для взаимно простых чисел a и n:

$$\text{если } b \equiv b' \pmod{\psi(n)}, \text{ то } a^b \equiv a^{b'} \pmod{n}.$$

10.3.2. Пример (Саломаа).

$$(e_i,\, d_i) = (1031,\, 31\,963\,885\,304\,131\,991),$$
$$q_i = 32\,954\,765\,761\,773\,295\,963 = 3\,336\,670\,033 \cdot 9\,876\,543\,211;$$
$$\psi(q_i) = 5\,492\,460\,958\,093\,347\,120 = 3\,336\,670\,033 \cdot 9\,876\,543\,210/6.$$

Даже этот пример с большими числами является нереалистичным с точки зрения практической надежности. В качестве примера мы используем малые числа, более подходящие для ручных вычислений.

В проектировании криптосистемы RSA используется секретная информация: мы начинаем с двух простых чисел

$$p_i' = 47, \quad p_i'' = 59.$$

Это дает

$$q_i = p_i' \cdot p_i'' = 47 \cdot 59 = 2773,$$
$$\psi(q_i) = \psi(2773) = \text{НОК}\,(46,\, 58) = 2 \cdot 23 \cdot 29 = 1334.$$

Теперь должно быть найдено такое число e_i, чтобы $\text{НОК}\,(e_i,\, 1334) = 1$. Существует много таких чисел, например,

$$e_i = 3,\, 5,\, 7,\, \ldots,\, 19,\, 21,\, 25,\, 27,\, 33,\, 35,\, 37,\, 39,\, \ldots$$

Если выбрано некоторое множество таких чисел $\{e_i\}$, то возможно даже многоалфавитное шифрование.

Возьмем $e_i = 17$, d_i получается из условия $e_i \cdot d_i = 1 \pmod{1334}$ с помощью алгоритма быстрого деления, который дает $d_i = 157$ (d_i не должно становиться слишком маленьким, так как в этом случае его можно найти методом проб и ошибок (разд. 10.4.3). Поэтому может быть предпочтительнее сначала выбрать подходящее d_i и по нему найти e_i.

Таким образом, схема шифрования задается равенством

$$E_i(m_j) = m_j^{17} \pmod{2773}.$$

Ввиду того, что $17_{10} = 10001_2$, возведение в степень 17 можно эффективно выполнить следующим образом:

$$\left(\left(\left(\left(m^2 \pmod{2773}\right)^2 \pmod{2773}\right)^2 \pmod{2773}\right)^2 \pmod{2773}\right) \cdot m \pmod{2773}.$$

Схема расшифрования задается равенством

$$D_i(c_j) = c_j^{157} (\mathrm{mod}\ 2773).$$

Кодирование буквенных символов: ⊔ (пробел)[12], a, b, ..., z биграммами 00, 01, 02, ..., 26, позволяет биграммы открытого текста закодировать числами из $\mathbb{Z}_{2773}$, так как $2626 < 2773$.

Например, сообщение

e r r a r e ⊔ h u m a n u m ⊔ e s t

сначала кодируется блоками длины два:

05 18 18 01 18 05 00 08 21 13 01 14 21 13 00 05 19 20,

а затем шифруется:

1787 2003 2423 0596 0340 1684 0340 0508 2109.

Идентичные блоки открытого текста приводят к идентичным блокам криптотекста — шифрование является поблочно одноалфавитным. Этот режим ECB (по аналогии с DES, разд. 9.6.3) должен быть защищен, по крайней мере, автоключом, как в режиме CBC. Даже периодическое, по-настоящему многоалфавитное шифрование, было бы значительно лучше.

10.4. Криптоаналитическая атака на RSA

Ничто не должно помешать криптоаналитику испробовать все классические методы атаки (см. часть II) против шифра RSA. Тем более, что этот шифр имеет несколько специфических слабостей.

10.4.1. Атака факторизацией числа q_i. Криптоаналитик, нашедший факторизацию числа $q_i : q_i = p_i' \cdot p_i''$, может вычислить

$$\psi(q_i) = 2 \cdot \mathrm{НОК}\left(\frac{p_i' - 1}{2}, \frac{p_i'' - 1}{2}\right)$$

и, зная e_i, вычислить d_i.

Для защиты метода RSA от такой атаки, т. е. чтобы сделать факторизацию числа q_i неподатливой (в действительности факторизация числа является более сложной функцией, чем проверка его простоты), должны быть выполнены следующие условия:

1) $q_i = p_i' \cdot p_i'' > 10^{200}$,

2) десятичные записи чисел p_i' и p_i'' имеют близкие длины,

3) ни p_i', ни p_i'' не должны быть малыми, не должны выбираться из какой-нибудь таблицы простых чисел и не должны иметь какой-нибудь специальный вид.

[12] Вопреки классическому обычаю, Ривест, Шамир и Адлеман не подавляют словарные пробелы, и литература по методу RSA в этом следует им.

Условие (1) препятствует (разд. 10.2.1) атаке перебором. Условие (2) делает невозможным представление q_i в виде разности двух квадратов (техника, восходящая к Ферма, 1643 г.):

$$q_i = p_i' \cdot p_i'' = \left(\frac{p_i' + p_i''}{2}\right)^2 - \left(\frac{p_i' - p_i''}{2}\right)^2$$

со значениями $(p_i' + p_i'')/2$, возрастающими от $\sqrt{q_i}$.

Условие (3) препятствует выбору в качестве возможных множителей достаточно малых простых чисел. Ни одна из этих атак, судя по имевшимся до сих пор сообщениям, успеха не имела. Вероятно, потому, что довольно легко можно предпринять защитные меры.

10.4.2. Атака итерацией. (Разд. 9.4.2.) Пусть $c^{(0)} = m_j$ — блок открытого текста, а $c^{(1)} = c_j = E_i(m_j)$ — блок криптотекста.

Образуем последовательность

$$c^{(k+1)} = E_i(c^k), \qquad k = 0, 1, 2, \ldots .$$

Наименьшее $k \geqslant 1$, для которого $c^{(k+1)} = c^{(1)}$, назовем показателем итерации s_m для $m = m_j$; показатель итерации указывает длину цикла, которому принадлежит m_j. Число $s_m - 1$ называется показателем возврата для $m = m_j$.

Пример 1. Как выше в разд. 10.3.2, положим $m_j = 0518$. Тогда

$$(e_i, d_i) = (17, 157), \quad q_i = 2773 = 47 \cdot 59; \ \psi(q_i) = 1334 = 2 \cdot 23 \cdot 29,$$

$$\begin{aligned}
c^{(0)} &= m_j = && 0518 && (= 11 \cdot 47 + 1)\\
c^{(1)} &= c_j = 0518^{17} \bmod 2773 &= 1787 && (= 38 \cdot 47 + 1)\\
c^{(2)} &= 1787^{17} \bmod 2773 &= 0894 && (= 19 \cdot 47 + 1)\\
c^{(3)} &= 0894^{17} \bmod 2773 &= 1364 && (= 29 \cdot 47 + 1)\\
c^{(4)} &= 1364^{17} \bmod 2773 &= 0518 &= m_j
\end{aligned}$$

Здесь мы пришли к открытому тексту с показателем возврата $s_{m_j} = 3$. Однако незаконный дешифровальщик не может этого знать; он находит следующий член последовательности:

$$c^{(5)} = 0518^{17} (\mathrm{mod} 2773) = 1787 = c_j.$$

Из следствия теоремы Кармайкла (см. разд. 10.3.1) имеем:

$$c^{(k)} = m_j^{17^k} (\mathrm{mod}\ 2773) = m_j^{17^k (\mathrm{mod}\ 1334)} (\mathrm{mod}\ 2773).$$

Но $17^{44} (\mathrm{mod} 1334) = 1$. Поэтому $c^{(44)} = m_j^1 \ (\mathrm{mod}\ 2773) = m_j$, и 44 является верхней границей для самого длинного периода, какой только может встретиться с $e_i = 17$.

Заметим, что 44 является делителем для

$$\psi(\psi(47 \cdot 59)) = \psi(2 \cdot 23 \cdot 29) = 2 \cdot 11 \cdot 14 = 308.$$

В действительности все 2773 элементов множества $\mathbb{Z}_{2773}$ разбиваются на

9 циклов длины 1 (неподвижные точки),
42 цикла длины 4 (включая цикл, начинающийся с 0518),
6 циклов длины 22,
56 циклов длины 44.

Пример 2.

$$(e_i,\, d_i) = (7,\, 23), \quad q_i = 55 = 5 \cdot 11; \quad \psi(q_i) = 20 = 2 \cdot 2 \cdot 5.$$

Этот пример достаточно простой для того, чтобы выписать все циклы:

9 неподвижных точек:

$$(0),\ (1),\ (10),\ (11),\ (21),\ (34),\ (44),\ (45),\ (54),$$

3 цикла длины 2:

$$(12,\, 23), \quad (22,\, 33), \quad (32,\, 43),$$

10 циклов длины 4:

$$(2,\, 18,\, 17,\, 8),\ (3,\, 42,\, 48,\, 27),\ (4,\, 49,\, 14,\, 9),\ (5,\, 25,\, 20,\, 15),\ (6,\, 41,\, 46,\, 51),$$
$$(7,\, 28,\, 32,\, 13),\ (16,\, 36,\, 31,\, 26),\ (19,\, 24,\, 29,\, 39),\ (30,\, 35,\, 40,\, 50),\ (37,\, 38,\, 47,\, 53).$$

Заметим, что $7^4 \pmod{20} = 1$. Значит, 4 является верхней границей длин периодов, которые могут встретиться с $e_i = 7$. Отметим также, что 4 совпадает с $\psi(\psi(55)) = \psi(20) = 4$.

Пример 3.

$$(e_i,\, d_i) = (3,\, 675), \quad q_i = 1081 = 23 \cdot 47, \quad \psi(q_i) = 506 = 2 \cdot 11 \cdot 23.$$

Теперь $3^{55} (\mathrm{mod}\, 506) = 1$. Значит, 55 является верхней границей длины периода, который может встретиться с $e_i = 3$. Заметим, что 55 равно

$$\frac{1}{2} \cdot \psi(\psi(1081)) = \frac{1}{2} \cdot \psi(506) = \frac{1}{2} \cdot 110 = 55.$$

Таким циклом длины 55 является цикл

$$\begin{array}{rrrrrrrrrrrrrr}
(512 & 768 & 430 & 531 & 629 & 98 & 722 & 683 & 209 & 284 & 995 & 653 & 16 & 853 \\
813 & 535 & 239 & 1051 & 25 & 491 & 190 & 55 & 982 & 439 & 54 & 719 & 676 & 568 \\
393 & 307 & 397 & 331 & 384 & 324 & 721 & 1041 & 860 & 1005 & 991 & 675 & 213 & 538 \\
660 & 807 & 606 & 627 & 101 & 108 & 347 & 192 & 581 & 354 & 867 & 2 & 8). &
\end{array}$$

Всего имеется 16 циклов длины 55, 12 циклов длины 5 и снова 9 циклов длины 1 (неподвижных точек), приводимых ниже:

$$(0),\ (10),\ (46),\ (47),\ (93),\ (1034),\ (1035),\ (1080).$$

Метод, защищающий RSA против атаки итерацией, подразумевает достижение большого показателя возврата для подавляющего большинства[13] элементов $m_j \in \mathbb{Z}_q$ $(q = q_i)$. Чтобы добиться этого, $\psi(\psi(q_i))$ должно быть настолько большим, насколько это возможно.

На самом деле справедлива следующая

Основная теорема об итерационной атаке. Для всех e_i, взаимно простых с $\psi(q_i)$, показатель итерации является делителем $\psi(\psi(q_i))$; эта граница величины показателя итерации достижима при подходящем e_i.

Доказательство основано на следствии теоремы Кармайкла:

$$c^{(k)} = m_j^{e_i^k} \pmod{q_i} = m_j^{e_i^k \pmod{\psi(q_i)}} \pmod{q_i} = \\ = m_j^{e_i^{k \pmod{\psi(\psi(q_i))}} \pmod{\psi(q_i)}} \pmod{q_i}.$$

При $k = \psi(\psi(q_i))$ имеем

$$c^{(\psi(\psi(q_i)))} = m_j^{e_i^{\psi(\psi(q_i))} \pmod{q_i}} = m_j^{e_i^0 \pmod{\psi(q_i)}} \pmod{q_i} = m_j^1 \pmod{q_i} = c^{(0)}.$$

Таким образом, величина $\psi(\psi(q_i))$ является периодом и потому кратна показателю итерации.

Для того, чтобы по крайней мере число

$$\psi(q_i) = \psi(p_i' \cdot p_i'') = 2 \cdot \text{НОК}\left(\frac{p_i' - 1}{2}, \frac{p_i'' - 1}{2}\right)$$

не стало малым, на p_i' и p_i'' должны быть наложены следующие условия:

4) Оба числа $(p_i' - 1)/2$ и $(p_i'' - 1)/2$ должны содержать большие простые множители.

5) НОД $\left((p_i' - 1)/2, (p_i'' - 1)/2\right)$ должен быть малым.

Условия (4) и (5) выполняются, если p_i' и p_i'' являются «сильными» простыми числами (разд. 9.5.2). Тогда числа $(p_i' - 1)/2$ и $(p_i'' - 1)/2$ просты и $\psi(q_i) = 2 \cdot (p_i' - 1)/2 \cdot (p_i'' - 1)/2 \approx q_i/2$.

Усилия по нахождению сильных простых чисел могут дать ценные результаты, но пока неизвестно даже, бесконечно ли множество таких чисел.

Кроме того, чтобы число $\psi(\psi(q_i))$ не стало малым, ввиду равенства

$$\psi\left(2 \cdot \frac{p_i' - 1}{2} \cdot \frac{p_i'' - 1}{2}\right) = 2 \cdot \text{НОК}\left(\left(\frac{p_i' - 1}{2} - 1\right)\Big/2, \left(\frac{p_i'' - 1}{2} - 1\right)\Big/2\right) = \\ = 2 \cdot \text{НОК}\left(\frac{p_i' - 3}{4}, \frac{p_i'' - 3}{4}\right)$$

на числа p_i' и p_i'' должны быть наложены следующие условия:

6) Оба числа $(p_i' - 3)/4$ и $(p_i'' - 3)/4$ должны содержать большие простые множители.

[13] Большего ожидать нельзя, так как всегда существуют даже неподвижные точки итерации; фактически Саломаа показал, что всегда существует 9 неподвижных точек.

7) НОД $((p_i'-3)/4, (p_i''-3)/4)$ должен быть малым.

Условия (6) и (7) выполняются, если вдобавок и числа $(p_i'-1)/2$ и $(p_i''-1)/2$ являются сильными простыми, т. е. числа p_i' и p_i'' являются дважды сильными простыми; тогда и числа $(p_i'-3)/4$ и $(p_i''-3)/4$ являются простыми числами и $\psi(\psi(q_i)) = 2\cdot(p_i'-3)/4\cdot(p_i''-3)/4$.

Дважды сильными простыми являются числа: 11, 23, 47, 167, 359, 719, 1439, 2039, 2879, 4079, 4127, 4919, 5639, 5807, 5927, 6047, 7247, 7559, 7607, 7727, 9839, 10799, 11279, 13799, 14159, 15287, 15647, 20327, 21599, 21767, ..., а также 2 684 999, 5 369 999 и 10 739 999.

Кроме числа 11, все дважды сильные простые числа имеют вид $24n-1$. Для дважды сильных простых чисел p_i' и p_i'' (для которых $p_i'\cdot p_i'' = q_i$), $\psi(\psi(q_i)) \approx q_i/8$.

10.4.3. Атака в случае малого e_i. Трудоемкость шифрования RSA мала, если мало число e_i (предельный случай $e_i = 3$). Это может оказаться благоприятным обстоятельством, если отправитель имеет ограниченные вычислительные возможности (например, в случае smart карты, а получатель, напротив, не страдает и от довольно большого d_i, например, при наличии мощного компьютера).

Следует избегать использования малого d_i, так как в этом случае незаконный дешифровальщик может воспользоваться атакой перебора. Использование малого e_i тоже опасно в случае, когда один и тот же блок открытого сообщения m_j, при помощи одной и той же степени $e_1 = e_2 = \ldots = e_s = e$, посылается разным получателям с (предположительно попарно взаимно простыми) q_1, q_2, ..., q_s, образуя криптотексты $m_j^e \pmod{q_1}$, $m_j^e \pmod{q_2}$, ..., $m_j^e \pmod{q_s}$.

Из этих перехваченных криптотекстов с помощью китайской теоремы об остатках может быть вычислено значение $m' = m_j^e \pmod{q_1, q_2, \ldots, q_s}$. Но поскольку m' меньше каждого из индивидуальных модулей, то выполняется равенство $m_j^e = m'$. Мы получили уравнение с известным m' и известным малым e (хоть и включающее довольно большие числа); оно может быть разрешено относительно m_j[14]. Правда, взлом не является полным: число d_i все еще остается неизвестным.

10.4.4. Риски. Существуют не только определенные блоки открытого текста, которых следует избегать, поскольку они приводят к очень коротким показателям возврата. Пример 2 показывает, что существуют также такие величины e_i, которые приводят к циклам малой длины. Некоторых чисел e_i следует избегать всегда: например, $e_i = \psi(q_i) + 1$ означает, что и $d_i = \psi(q_i)$, так что $E_i(m_j) = D_i(m_j)$, т. е. тождественны, так что все m_i становятся неподвижными точками.

Существуют другие удивительные находки: если для данного произведения $q_i = p_i'\cdot p_i''$ двух простых чисел p_i' и p_i'' можно вычислить $\psi(q_i)$, то q_i может быть разложено на множители. Действительно, если $(p_i'-1)/2$ и $(p_i''-1)/2$

[14] Винер (M. J. Wiener), *Cryptanalysis of short RSA secret exponents*. EUROCRYPT'89 Proceedings. Lecture Notes in Computer Science 434, Springer, 1990.
Также: IEEE Transactions on Information Theory, v. 36, № 3, May 1990, pp. 553–558.

взаимно просты, то уравнения $q_i - 2\psi(q_i) + 1 = p_i' + p_i''$ и $q_i = p_i' \cdot p_i''$ определяют оба множителя.

10.4.5. Недостатки. Метод RSA считается практически надежным при соблюдении указанных выше условий. По крайней мере, ни о каких серьезных успешных атаках на него ничего не известно.

Но метод RSA имеет следующие недостатки:

1) RSA нуждается в довольно длинных ключах q_i, — в ближайшем будущем из 1024 или большего числа бит.

2) RSA медленнее, чем DES более чем в тысячу раз.

10.5. Секретность или аутентификация?

Перед системой шифрования с открытым ключом, ввиду общедоступности множества ключей, возникает проблема, которую классические симметричные методы долгое время игнорировали. Пользователей этих методов интересовали лишь пассивные действия противника, который читал или подслушивал шифрсообщения, передаваемые по слабо защищенным каналам связи, вроде телеграфных, радио, оптических или акустических. Задача шифрования состояла в том, чтобы сделать криптоанализ перехваченных сообщений как можно более трудным. Возможность активного влияния на канал связи не рассматривалась в качестве серьезной опасности угрозы вражеского проникновения в него считалась просто невозможной. Это было беспечностью.

Однако радиоконтакты со шпионами выявили человеческую сторону этой проблемы: шпион может быть взят в плен, а с его передатчиком может оперировать кто-то другой. Такая *Funkschpiel* (радиоигра, *нем.*) встречалась в 1942 г. и 1943 г. между немцами и англичанами (включая одного датского подпольного агента). Для исключения подобных ситуаций всегда нужно требовать подписи оператора, хотя даже это не помогает, если оператор «перевербован». Если, однако, оператор работает под контролем противника, он может пропускать элементы «секретного контроля», которые предположительно он должен регулярно вставлять в сообщение в качестве аутентификатора[15]. Все это совершенно аналогично обычному использованию подписи. В важных вопросах аутентификация так же важна, как и секретность.

Существует глубокий конфликт между секретностью и аутентификацией, как это показывает следующий пример. Некоторое сообщение, имеющее характер тревоги, может быть записано и позднее передано вновь, вызвав ложную тревогу. Этого можно избежать, вставив пометку о времени сообщения, но в этом случае возникнет компромисс открытый текст-криптотекст (разд. 11.2.5) с опасностью взлома шифра. Секретность и аутентификация — разные вещи, и из одной не следует другая.

Проиллюстрируем этот конфликт ролью избыточной информации. Зашифрованное сообщение лучше защищено против криптоанализа, если оно содержит мало избыточной информации; и лучше защищено против поддел-

[15] Подобная ситуация хорошо показана в фильме «Вариант Омега» с О. Далем в главной роли. — *Прим. ред.*

ки, если избыточной информации много. Хороший пример дают банкноты и рукописные подписи. Секретность антагонистична аутентификации, и для выполнения обоих требуется два мероприятия, независимых друг от друга. Это можно видеть в определении метода подписи (разд. 10.1.3), противопоставляемого методу обеспечивающему только секретность.

Методы шифрования, используемые для цифровой подписи, предлагают дополнительную идентифицирующую информацию в соответствии с заранее запланированным этикетом («протоколом»), а также включают предварительное исправляющее ошибки (разд. 4.4.6) кодирование.

Мощь асимметричных методов проявляется, в частности, в идентификации и потому они полезны при распределении ключей, — наиболее опасной части управления ключами. Это с самого начала подчеркивали Диффи и Хеллман. Большое количество времени, требуемое асимметричными методами (может быть, большее, чем симметричными методами) более чем оправдано, при их использовании для цифровой подписи. Длины подписей и ключей обычно малы по сравнению с длинами сообщений. Асимметричные системы (или системы шифрования с открытым ключом) и классические симметричные системы не являются антагонистическими, они дополняют друг друга. В международном банковском деле стойкость шифрования данных невелика, зато надежности аутентификации уделяется большое внимание.

В США Digital Signature Standard, DSS (стандарт цифровой подписи, *англ.*) Национального института стандартов и технологий (N.I.S.T.) основан на Digital Signature Algorithm, DSA (алгоритме цифровой подписи, *англ.*), который, посягая на патенты Шнорра и Эль Гамаля, использует в качестве однонаправленной функции дискретный логарифм (разд. 10.2.5). Этот выбор был подвергнут критике, так как по мнению некоторых исследователей в качестве стандарта предпочтительнее использовать алгоритм RSA.

Все односторонние функции, упоминавшиеся до сих пор, существенным образом опираются на модулярную арифметику. Другая однонаправленная функция, имеющая значительный математический интерес, возникла из «проблемы рюкзака» — одной из задач целочисленного программирования.

10.6. Надежность систем публичного ключа

Шеннон, конечно же, не хотел, чтобы его положение: «Противнику известна используемая система» понималось в том смысле, что противнику надо дать полное описание используемого шифра. Название «системы шифрования с открытым ключом» лучше, чем название «системы с публичным ключом», говорит о том, что ключи расшифрования и остальная информация держится в секрете. Эта открытость (метода шифрования) имеет технические основания, а вовсе не политические, и настоятельная необходимость превращается в реальность (в наше время то же происходит и с симметричными методами, которые не нуждаются в открытости). В обществе выражение «публичный ключ» может создать впечатление, что криптоанализ сейчас больше чем когда-либо является общедоступной областью.

Конечно это не так. Криптоанализ до сих пор остается областью, покрытой тайной. Как бы то ни было, я не могу удержаться, чтобы не заметить, что используемые сейчас системы с открытым ключом должны доставлять огромную радость профессиональному криптоаналитику, который может использовать для их взлома не только специально разработанные, но и все известные классические методы атаки. Это возможно, в частности, в том случае, когда система с открытым ключом работает слишком долго, или когда при передаче больших объемов информации одни и те же ключи используются по несколько раз. Иногда разумные идеи приводят к иллюзорным сложностям; например, применение двойных сильных простых чисел может открыть широкую дорогу криптоаналитической атаке[16)].

Пользователи часто делают заключение о безопасности той или иной системы шифрования только на основе ее комбинаторной сложности. Однако большинство результатов в теории сложности — довольно трудного раздела математики — относится к верхним оценкам сложности решения задач, нижние оценки, например, для задачи факторизации целых чисел, в настоящее время не известны.

Устранение классического шифровальщика и замена его машинисткой с компьютером понижает безопасность: исключение ошибок шифрования, допускаемых шифровальщиками, в гораздо большей степени компенсируется недостатком опыта и проницательного интеллекта, которые являются единственным средством против опасных криптологических ошибок.

Таким образом, нельзя ожидать, что опытные криптоаналитики, особенно работающие в государственных спецслужбах, не смогут взломать системы шифрования с публичным ключом. Однако узнать о таком взломе будет трудно, так как профессионалы сдержанны и не хвастают результатами своей работы.

[16)] На это указал Герольд (Anton Gerold).

Глава 11

Надежность шифрования

Даже в криптологии молчание — золото.

Лоуренс Д. Смит.

Пароли служат для выбора метода из класса методов, а ключи — для выбора схемы шифрования из системы шифрования. Не будет ошибкой предположить, что противнику известно, какой именно метод выбран — их существует не так уж много, и большинство криптографов знакомо лишь с немногими. «Основной закон криптологии», согласно Керкхоффсу[1)], сформулировавшему его словами «*il faut qu'il puisse sans inconvénient tomber entre les mains de l'ennemy*», был выражен более кратко Шенноном в 1949 г. так: «враг знает используемую систему». Отсюда следует, что нужно быть особенно внимательным в выборе ключа. Серьезной ошибкой является использование очевидных слов. Порта принадлежит следующее недвусмысленное предостережение: «чем дальше удалены ключевые слова от области общего знания, тем больше надежность, которую они обеспечивают». Использование ключей вряд ли стало общей практикой прежде чем неуполномоченные лица начали добиваться дешифрования сообщения путем угадывания ключевого слова.

Порта описывает случай дешифрования им сообщения за пять минут благодаря угадыванию ключа OMNIA VINCIT AMOR (любовь побеждает все, Вергилий, *лат.*). Джованни Баттиста Аргенти тоже имел счастье угадать ключевое слово IN PRINCIPIO ERAT VERBUM (вначале уничтожь слово, *лат.*). Такие слова, как TORCHE (факел, *фр.*) или LIBERTY (свобода, *англ.*), GLOIRE (слава, *фр.*) и PATRIE (отечество, *фр.*), KAISER (император, *нем.*) и VATERLAND (отечество, *нем.*), выражающие благородные патриотические чувства, может быть очень хороши для поднятия морали, но вряд ли годятся в качестве криптографических ключей. (Удивительно, как много людей выбирают свои имена и даты рождения в качестве компьютерного пароля. Наверное, они неспособны запомнить что-либо другое.)

[1)] Керкхоффс (Auguste Kerckhoffs) (1835–1903 гг.) — фламандский профессор, написал книгу La cryptographie militaire, 1883 г.

11.1. Криптографические ошибки

Под ошибками мы понимаем нарушение безопасности, т. е. не только применение очевидного ключа, но все, что делает более легкой жизнь неуполномоченного дешифровальщика.

11.1.1. Ошибки шифрования. Прежде всего, это ошибки шифрования. Они делают работу законного дешифровальщика трудной или даже невозможной. В последнем случае катастрофа совсем рядом: он вынужден попросить повторить сообщение. Если дословно то же сообщение шифруется тем же ключом (на этот раз правильно), то оба сообщения легко сравнить, и они будут совпадать всюду, кроме точек, где произошла ошибка. Это допускает «разностный криптоанализ» полученного «компромисса открытый текст-открытый текст». Если же для того же повторного сообщения используется другой ключ (криптотекст-криптотекст компромисс), то в результате соответствующих процедур иногда можно этот ключ получить — даже если это был прогрессивный ключ, в котором алфавиты не повторялись. Может показаться невероятным, но во время Второй мировой войны немцы часто передавали один и тот же приказ разным подразделениям, использующим разные ключевые сети, применяя разные методы шифрования или ключи — идентичные длины неизбежно вызывали подозрение. Единственное разумное решение в данном случае состоит в переписывании сообщения с использованием новых слов и фраз. Даже русский метод (разд. 3.4.2) разрезания сообщения где-то в середине и соединения частей в обратном порядке не может помочь в таких случаях.

11.1.2. Другой классической технической ошибкой является повторение зашифрованного сообщения открытым текстом, например, потому что получатель еще не получил новый ключ. Дело не только в том, что кто-нибудь может прочитать сообщение: метод шифрования и ключ теперь могут быть восстановлены. Это может скомпрометировать не только ключ на этот день, но также основной метод, используемый для построения или выбора дневного ключа. По этой причине, «горе тому, кто передает открытый текст» — было железным правилом лейтенанта Ягера (разд. 4.4), который был любимцем криптоаналитических групп союзников. Очевидно, что кульминацией в жизни профессионального криптоаналитика является работа с каким-нибудь компромиссом, а также понятно, что секретные службы применяют все возможные хитрости, чтобы спровоцировать такой случай. В 1941 г. один высший японский государственный служащий ухитрился передать американскому послу Грею заметку, в которой было сказано, что один член японского правительства хочет послать сообщение американскому правительству, но боится, что военные лидеры смогут это узнать, и потому просит переслать его, закодировав самым секретным дипломатическим кодом. Это был код М-138-А, и так зашифрованный текст известного сообщения попал в эфир. Несмотря на это, японцы оказались не в состоянии взломать М-138-А.

Подобная история случилась во время дела Дрейфуса. Когда Альфред Дрейфус в 1894 г. был арестован по самым нелепым обвинениям, и газета *La*

Libre Parole радостно протрубила эту новость, генерал Паниццарди, итальянский военный атташе в Кэ д'Орсе послал телеграмму в Рим. Французские криптоаналитики, которым была передана ее копия, имели причину полагать, что Паниццарди использовал коммерческий код Баравелли (разд. 4.4.3), оперировавший с группами из одной, двух, трех и четырех цифр, и что затем этот код был подвергнут перешифрованию. Поиск последовательности /dreyfus/, которая должна быть закодирована как 227 1 98 306, привел их к шаблону 527 3 88 706, и так они узнали, что перешифровка касалась только первого знака каждой группы (это достигалось перенумерацией страниц кодовой книги). Им удалось дешифровать все сообщение, кроме четырех последних групп. Возникло подозрение, что эти 4 группы означали: «*ufficiale rimane prevenuto emissaria*» (чиновник, оставаться, предупрежден, шпион, *итал.*), что толковалось Сандхерром, шефом секретной службы, как доказательство вины Дрейфуса. Весь следующий день криптоаналитики работали над системой перенумерации страниц, и это привело к другому дешифрованию: «*ufficialmente evitare commenti stampa*» (официально, избегать, сплетни, пресса, *итал.*). Это реабилитировало Дрейфуса, но Сандхерра не удалось убедить. «Такие вещи всегда в чем-то неточны» — комментировал он. Поэтому у Маттона, одного из подчиненных Сандхерра, возникла идея всучить Паниццарди какое-нибудь сообщение. Двойной агент подбросил ему в виде утечки информации текст, который выглядел как важное сообщение, и Паниццарди передал его в Рим почти дословно. Криптоаналитики, не осведомленные о том, следствием чего оно было, почти сразу же дешифровали это сообщение. Теперь Маттон убедился, что он был прав. Тем не менее сфабрикованная версия была представлена в суд, и Дрейфуса осудили. Справедливость восторжествовала лишь спустя 12 лет (в 1906 г. Дрейфус был оправдан). Однако Франция еще не покончила со скандалом Дрейфуса: в феврале 1994 г. французский министр обороны Леотар уволил главу Исторического архива вооруженных сил полковника Гожака за публикацию «недопустимых тенденциозных исследований» по делу Дрейфуса.

Австро-Венгерская империя также добилась своего триумфа. После того, как криптологическая команда Фигля проанализировала 150 слов одного итальянского дипломатического кода, используемого между Римом и Константинополем, она шаг за шагом увеличивала свои знания с помощью контрабандных фрагментов информации, представляющих важное военное значение, почерпнутых из итальянских газет, публикуемых в Константинополе. В течение месяца они смогли расширить свой словарь до 2000 слов.

Еще одним из простейших методов дешифрования является метод, известный русским как «кража открытого текста из посольства». Итальянцы тоже имели свою *squadra penetrazione* (бригаду проникновения, *итал.*). После такого похищения дипломат обычно спешил уверить свое правительство, что код, находившийся в пользовании и теперь скомпрометированный, был не очень хорошим.

Примером кода, вышедшего из употребления, был код американского госдепартамента GRAY (означающий серый цвет, а не метод Грея бинарного

кодирования). Когда в конце Первой мировой войны пришла пора заменить устаревшие и уже скомпрометированные коды RED, BLUE и GREEN, никто не думал, что они останутся в использовании еще на два десятилетия. Чиновники министерства иностранных дел были так хорошо знакомы с этим вопросом, что смогли произнести импровизированные речи, используя элементы кода GRAY. 6 декабря 1941 г. президент Франклин Рузвельт послал меморандум Корделлу Халлу: «Дорогой Корделл, передай это Грею (американскому послу в Токио) — я думаю, можно использовать серый код — это сэкономит время, и я не возражаю, если это позволит читать наши сообщения. ФДР». Было слишком поздно, чтобы достичь желаемого эффекта; потребовалось время на дешифрование текста, — и эта персональная мирная инициатива, которую Рузвельт хотел сообщить Тенно, в любом случае не могла бы предотвратить атаку на Перл Харбор.

11.1.3. Эти эпизоды обнаруживают некоторый общий метод криптоанализа, метод вероятного слова. Такие слова часто основываются на текущих событиях; в подобном случае сообщение должно быть перефразировано (чтобы избежать вероятных слов). В Первую мировую войну французские войска часто совершали атаки на немецкие позиции просто для того, чтобы вызвать некоторые «вероятные слова» в немецких радиопередачах — хорошо еще, что солдаты редко знали, ради чего они рискуют своей жизнью.

Во Второй мировой войне англичане топили светящиеся бакены, которые отмечали путь через заминированный противником вход к Кале только для того, чтобы вызвать немецкие сообщения, содержащие вероятное слово /leuchttonne/ (светящаяся бочка, *нем.*) (разд. 14.1).

Помимо таких слов, как атака, бомбардировка и т. д. военные сообщения содержали сокровища бросающихся в глаза слов и стереотипных фраз таких, как штабы, генеральный штаб, дивизия и радиостанция. Одно и то же сообщение, повторяемое ежедневно, даже если это лишь сообщения «нечего сообщать», — может иметь разрушительный эффект. Это дает зацепку для применения метода вероятного слова, подобного словам любовь, сердце, огонь, пламя, горение, жизнь, смерть, которые Порта выписал как неизменные строительные блоки любовных писем. Стереотипные фразы могут свести к нулю выгоды изменения ключа: новый ключ может быть быстро выведен из повторяющихся последовательностей. Не каждый, разумеется, будет столь удачлив, как лейтенант Бертольд из службы G.2A.6 Американских экспедиционных сил, который перехватил радиопередачу в 07:40 11 марта 1918 г., состоявшую из серии цифр и зашифрованную, очевидно, каким-то новым ключом; а несколькими часами позднее он услышал сообщение такой же точно длины, но в буквах — адресат еще не получил новых шифровальных инструкций и попросил повторить передачу в старом ключе. Подобный же компромисс имел место с ручным шифром PLAYFAIR, используемым Немецким африканским корпусом, когда 1 января 1942 г. ключ был изменен.

Далеко идущие последствия имел аналогичный случай, когда существование только что поступившей 4-роторной шифромашины ENIGMA было разоблачено в конце 1941 г. несколькими тренировочными передачами, парал-

лельными с сообщением, зашифрованным 3-роторной ENIGMA. В результате Блетчли Парк оказался в состоянии вычислить прокладку проводов нового («греческого») ротора β еще прежде, чем новая ENIGMA была официально введена в строй 1 февраля 1942 г.

Даже в мирное время важно владеть профессией: потому что никто не знает ничего лучшего, чем шаблонные тексты и избитые фразы, передаваемые на маневрах. Если формулировка в достаточной мере лишена воображения, вся криптосистема может быть раскрыта прежде, чем будет сделан единственный выстрел. Как писал Хюттенхайн, «ошибочно делать главным методом шифрования метод, который применялся достаточно долго в ограниченном масштабе». Когда речь идет о стереотипных началах и концах, которых иногда не удается избежать (вспомним «Для Мёрфи», разд. 4.4), то даже метод «русского соединения» мало помогает, хотя неопытному взломщику кодов он может помешать найти правильный путь.

11.1.4. Сам факт передачи сообщения может быть существенным. Зная, что каналы связи будут перегружены во время серьезной военной операции, штабные офицеры обычно посылают свои персональные сообщения заблаговременно, на несколько дней раньше. Это называется эффектом нижнего белья. Поэтому, если бы позволяли условия, каналы связи держались бы постоянно открытыми, и тогда в течение спокойных периодов посылалось бы «фальшивое наполнение» — не тестовые фразы или отрывки из газетных статей, но любые несообразные и непериодические последовательности букв, если можно, то случайный текст, а еще лучше синтетический язык из мультиграмм, частота букв в которых была бы аналогична какому-либо естественному языку («traffic padding», поточная набивка, *англ.*). Длинные непериодические последовательности могут порождаться, начинаясь в случайных или нерегулярных местах текста, состоящего, скажем, из 10 000 слов. Еще лучше, применять предложенный Кюпфмюллером в 1950-х гг. метод «n-граммной аппроксимации». В этом методе текст получается из главного текста с помощью регистров сдвига. Берутся последние $n-1$ символов, и следующее слово ищется в главном тексте из условия, чтобы оно содержало эти последовательные символы; затем присоединяется следующий символ и указанный процесс повторяется.

Ниже приводится синтетический бессмысленный текст, полученный тетраграммной аппроксимацией из первой главы известного романа Томаса Манна.

> *thomas ist daher mit mein hand zeigen augen von geschaeftig im kreissigen mauemdisellschaeftwar zur seligen durchterlich hier familie hierheben herzigkeit mit eindrinnen tonyzu plaudertfuenf uhr erzaehlungich regeshaehm die konnte neigte sie dern ich was stuetzte heissgetuebrige wahrend tause*

Защита сообщений путем «набивки» приводит к резкому увеличению нагрузки на неуполномоченного дешифровальщика и тем самым замедляет дешифрование им подлинного сообщения. Но, с другой стороны, предполагае-

мый адресат должен быть весьма бдительным, чтобы не пропустить какое-нибудь редкое подлинное сообщение в этом потоке мусора.

11.1.5. Заполнение «иксами» (/x/), повторение какого-либо слова, или применение сдвоенных букв может представлять определенный риск. Вместо этого следует перефразировать предложение или использовать синонимы или омофоны (выбранные случайно!) — это включает и использование пустышек, скрывающих частичные повторения вблизи омофонов. Одно из основных правил профессиональной криптологии заключается в подавлении не только пунктуации, но также и пробелов между словами. И сколь же шокирующим является факт, что общей практикой в немецком *Вермахте* (если не было точных противоречащих приказов) было вставлять /x/ для точки, /y/ для запятой, /j/ для кавычек, и даже /xx/ для двоеточия и /yy/ для точки с запятой. Иногда числа, закодированные буквами, были огорожены знаками /y/ ... /y/. Важные слова дублировались: /апап/, /vonvon/ для «к» и «из», например. Применялись даже утроения букв: /bduuu/ для *Befehlshaber der U-Boote*, /okmmm/ для *Oberkommando der Marine* (главное командование военно-морским флотом, *нем.*). Правда, с другой стороны, немецкая буква «х»: /ch/ часто заменялась на /q/. Вместе с неизбежными «по приказу фюрера» и /Heil Hitler/, которые никто не отваживался опускать, такие глупые обычаи немцев были великолепными помощниками англичан во взломе ENIGMA. Англичане к ним настолько привыкли, что выходили из себя, когда дешифрование передачи ENIGMA приводило в начале и конце сообщения к ничего не значащим последовательностям, называемым англичанами *quatsch* (ерунда, *нем.*). Таким вещам гораздо большее внимание уделялось во времена Ардженти, чем в XIX и XX столетиях, когда люди чересчур уверовали в невозможность взлома перешифрованных кодов и других сложных методов. Подавление удвоения букв было одной из намеренных «ошибок» переписывания, которые рекомендовал Ардженти. Как мудро заметил Порта в 1563 г.: «Лучше пусть считают написавшего безграмотным, в противном случае придется слишком дорого расплачиваться за разоблачение его планов». К сожалению, чем старше по званию офицер, тем реже можно ожидать, что он продемонстрирует проницательность, необходимую для того, чтобы мириться с безграмотными текстами. Идеальный шифровальщик должен обладать холодным интеллектом, соединенным с поэтическим воображением и полным пренебрежением к традиционной орфографии. Конечно, заманчиво зашифровать «радио» и «станцию» отдельно или даже шифровать их побуквенно, подобно одному австрийскому шифровальщику, который был слишком ленив, чтобы найти в словаре необходимую комбинацию; и этим воспользовался Луиджи Сакко, взломав шифр в 1918 г. То же самое происходит и в случае злоупотребления нулями в качестве пробелов между словами; некоторые члены французского *résistance* (сопротивления, *фр.*) использовали слово «tabac» (беда, *фр. жарг.*), когда требовалось подкрепление, в результате была взломана двойная перестановка. Малейшая небрежность может иметь ужасные последствия. В эпизоде Бертольда (разд. 11.1.3) его шеф Мурман, который разоблачил планы немецкой атаки 21 марта 1918 г., написал командованию:

«Посылка одного только этого сообщения, безусловно, должна стоить жизни тысячам немцев».

11.1.6. Признаком хорошего офицера связи является то, что он объясняет своим подчиненным, какую роль может сыграть в руках врага малейшая ошибка шифрования, и что он контролирует их усилия. Живерж писал: «Хорошо кодировать — или не кодировать вообще. Передавая просто открытый текст, вы даете врагу всего лишь какой-то кусок информации. Плохо же кодируя, вы позволяете ему прочесть всю вашу корреспонденцию, а также корреспонденцию ваших друзей». Однако этот совет, сделанный из лучших побуждений, не следует интерпретировать буквально, — что лучше уж передавать радиосообщения без всякого их шифрования. Такая ситуация имела место в конце августа 1914 г. в русской армии в Восточной Пруссии потому, что войска еще не получили новые кодовые книги, а телефонные линии были перегружены или отсутствовали. В результате Гинденбург и Людендорф одержали победу под Танненбергом и стали национальными героями. В другом экстремальном случае немцы шифровали погодные сообщения во время Второй мировой войны; когда преобладал западный ветер в Европе, это часто снабжало англичан «вероятным словом». Немцам следовало бы лучше передавать такие малозначащие сообщения открытым текстом.

Рорбах рекомендовал, оценивая надежность какого-либо метода, рассматривать и степень его защищенности от ошибок шифровальщика, следуя принципу «человеку свойственно ошибаться». Принципы надежной разведки и контрразведки, конечно, при этом тоже играют важную роль.

11.1.7. Применение легко запоминаемых паролей и ключей дает незаконному дешифровальщику дополнительное подтверждение, что он добился успеха в восстановлении ключа. В частности, это верно, если ключ имеет какую-то особую важность для отправителя сообщения.

11.1.8. Нельзя недооценивать неудобства, возникающие из-за необходимости поддерживать криптографическую надежность. Регулярное изменение ключей обеспечивает работой всех, имеющих к этому отношение. Даже в этих условиях трудно понять, почему Госдепартамент США использовал до 1917 г. такие короткие ключевые слова как PEKIN и POKES, хотя еще Порта применял длинные CASTUM FODERAT LUCRETIA PECTUS ALGAZEL; Ардженти применял такие ключи, как FUNDAMENTA EIUS IN MONTIBUS SANCTIS или GLORIA DICENTUR DE TE QUIA POTENTER AGIS. Как писал Виженер, «чем длиннее ключ, тем труднее взломать шифр». Необходимое условие многоалфавитного шифрования состоит в квазинепериодичности; другими словами, если ключ является периодическим, он не должен быть существенно короче сообщения. В случае необходимости длинное сообщение должно быть разрезано, и его части зашифрованы разными ключами. Предупреждение Хитта, что «ни одно сообщение не является надежным без того, чтобы ключевая фраза была сравнимой по длине с самим сообщением» не означает, что сообщение может быть столь же длинным, как и период ключа — независимо от того, используется или нет шифровальная машина. Сообщения более чем из 1000 символов в любом случае рискованны, так как автоматиче-

ские приемы дешифрования для машины М-209, например, хорошо работают с сообщениями порядка 800 символов или больше (чистый криптоанализ, разд. 22.2.1). Сообщения из 200–300 символов обычно при таком подходе надежны; максимальная длина, допускаемая М-209, составляет 500 символов. Предел в 180 символов применительно к машине ENIGMA вырос до 250, когда были введены роторы VI-VIII (разд. 8.5.3). Применение индивидуального ключа представляет собой дополнительное организационное усилие и требует дополнительной надежности и мер против шпионажа. Существует много ситуаций, где это становится невозможным, например, изолированные позиции, куда (надежная) доставка ключей не может быть гарантирована, или где запас ключей может попасть в руки врага и быть использован в целях обмана.

11.1.9. В этом пункте рассматривается вопрос о том, как принимающая станция может судить, исходит ли полученное радиосообщение от законного партнера или от самозванца. Криптографические меры аутентификации, упомянутые в разд. 10.5, можно дополнить стеганографическими методами (контроль надежности) вроде вставных нулевых символов в специфических местах криптотекста или создания намеренных ошибок правописания в согласованных точках открытого текста, не говоря уже о «почерке» — индивидуальном стиле передачи («radio fingerprint») оператора, который ведет *Funkspiel* (радиоигру, *нем.*).

11.1.10. Стоит упомянуть самый банальный и наиболее грубый криптоаналитический метод: похищение криптодокументов путем шпионажа, воровства, разбоя или получения их в качестве военных трофеев после сражения. Лучшим способом защитить себя от этого является очевидное, но часто игнорируемое: «*чего больше не существует, то не может попасть в чужие руки*!». Карл Вейерштрасс имел мужество поступить так с письмами Софьи Ковалевской. Это правило особенно приложимо к индивидуальным ключам, после использования они должны немедленно уничтожаться (разд. 8.8.2). Планирование многократного использования индивидуальных ключей (разд. 8.8.7) абсурдно.

11.1.11. Избитым стало выражение, что война может принести богатую добычу. Это, в частности, приложимо к криптологическим материалам; например, 12 февраля 1940 г. немецкая субмарина U-33 была захвачена Британским королевским флотом. В других отношениях безупречный радиооператор Кумпф забыл бросить за борт роторы шифромашины ENIGMA. Поляки уже работали над коммутацией первых пяти ее роторов, но роторы VI и VII были в новинку для англичан. В августе 1940 г. был захвачен также и ротор VIII. 26 апреля англичане захватили немецкий корабль «Полярес». Там они обнаружили соответствующие открытый текст и криптотекст за предшествующие 4 дня (правда, этого оказалось недостаточно, чтобы полностью взломать военно-морскую шифромашину ENIGMA). Оперативные инструкции захваченной в июне субмарины U-13 тоже были мало полезны. Прорыв произошел в следующем году: 4 марта 1941 г. при захвате «Кребса» в норвежском Вест-фиорде удалось добыть не только два уже известных ротора, но также полные ключи на предшествующий месяц. Это позволило в Влет-

чли Парк прочитать в марте 1941 г. все февральские сообщения немецкого флота. В результате стало возможным восстановить применявшиеся таблицы биграмм.

7 мая специальная атака на метеорологическое судно «Мюнхен» принесла полные ключи на июнь (ключи же на текущий месяц и сама шифромашина были выброшены за борт согласно инструкции), а также *короткий «погодный ключ»*. Субмарину U-110 заставили глубинными бомбами подняться на поверхность у западного побережья Ирландии 9 мая 1941 г. Трофеи включали, кроме еще одной шифромашины ENIGMA, золотое сокровище правил пользования, включая таблицы биграмм Баха для кодирования индикатора (разд. 4.1.2) и *тетрадь коротких сообщений*. И, наконец, произошла вторая запланированная атака на метеорологическое судно «Лауенбург» 28 июня 1941 г. Немцы ухитрились выбросить за борт машину ENIGMA, но англичане захватили полные ключи на июль. Это позволило Блетчли Парк взломать 3-роторную машину ENIGMA немецкого ВМФ, и, начиная с этого момента, англичане могли регулярно прослушивать радиосвязь с субмаринами — с запозданием всего на несколько часов. Потери британского флота стали соответственно много меньше.

Введение 4-роторной машины ENIGMA 1 февраля 1942 г. изменило ситуацию, но в декабре того же года передачи снова стали регулярно дешифроваться и союзники начали одерживать верх в подводной войне. Это произошло после нового захвата подводной лодки, которому способствовала невероятная тупость со стороны Мэртенса и Штуммеля, глав секретной службы B-Dienst немецкого ВМФ. Захват подводной лодки U-559 в Порт Саиде 30 октября 1942 г. снабдил англичан новым изданием *тетради коротких сообщений* и вторым экземпляром *короткого погодного ключа*, которые и сами по себе были прекрасной находкой. Вдобавок Арчер сумел 13 декабря 1942 г. дешифровать одно сообщение, которое показало, что когда 4-роторная машина ENIGMA использовалась для связи с береговыми станциями, имевшими лишь 3-роторные машины, ее четвертый ротор («греческий») ставился в нулевую позицию. Таков был обычай, который делал связь возможной. Явной глупостью было то, что трехбуквенная кольцевая установка 3-роторной машины ENIGMA была всегда такой же, как и первые три буквы кольцевой установки 4-роторной машины ENIGMA. Это было вовсе не обязательно, а делалось только для удобства создания ежемесячных приказов. Отсюда вытекало, что если враг (англичане) знает установки для 3-роторной машины ENIGMA, то требовалось всего 26 попыток, чтобы найти установку греческого ротора. Таким образом, начиная с 13 декабря 1942 г. англичане окончательно взломали 4-роторную машину ENIGMA, работавшую в ключевой сети TRITON подводных лодок немецкого ВМФ (созданной в 1941 г.). И даже введение второго дополнительного греческого диска 1 июля 1943 г. было недостаточным, чтобы помешать англичанам читать сообщения, зашифрованные с помощью ENIGMA, до самого конца войны. Англия, конечно, тоже имела потери в этом соперничестве. В 1940 г. немецкий вспомогательный крейсер «Комет» захватил биграммные шифры и кодовые книги торгового флота (разд. 4.1.2, 4.4.5).

Союзники об этом ничего не знали до тех пор, пока не изучили немецкие архивы после войны.

11.1.12. Но не только такие важные вещи помогают врагу; даже малейшие детали могут случайно выдать информацию огромной важности. В августе 1941 г. в руки англичан на побережьи Исландии попала немецкая субмарина U-570 почти без единой царапины. Деревянный ящик для ENIGMA был пустым, но в нем оказалось гнездо для 4-го ротора. Это было подтверждением того, о чем англичане уже подозревали, учитывая упоминания о четвертом роторе ENIGMA в захваченных руководствах, — что ввод этой версии неизбежен. Таково обилие мелких деталей, вытканных на гобелене, который хранит в себе состояние криптоанализа. Каждый обрыв нити задерживает дешифрование на некоторое время, возможно, навсегда.

11.2. Принципы криптологии

> Эта машина (ENIGMA) была бы неуязвима,
> если бы она использовалась правильным образом.
>
> *Гордон Уэлчмен*, 1982 *г.*

> Ни одна шифровальная машина не может делать свою работу правильно, если она используется небрежно. Во время Второй мировой войны особенно много небрежности имело место на стороне Оси (т. е. Германии и ее союзников).
>
> *Сифер Девур, Луи Кру*, 1985 *г.*

В течение веков криптология накопила бесценный опыт — даже открытая литература свидетельствует об этом. Эти знания, рассеянные среди множества источников, можно выразить в виде нескольких принципов криптографической работы, следование которым необходимо в том числе и для защиты против незаконного дешифрования. В нашу эру компьютеров эти принципы становятся особенно важными.

11.2.1. К естественным способностям человека относятся смелость, стойкость и противодействие опасности. Но эти положительные качества имеют тот побочный эффект, что человек склонен переоценивать свои возможности. Поэтому

Принцип № 1: Нельзя недооценивать противника.

Как мы видели, немецкие криптографы не подозревали, что Союзники могут проникнуть в их криптосистемы. Были отдельные случаи[2)] опасения, но официальное мнение о надежности упрямо сохранялось. И лишь воен-

[2)] Уже в 1930 г. лейтенант Лукан (Henno Lucan), второй офицер связи линкора Эльзас, указал на одну слабость шифромашины ENIGMA. С вводом штепсельного коммутатора эти опасения, казалось, были устранены.

но-морской флот был единственной военной службой, которая решила усовершенствовать свои криптомашины перейдя 1 февраля 1942 г. от 3-роторной машины ENIGMA к 4-роторной ENIGMA, в которой было 8 сменных роторов вместо 5, используемых Армией и Военно-воздушным флотом. Тем самым признавалось, что необходимо увеличить надежность шифромашин. Однако немецкий Генеральный штаб был уверен в победе и был интеллектуально не готов прислушаться к серьезным предостережениям. И даже в Военно-морских силах существовала глубоко укоренившаяся уверенность в невзламываемости машины ENIGMA. Еще в 1970 г. контр-адмирал Бонатц, когда-то штабной офицер секретной службы B-Dienst Военно-морского флота, в своей книге выразил уверенность в том, что Союзники, несмотря на то, что они захватили несколько экземпляров ENIGMA, не могли взломать немецкую криптосистему, разве что они могли читать немецкие сообщения в течение некоторого ограниченного времени.

И не одни немцы были столь самоуверенны. Американские криптологи тоже не могли вообразить, что Рорбах уже взломал их M-138-A. Агентство секретной связи американской армии безуспешно пыталось взломать свою новую роторную машину M-134-C (SIGABA). Что бы это могло означать? Почему немцы не могут работать так же хорошо, как и англичане, которые сумели взломать ENIGMA? Показательно, что именно Рузвельт, самый интеллектуальный среди союзных лидеров, всегда не очень доверял утверждениям криптологов. Может быть, он лучше других знал присущее человеку свойство игнорировать то, что ему неприятно?

Королевскому ВМФ потребовалось целых три года, чтобы убедиться в том, что немецкая секретная служба B-Dienst читает некоторые из их шифров. Взлом машины ENIGMA дал окончательное доказательство того, что, по крайней мере, военно-морской шифр №3, главная криптосистема войсковых соединений конвоя в северной Атлантике (немецкое кодовое название Франкфурт), был взломан немцами. Его заменили военно-морским шифром №5, и таким образом с середины июня 1943 г., как стало известно после войны, немцы были отсечены от источника информации. Что могло случиться, если бы немцы подобным же образом узнали о ненадежности своих машин ENIGMA?

Может быть, ничего, как показывает чрезвычайно глупый случай постоянной недооценки англичан контр-адмиралом Мэртенсом и его начальником штаба Штуммелем. Это случилось в середине 1943 г.: дешифрование сообщений союзного конвоя показало, что американцы предполагают наличие двадцати немецких подводных лодок в узком квадрате карты. Действительно, «волчья стая» *Meise* (синица, *нем.*) из 18 субмарин находилась в этом квадрате. Командующий подводными лодками Гроссадмирал Дениц (1891–1980 гг.) приказал Мэртенсу расследовать этот случай так же, как он сделал в 1941 г., когда была захвачена субмарина U-570. Снова Мэртенс реабилитировал ENIGMA. Он сказал, что английская противолодочная служба сама передала заявление, что информация союзников о дислокации подводных лодок получена с помощью пеленгации. Мэртенс спас свою голову, ложно объяснив, что лодки обнаружили с помощью H2S (немецкое кодовое назва-

ние Rotterdam-Gerät, роттердамский прибор, *нем.*), радара, работавшего на длине волны 9.7 см, обнаруженного 2 февраля 1943 г. в сбитом над Роттердамом английском бомбардировщике. Дениц принял это к сведению, но сохранил подозрения, и, в конце концов, уволил Мэртенса после инцидента с конвоем SC 127 12 марта 1944 г., который был следствием измены или отсутствия надежного шифрования. Сегодня, известно, что Мэртенс стал жертвой ловкой дезинформации англичан.

Русские тоже сумели проникнуть в шифросистему ENIGMA. Они подняли затонувшую 30 июля 1944 г. в Финском заливе подлодку U-250 и захватили машину ENIGMA. Мнения о том, как далеко зашли русские успехи в этом направлении, разделились. В то время как в одном из немецких документов от января 1943 г. констатируется: «Достоверно известно, что в некоторых частных случаях русские добивались успеха в дешифровании ENIGMA», — Томас в 1978 г. заявил, что после десяти лет детального изучения этого вопроса он нигде не нашел никаких свидетельств, доказывающих, что русские когда-либо могли дешифровать немецкие радиосообщения ENIGMA.

«Вопрос о том, проникли ли русские в криптосистемы США, дебатировался довольно часто, в частности, после того, как Левин, журналист русского происхождения, специализировавшийся в советологии, из многочисленных разговоров, которые он вел с генералом Вальтером Кривицким, дезертировавшим главой советской разведки в Восточной Европе, стал с середины 1939 г. убежден в том, что советские криптоаналитики читали американские коды» (Кан).

11.2.2. Изобретатели криптосистем могут переоценивать надежность своих систем. «Почти каждый изобретатель криптосистемы убежден в неразрешимости его детища», писал Кан. Довольно трагикомичным примером является Базерье. Работая для французского правительства и армии, он взломал несколько новых шифров, дешифровав данные ему тестовые примеры. Затем он изобрел свою собственную шифросистему и сразу окрестил ее абсолютно надежной. Маркиз де Виари, чье изобретение Базерье вдребезги разбил незадолго перед тем, взял реванш. Он даже создал специальный метод криптоанализа (разд. 14.3), применимый к широкому классу аппаратов, от Джефферсона и Базерье до М-94 и М-138-А, каждый из которых использует семейства неродственных алфавитов. И вот здесь мы приходим к принципу Керкхоффса.

Принцип № 2: Если кто и может судить о надежности криптосистемы, то только криптоаналитик.

Это мог сказать еще Порта, однако этот принцип был сформулирован только Керкхоффсом в 1883 г. Он критиковал «арбитров», обсуждавших надежность метода шифрования путем подсчета числа столетий, потребных для полного перебора всех возможных комбинаций. В действительности же такие комбинаторные подсчеты могут дать лишь верхнюю оценку усилий, необходимых в худшем случае, для грубейшего из всех криптоаналитических методов, а именно переборного поиска, называемого также «атакой грубой силы».

Поэтому повсюду в цивилизованном мире правительственные службы (а также некоторые неправительственные) имеют двойную обязанность: проектировать надежные криптосистемы, а также пытаться взломать их. «Со взломщиками кодов и создателями в одном и том же агентстве, АНБ, проводит больше криптографических экспертиз, чем любая другая организация в стране, общественная или частная», — писал Бейкер не без гордости. Он известный юрист, совсем недавно был главным юристом АНБ. Его похвала звучит даже лучше, чем похвала из независимого источника.

11.2.3. Керкхоффс был одним из первых, кто занимался криптографией с практической точки зрения. Обсуждая вопрос о легкости обработки (чтобы разобрать ее позднее), он писал: «Полезно проводить различие между криптосистемой, нужной для краткого обмена сообщениями между несколькими отдельными людьми, и методом криптографии, предназначенным для обмена корреспонденцией между разными армейскими командирами в течение неограниченного срока». Он провел различие между криптосистемой как классом методов (по-французски *système*) и ключом в узком смысле, и постулировал, как упоминалось выше: «*Il faut qu'il puisse [le système] sans inconvénient tomber entre les mains de l'ennemy*» (никакое затруднение не должно возникнуть, если система попадает в руки врага, *фр.*). Это привело Шеннона к более точной формулировке:

Принцип № 3: При обсуждении надежности шифрования класса методов следует принимать во внимание, что противнику известен этот класс методов. («Противнику известна используемая система», *Шеннон.*)

О. Хорак выразил эту мысль другими словами: «Безопасность слабого метода шифрования не увеличится, если сохранить метод в секрете».

По практическим причинам в некоторых ситуациях одни методы используются чаще других. В частности, врожденный консерватизм аппарата управления имеет определенные предпочтения, которые невозможно скрыть от противника («философия шифрования»). Кроме того, грубое разделение методов шифрования, таких как перестановки, одноалфавитное или многоалфавитное шифрование, возможно при помощи простых тестов. Существуют также так называемые правила большого пальца, подобные критерию Сакко о том, что короткий (скажем, не более чем из 200 символов) криптотекст, включающий все символы алфавита, по всей вероятности, зашифрован многоалфавитно.

Машины и другие устройства, в том числе и зашифрованные документы, могут во время боя попасть в руки врага или быть украдены. Это относится и к машинам типа ENIGMA. Строго следуя доктрине Керкхоффса, ENIGMA должна была бы к началу Второй мировой войны получить пятый ротор; причем роторная позиция (Walzenlage) должна была бы изменяться с первого же дня каждые 6 часов (а не всего трижды в день, начиная с 1942 г.), и каждый квартал множество роторов должно было бы полностью меняться. Правда, это было бы не слишком легко, учитывая десятки тысяч использовавшихся аппаратов ENIGMA, но именно так следовало бы сделать.

Но, как писал Кан, «у немцев не было монополии на криптографические неудачи, и в этом отношении англичане были ничуть не логичнее немцев». И американцы были столь же нелогичны. Их шифромашина М-209, сконструированная Хагелином и построенная по лицензии, была значительно менее надежна, чем ENIGMA, она использовалась также итальянским ВМФ (С-38m), партнером Германии. Не удивительно, что немцы в Северной Африке в 1942 и 1943 годах часто знали цели и время американских атак. А англичане, которые могли читать С-38m, знали все, что им было нужно о снабжении фельдмаршала Эрвина Роммеля.

11.2.4. Желание криптографа не сделать дешифрование слишком легким для противника приводит к усложнению известных методов. Композиция методов (гл. 9) долгое время использовалась именно с этой целью, главным образом комбинируя существенно разные методы, вроде перестановки одноалфавитной подстановки или перешифрования кода многоалфавитным шифром. Некоторые криптоаналитические методы, однако, часто бывают нечувствительными к таким усложнениям. В лучшем случае ничего не приобретается, а в худшем случае такая комбинация дает нечто непредвиденное. Живерж (1924 г.) сформулировал

Принцип № 4: Поверхностные усложнения могут быть иллюзорными, так как они могут дать криптографу ложное ощущение надежности.

В типичном случае кто-то исключает с самыми лучшими намерениями, но совершенно напрасно, тождественное преобразование в качестве шага шифрования ВИЖЕНЕР. При таком подавлении, которое не должно оставить незатронутой ни одну букву, ни одна буква не может представлять себя. Но это свойство позволяет для достаточно длинного вероятного слова найти несколько позиций, где оно может встретиться (гл. 14). То же свойство справедливо для систем с одноцикловыми алфавитами, для всех цилиндрических устройств от Джефферсона и Базерье до М-94 и для всех ленточных устройств от Хитта до М-138-А. Более того, все криптосистемы с настоящими взаимообратными алфавитами, обладают этим свойством, включая ENIGMA, благодаря изобретению отражающего ротора, который можно назвать шедевром *иллюзорной сложности*. По этому поводу Уэлчмен заметил: «Было бы возможно также, хоть это и трудно, сконструировать машину, подобную ENIGMA, с элементами самошифрования, которая выводила бы из строя большинство наших методологических находок, включая «females» (разд. 19.6.2.1).

11.2.5. Наконец, последним и, возможно, наиболее важным пунктом является человеческий фактор. Надежность шифрования напрямую зависит от шифровальщика. Незаконный дешифровальщик живет за счет ошибок, упомянутых в начале этой главы.

Прежде всего, существуют компромиссы.

Компромисс открытый текст—криптотекст: повторение передачи открытым текстом.

Криптотекст—криптотекст компромисс ключей: передача двух «изологов», т. е. одного и того же открытого текста, зашифрованного с разными ключами (в частности, публичные ключи допускают этот компромисс).

Открытый текст—открытый текст компромисс ключей: передача двух разных открытых текстов, зашифрованных с одним и тем же ключом.

Кроме того, существуют элементарные правила хорошего криптографического языка: не применять сдвоенные буквы и частые буквенные комбинации вроде /ch/ и /qu/, подавлять знаки пунктуации и особенно подавлять пробелы между словами, использовать омофоны и пустышки профилактически — против атак вероятного слова. Открытый текст, хорошо приготовленный для шифрования, орфографически неправилен, лингвистически скуден и стилистически ужасен. Какому командующему генералу понравилось бы таким образом формулировать свои приказы, какой посол направил бы такое сообщение своему правительству? Ответ прост: они не должны это делать сами, за них это должны делать их криптоофицеры. Как Рузвельт, так и Черчилль подчинялись во время Второй мировой войны этим потребностям криптографической надежности. Только Мэрфи не подчинялся.

Вдобавок, послы и генералы обычно не любят наблюдать за своими шифровальщиками; на самом деле, большинство из них даже не понимает их нужд и криптологически невежественны. Когда Уитстоун изобрел специальную биграммную подстановку, которую позже назвал ПЛЕЙФЕЙР (разд. 4.2.1), он не мог преодолеть неприязни Foreign Office к сложным шифрам. Генералы Наполеона шифровали свои сообщения лишь частично, и так же делали итальянцы до битвы при Изонцо в 1916 г.

Поэтому важным принципом службы связи является то, что радиоперехват и надзор за своими собственными подразделениями по меньшей мере так же важен, как и прослушивание противника. К этому Эрих Хюттенхейн прибавил: «*Ein Verbündeter, der keine sicheren Chiffrierungen verwendet, stellt ein potentielles Risiko dar*» (союзники, которые не используют надежную криптографию, представляют собой потенциальный риск, *нем.*).

Часто говорят, что *ошибка криптографа — единственная надежда криптоаналитика*. Эта надежда оправдывается: всегда существуют нервные стрессы шифровальщиков на дипломатической и военной службе, и, вероятно, ошибки шифрования часто случаются. Чем более усложненным является метод, тем более искаженным будет открытый текст после окончательного дешифрования. При недостатке времени опасность повтора сообщения без аккуратной перефразировки кажется неизбежной. Совет Живержа был таким: «*Chiffrez bien, ou ne chiffrez pas*» (шифровать хорошо — или не шифровать вообще, *фр.*). Совет Рорбаха — следующим:

Принцип № 5: При оценке надежности шифрования класса методов должны приниматься во внимание криптографические ошибки и другие нарушения правил безопасности.

Хороший криптолог знает, что он не должен рассчитывать ни на что, даже на постоянные ошибки противника. Он должен быть особенно критичен

к своим собственным возможным ошибкам. Наблюдение за его собственными шифровальными привычками каким-нибудь *advocatus diaboli* (адвокатом дьявола, *лат.*) абсолютно необходимо. Это слишком ясно показали действия немцев во Второй мировой войне. Милнер-Барри написал: «Если бы не человеческие ошибки, осложненные единственным неудачным техническим решением, то ENIGMA была бы совершенно надежной машиной».

Этот недостаток, казавшийся разумной идеей (7.3.2), состоял в придании шифрованию взаимнообратного характера. В результате за облегчение шифрования была заплачена высокая цена.

11.3. Критерии Шеннона

Если кто-нибудь готов следовать данному выше совету, он доложен решить вопрос о том, какой метод выбрать. Ответ зависит, с одной стороны, от степени желаемой надежности, а с другой, от вкладываемых усилий. Клод Шеннон составил следующий список[3)] из пяти критериев для определения класса криптографических методов:

(1) Степень требуемой надежности шифрования	Какое количество информации получит противник из определенного объема перехваченного материала.
(2) Длина ключа	Как короток ключ, насколько просто обращение с ним?
(3) Практическое выполнение шифрования и расшифрования	Какой объем работы нужно выполнить?
(4) Разрастание криптотекста	Насколько криптотекст длиннее открытого текста?
(5) Распространение ошибок шифрования	Как далеко распространяются ошибки шифрования?

Эти критерии настолько противоречивы, что неизвестно ни одной криптосистемы (и, вероятно, такая система просто не может существовать), которая удовлетворяет всем этим требованиям. С другой стороны, ни один из критериев нельзя игнорировать.

Если опустить пункт (1), то приемлем даже открытый текст. Если полностью опущен пункт (2), то приемлем индивидуальный ключ. Если отсутствуют пункты (3) и (4), то существуют экзотические криптосистемы, которые удовлетворяют всем остальным пунктам. Если исключен пункт (5), то методы, основанные на тщательном перемешивании, будут удовлетворять всем остальным критериям.

[3)] Клод Шеннон (Claude E. Shannon *A Mathematical Theory of Cryptography*, Internal Report, September 1, 1945. Опубликовано в книге: Communication Theory of Secrecy Systems. Bell System Technical Journal, v. 28, pp. 656–715 (October 1949).)

Современная криптография имеет тенденцию в зависимости от ситуации использовать индивидуальные ключи (которые требуют постоянного и надежного механизма распределения ключей) или применять смешанные методы (что требует отсутствия помех, т. е. исправляющих ошибки каналов связи).

В ситуации крайней секретности, например, при переговорах между главами государств в чрезвычайных обстоятельствах, использование индивидуальных ключей является вполне нормальным, так как в подобных ситуациях объем передаваемых сообщений не очень велик. Но и при передаче больших объемов информации индивидуальные ключи также могут использоваться.

Уинтерботам, ответственный за безопасность материалов, полученных дешифрованием радиосообщений немецких ENIGMA и SZ42 Блетчли Парк, рассылал их под названием ULTRA по полевым соединениям, и при этом добивался, чтобы они были защищены индивидуальными ключами. Это показывает, с одной стороны, насколько важны были эти секретные материалы, если для их передачи использовалась трудоемкая связь с индивидуальными ключами, а с другой стороны, как надежно англичане оценивали индивидуальные ключи, что, конечно, справедливо. Могли ли немецкие офицеры требовать соблюдения таких же строгих мер безопасности?

В области коммерции совместно с DES последние два десятилетия использовались различные смешанные методы. Непрерывная критика надежности представленного *de facto* стандарта, наверное, уже смягчилась бы, если бы была увеличена длина ключа, которую многие рассматривают как чересчур короткую.

11.4. Криптология и права человека

Поскольку криптографические методы используются, то даже любители пытаются их взломать. Но сегодня любитель даже с доступом к компьютеру средней мощности поймет, что трудно проникнуть в шифр, который удовлетворяет профессиональным стандартам. Агентство Национальной безопасности США, однако, желает сохранить право надзора над каждым коммерческим каналом связи, который попадает под подозрение. Можно ожидать, что американское правительство не разрешит разведывательным службам своего потенциального противника — такие еще существуют — чтобы они построили коммуникационную сеть под крышей частного коммерческого предприятия. Прошли времена, когда Стимсон, Государственный секретарь при президенте Эдгаре Гувере, мог расформировать Черный кабинет Госдепартамента (1929 г.), а потом в своей автобиографии (1948 г.) оправдать себя следующим образом: «Джентльмены не читают писем друг друга». Даже президент Картер не выказал никаких угрызений совести. Может быть, это показывает, что американцы при Картере не преуспели в чтении русских сообщений? Конец Холодной войны означал лишь сокращение, а вовсе не прекращение скрытой опасности шпионажа.

11.4.1. С криптологией имеют дело не только дипломатические и военные администрации разных стран. Не следует забывать, что криптология оказыва-

ет сильное влияние на постоянный конфликт интересов между гражданином или личностью и государством, представляющим общество в целом. С одной стороны, существует неоспоримое право гражданина (или организации) защищать свои личные (или ее частные) интересы при помощи эффективных криптосистем, а с другой стороны, существует конституционный долг государства защищать свою внутреннюю и внешнюю безопасность, которая может потребовать проникновения в некоторые шифрованные сообщения с разведывательными целями.

Позиция государства была выражена Хокинсом, помощником президента США по национальной безопасности, 3 мая 1993 г. следующим образом: «Принуждение закона и безопасность национальных коммуникаций убеждают, что если право народа на секретность торжествует и разрешается свободное применение криптографии, то преступники и шпионы будут избегать нашего прослушивания». Секретность писем не абсолютна даже в цивилизованных странах, и в случаях, регулируемых законом, она может быть приостановлена для блага государства, но не для блага частных лиц. Шифрованные сообщения не являются исключением — именно использование криптографии порождает начальное подозрение.

С другой стороны, именно в США, где большинство граждан смотрят на обладание огнестрельным оружием как на их конституционное право, обладание криптографическими средствами не рассматривается как государственная монополия. Европа с ее во многом отличной историей, не ушла так далеко вперед в этом отношении.

Диффи описал это положение короткой формулой: «...личная секретность против правительственной секретности». В Европе имеется достаточно причин, чтобы настаивать на свободе от авторитарного государства. Таким образом, есть необходимость найти внутри общей структуры каждой политической конституции средства, способные регулировать правительственный криптоанализ; его граница должна быть определена. Странно, но более крупные государства имели здесь большие трудности в достижении результатов, чем более мелкие; Австрия, например, является в этом отношении более продвинутой страной, чем Германия.

Какое-то решение здесь необходимо также в интересах мировой торговли. В США правило торговли криптологическим оборудованием с иностранными партнерами таково: «Шифрование с целью аутентификации сообщения возможно, тогда как шифрование с целью сохранения информации секретной заставляет удивленно поднять брови» (Бернштейн). В развитие этой темы: Циммерман, по словам его юриста, в 1994 г. был обвинен в нарушении *International Traffic in Arms Regulations*, потому что он разместил в Интернете и таким образом сделал свободно доступной криптосистему PGP (*Pretty Good Privacy*, разд. 9.6.5), которая считалась военным оборудованием («криптографические устройства, а также классические и неклассические данные, относящиеся к криптографическим устройствам, Категория XIII»). Обвинение было снято в 1996 г., однако ситуация остается по-прежнему неудовлетворительной.

11.4.2. Чтобы урегулировать этот конфликт между защитой частной сферы законопослушного гражданина, гарантирующего конфиденциальность своего сообщения, с одной стороны, а с другой — выполнением конституционных функций государства, были предложены различные схемы.

1. Ограничение использования криптосистем в гражданской области требованием официального одобрения, — либо в индивидуальных случаях, либо определенных систем, распространяемых коммерческим поставщикам (запрет отдельных методов, исключая применение индивидуальных ключей, недостаточен, так как лишь провоцирует обман).

2 a. Ограничение надежности шифрования регулированием доступности соответствующих криптосистем в гражданской области. Агентство, производящее криптосистемы, может обеспечивать гражданам некоторые криптоаналитические гарантии, и тем самым увеличивать стимул для добровольного подчинения (коммерческий поставщик с государственной поддержкой на рынке может быть лидером).

2 b. Аналогично (2 a), но в соединении с запретом использования других криптосистем в гражданской области.

3. Система (*escrow system*) условного депонирования, т. е. условного вручения документа, депонированного (хранящегося) у третьего лица, требующая депонирования полных данных для каждой криптосистемы, используемой в гражданской области. При этом агентство условного депонирования является независимым и требуется сохранение конфиденциальности.

Дальнейшие предложения могут быть комбинацией указанных выше вариантов. Есть основания ожидать, что разные демократические государства придут к разным решениям в пределах своего суверенитета. Во Франции, например, уже установлено решение в направлении пути (1), который мог бы рассматриваться как недемократический, Нидерланды некоторое время склонялись к такому же решению. В Германии одно время имелась выраженная тенденция к решению, подобному (2a), с недавно учрежденным органом BSI, подчиненным МВД; можно ожидать, что это приведет к решению типа (2b). Эта либеральная позиция была в июне 1999 г. подтверждена новым федеральным правительством. Еще не очевидно, какое решение будет принято Соединенным Королевством с его «Official Secrets Act» (законом о государственной тайне). В США в 1993 г. администрацией Клинтона защищалось решение в направлении пути (3) (*key escrow system*[4]) — ключевая система условного депонирования, *англ.*, см. разд. 11.4.3). Это вызвало громкий протест, и после дискуссии в 1996 г. появилась модификация в направлении (2a). Для Европейского Союза ни схема (1), ни схема (3) не являются наиболее вероятным выбором.

Более того, можно себе вообразить, что означали бы такие разные подходы для международных коммерческих поставщиков. Уже сейчас торговля в области криптографических устройств пересекает международные грани-

[4] *Escrowed Encryption Standard* (EES), Federal Information Processing Standards Publication (FIPS PUB) 185, Feb. 9, 1994.

цы с большим трудом: «Международное использование шифрования ввергает пользователя в легальную трясину импорта, экспорта и регулирования секретности, которые часто невразумительны, иногда противоречивы» (Бернштейн). Даже простое международное путешествие с ноутбуком уже является потенциально наказуемым. Заместитель помощника государственного секретаря Марта Харрис заявила 4 февраля 1994 г.: «Мы не будем больше требовать, чтобы американские граждане получали, как прежде, экспортную лицензию для временного персонального использования криптографической продукции за пределами США. В прошлом это требование вызвало задержки и неудобства во время деловых поездок».

11.4.3. Система EES (Escrowed Encryption Standard, шифрового стандарта условного депонирования, *англ.*) основана на алгоритме шифрования SKIPJACK (разд. 9.6.5), использованном в специализированном процессоре CLIPPER. Чтобы обеспечить условное депонирование двумя агентами, имеются две уникальных ключевых процессорных компоненты. Эти компоненты передаются правительственному чиновнику для электронного наблюдения лишь в соответствии с изданными и одобренными Генеральным прокурором процедурами. Ключевые компоненты необходимы для создания сложением по mod 2 уникального процессорного ключа.

80-битный текст («сессионный ключ») KS, согласованный с партнерами или распределенный по секретной схеме, служит, как и в DES, для установки начального положения шифрующего блока c_0, который может быть применен одноалфавитно способом, соответствующим *Electronic Code Book* (электронной кодовой книге, *англ.*), или преобразован в автоключ способом, соответствующим *Cipher Block Chaining* (разд. 9.6.3). Самым важным является то, что этот процессор имеет регистр памяти, LEAF («Law Enforcement Access Field»), в котором текущий сессионный ключ сохраняется в зашифрованном виде; для его получения можно использовать уникальный процессорный ключ. Поэтому получив судебный ордер и используя специальное дешифрующее приспособление, можно проинспектировать нужный канал. Всякий раз новый обмен информацией начинается с новым сессионным ключом; расшифровывающее устройство должно быть способно выделить из LEAF и расшифровать этот сессионный ключ. За исключением начальной задержки с получением ключей, выделенные сообщения могут расшифровываться в режиме реального времени в течение всего срока наблюдения. При этом даже голосовая связь может быть изучена в цифровой форме.

В отличие от DES, алгоритм SKIPJACK сохранялся в секрете в силу полномочий «защищать LEAF», несмотря даже на то, что для стойкости против криптоаналитической атаки не требуется, чтобы алгоритм сохранялся в секрете. Кроме того, алгоритм SKIPJACK был зарегистрирован как СЕКРЕТНЫЙ, НЕ РАЗРЕШЕННЫЙ ДЛЯ ИНОСТРАННЫХ ГРАЖДАН. Поэтому он не годился в качестве международного стандарта *de facto*. Но в 1998 г. эти ограничения, наконец, были сняты.

На самом деле, «секрет» LEAF является довольно примитивным. Он легко доступен, так что незаконный дешифровальщик может найти его, приложив

достаточные усилия. Деннинг рассмотрела некоторые практические вопросы, возникающие при этом. Микали («Fair Cryptosystems», патент США 5 276 737, от 4 января, 1994) предложил улучшенные криптосистемы, использование которых делает невозможным злоупотребления.

Для диалоговых коммуникаций, работающих в режиме реального времени, Бет и другие в 1994 г. предложили создать наблюдательное агентство, активно участвующее в протоколе, используемом отправителем и получателем для установления сессионного ключа, но так, чтобы два партнера не могли обнаружить участия агентства. Новшество этого подхода, однако, состоит в такой возможности, что в случае отсутствия перехвата информации провайдер сети может доказать этот факт.

11.4.4. Недоверие, которое некоторые граждане (или организации) испытывали к государственной власти, не уменьшилось в связи с недавними случаями, например, в США с действительным или подозреваемым посягательством АНБ на развитие алгоритмов шифрования, подобных DES. Отмечалось, что самовольно взятая NSA на себя ответственность за одобрение и рекомендацию криптографических алгоритмов «подобна назначению лисицы присматривать за курятником».

В 1957 г. появились сообщения о тесных контактах между Уильямом Фридманом и Борисом Хагелином, возбудившие некоторые подозрения.

Рис. 80. Борис Хагелин (1892–1983)

Наконец, есть третья группа, стоящая вне философии равновесия между конституционными правами на личное и государственным правом на обеспечение правопорядка. Эта группа не может быть проигнорирована, ввиду ее экономической важности — это коммерческий поставщик. В интересах этой группы было иметь хорошие отношения как с гражданами (потенциальными покупателями), так и с государством (надсмотрщиком, а иногда также клиентом). В лучшем случае, поставщик был честным посредником между двумя другими группами.

Однако эта роль затрудняется некоторой нечестностью государственных властей, воздействующих на коммерческих поставщиков созданием запретительных мер на их торговлю с иностранными партнерами, которые не сохраняются при их внутренней торговле с самим государством. Это плохо согласуется с правилами мировой свободной торговли.

11.4.5. Трудно бороться с ощущением, что криптология в начале третьего тысячелетия все еще остается в *Черном кабинете*. Государственные власти в этом вопросе совершенно непроницаемы и продолжают цепляться за последние остатки своего могущества. Однако существует прочный фундамент прав, от которых государственные власти не могут отказаться, так как должен соблюдаться баланс власти. Не только единственная остающаяся сверхдержава, США, но также и менее сильные государства Европы поймут необходимость

того, чтобы гражданские и коммерческие криптография и криптоанализ пришли к соглашению с государством. Соединенное Королевство с его давними традициями демократии, до сих пор непроницаемо секретное в вопросах криптоанализа, придерживается девиза, что тот, кто не защищает свою собственную безопасность, подвергает опасности безопасность своих друзей.

Но требования гражданского и коммерческого мира должны восприниматься серьезно. Политически было бы трудно принять то, что в США патентные приложения для криптосистем были заблокированы на основании Закона о секретных изобретениях 1940 г. или Закона о национальной безопасности 1947 г. Подобным же образом, щепетильность по отношению к защите частных и персональных данных является политическим фактором, и ее нельзя игнорировать. В США политика, основанная на внутренних ограничениях, до сих пор была неубедительной, как показал скандал вокруг FIDNET (Federal Intrusion Detection Network) и CESA (Cyberspace Electronic Security Act, закон о безопасности электронного кибернетического пространства, *англ.*), 1999 г.

11.4.6. В декабре 1998 г. в рамках *Вассенарского соглашения по контролю над экспортом обычных вооружений, товаров и технологий*, были достигнуты некоторые принципы довольно либерального экспортного контроля продуктов криптографии. В частности, экспорт 64-битных алгоритмов шифрования не требовал контроля для стран-членов Вассенарского соглашения.

Затем 16 сентября 1999 г. в США администрация Клинтона объявила о своем намерении дальнейшей либерализации, допускающей «экспорт и реэкспорт некоторых предметов шифрования или средств программного обеспечения для индивидуального пользования, коммерческих фирм и других неправительственных конечных пользователей и всех предназначений». Эта новая политика упростит американские правила экспорта продуктов шифрования. Она опирается на следующие три принципа: заблаговременный (до сбыта) технический обзор продуктов шифрования, упрощенная система послеэкспортного отчета, и процесс, позволяющий правительству пересмотр решения об экспорте сильных шифров иностранным правительствам. Эти послабления вызвали понимание со стороны американской промышленности. «Ограничения для государств, поддерживающих терроризм, их граждан и других запрещенных организаций этим правилом не изменяется». Это может утешить американский Департамент юстиции. Насколько важно это ограничение, было продемонстрировано террористической атакой на США 11 сентября 2001 г.

В целом, американское правительство полагает, что «эта новая политика продолжает служить всей области национальных интересов: поддержке обеспечению правопорядка и национальной безопасности, защите частной жизни и содействию электронной коммерции». 12 января 2000 г. Экспертное управление США (ВХА) опубликовало новые либеральные правила, регулирующие экспорт программных продуктов.

11.4.7. Остается надеяться, что эти новые правила станут эффективными. Надо надеяться, что победит здравый смысл. В настоящее время главная цель научной работы в гражданской и коммерческой криптографии состоит

в том, чтобы *найти нижние оценки сложности незаконного дешифрования для точно определенного типа компьютера, при реальных предположениях о недостатке дисциплины среди непрофессиональных пользователей.*

Это полезная задача, решение которой может дать пользователю криптосистемы гарантированную надежность. Это включает и необходимый для открытого источника код, поскольку каждая криптосистема с каким-либо неопубликованным алгоритмом может содержать неприятные сюрпризы.

Часть II

Криптоанализ

Il ne faut alors ni se buter, ni se rebuter, et faire comme en politique: changer son fusil d'épaule.

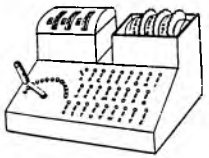
[Не следует пытаться пробить стену лбом или паниковать. Поступи, как политик: измени свою точку зрения, *фр.*]

Этьен Базерье, 1901 *г.*

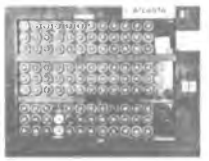
Дешифрование, по-моему, самое чарующее из искусств и, боюсь, что я потратил на него гораздо больше времени, чем оно того заслуживало.

Чарльз Бэббидж, 1864 *г.*

Введение: аппаратура



Циклометр
(*Польша*)



BOMBE
(*Великобритания*)

COLOSSUS
(*Великобритания*)

Процитированный выше Базерье с галльским шармом предостерегает от переоценки систематики криптоаналитических методов, которые мы обсуждаем в этой второй части.

Если исключить простой, но, как правило, трудноосуществимый метод исчерпания (перебора), все эти методы основаны на свойствах, присущих языку, которые трудно искоренить даже с помощью самого утонченного шифрования. При систематической работе находят и используют определенные инварианты криптологических методов. Как нахождение шаблонов, так и частотный анализ основаны на инвариантах одноалфавитных (односимвольных и многосимвольных) шифров. Но даже и многоалфавитный шифр оставляет инвариантным один лингвистический параметр, называемый *Kappa*. Это позволяет свести повторно используемый многоалфавитный шифр к одноалфавитным шифрам. Для перестановки, которая является полиграфическим шифром особого рода, использование этой особенности ведет к применению контактных частот.

Согласно Фридману, криптоанализ включает в себя определение используемого языка, определение общей криптосистемы, определение особого ключа и определение открытого текста; обычно именно в таком порядке.

Криптоанализ требует применения надлежащих средств в надлежащем месте. Живарж выразил это более радикально: *Certains rasoirs excellents sont pourtant tout à fait dangereux dans les mains d'un singe.* [Некоторые превосходные бритвы тем не менее опасны в руках обезьяны, *фр.*]

Следует заметить, что, как правило, активные криптоаналитические атаки против правительственных или коммерческих каналов связи наказуемы, тогда как атаки против частных каналов могут спровоцировать лишь процесс о возмещении ущерба. Однако, поскольку мы нуждаемся в знании криптоаналитической методики, чтобы быть способными размышлять о безопасном применении криптографических методов, в частности, об избежании иллюзорных сложностей, — мы надеемся, что нас извинят за профилактический подход к обсуждению криптоанализа.

Зачастую криптоанализ — это вопрос не только физического напряжения, но также наличного времени. Многие новости становятся бесполезными, как только они устаревают, а в некоторых областях они устаревают очень быстро. Бисли в книге *Very special intelligence*, London, 1977 г. заметил в этой связи: «Однако следовало бы подчеркнуть, что для практического использования криптоанализ должен быть быстрым». А Фридман заявил: «Лучшее, чего можно было бы ожидать, это чтобы степень безопасности была достаточно велика, чтобы отсрочить раскрытие информации противником на столь длительный срок, что к моменту раскрытия эта информация потеряла бы для него всю свою ценность».

Иногда решения криптограмм покоились на основаниях, ненамного превосходящих простую догадку. Требования к достаточной достоверности дешифрования могут колебаться в зависимости от ситуации — от рационального восстановления 90% открытого текста (Мейер—Матиаш) и до полного восстановления не только самого сообщения, но также ключа и всей криптосистемы (Рорбах).

Шифровальные схемы и машины существуют не долее периода их непосредственной полезности, а затем сразу выходят из употребления. Спланированное устаревание (разд. 2.1.1) — хороший прием, который минимизирует риск в случае кражи или захвата криптографического оборудования. Применяемая временная шкала очень сильно зависит от передаваемой информации: десять часов для управления артиллерийским огнем соответствует десяти годам в дипломатической переписке. Но даже устаревшие книги кодов или шифровальные таблицы содержат информацию, разглашение которой нежелательно. Практический криптоанализ извлекает пользу из множества крохотных деталей об особенностях противника, его обычаях, его предпочтениях — и тщательно их регистрирует. В этом смысле криптоанализ кормится уже полученными им результатами.

Криптоанализ в значительной степени основан на ошибках шифрования, в том числе и на мелких ошибках противника. Сакко саркастически заметил в этой связи: *Les chiffreurs se chargent suffisamment d'aider l'ennemi.* [Шифровальщики в значительной мере занимаются помощью врагу, *фр.*]

Асимметрические криптосистемы допускают публикацию шифровальных ключей (открытый ключ) и указывают на либерализацию криптографии (публичная криптография), которая в свою очередь вызывает рост общественной осведомленности о криптоанализе вообще — несмотря на то, что криптологические службы правительственных агентств смотрят на эту сторону публич-

ной криптографии довольно неодобрительно. Эти агентства не рассматривают себя в качестве инструментов для образования масс. В качестве примера процитируем высказывание Циммермана: «...авторы многих программ шифрования говорят, что они никогда не слыхали о методе CBC или CFB. Сам факт, что они даже не изучили толком криптографию, чтобы знать эти элементарные понятия, является обескураживающим». Криптоанализ не собирается умирать!

К тому же криптоанализ уже вышел на свободный рынок. Компания Access Data Recovery (87 East 600 South, Orem, Utah 84058, USA) — и это не единственный пример — продает за какие-то сто долларов программу, разработанную Томпсоном, который взломал шифры, используемые в Word Perfect, Lotus 1–2–3, MS Excel, Symphony, Quattro Pro, Paradox и MS Word — и сделал это не методом перебора, а при помощи настоящих криптоаналитических методов. Люди покупают эту программу для восстановления забытых паролей, а полицейские офицеры используют ее при чтении конфискованных данных.

Глава 12

Комбинаторная сложность перебора

Gewöhnlich glaubt der Mensch, wenn er nur Worte hört, es müsse sich dabei doch auch Was denken lassen.

[Обычно человек, как только он слышит слово, полагает, что при этом надо еще также, чтобы было позволено думать, *нем.*]

Гёте

Критерием комбинаторной сложности шифра является мощность класса методов, соответствующая числу имеющихся ключей. Как мера надежности против дешифрования, она дает верхнюю границу работы, требуемой для исчерпывающего поиска, в предположении, что класс методов известен (Принцип Шеннона: «Противнику известна используемая система»).

Мы будем часто применять формулу Стирлинга[1]:

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\left(1 + \frac{1}{12n - \frac{1}{2}} + O\left(\frac{1}{n^3}\right)\right) = \sqrt{2\pi e}^{\,n+1/2}\left(1 + \frac{1}{12n - \frac{1}{2}} + O\left(\frac{1}{n^3}\right)\right)$$

с числовыми значениями

$$\begin{aligned} \sqrt{2\pi} &= 2.506628275\ldots, \\ e &= 2.718281828\ldots, \\ \sqrt{2\pi e} &= 4.132731353\ldots \end{aligned}$$

Также будем использовать ее логарифмический вариант[2] (с основанием логарифма, равным 2):

$$\operatorname{ld} n! = \left(1 + \frac{1}{2}\right)(\operatorname{ld} n - \operatorname{ld} e) + \frac{1}{2}(\operatorname{ld} \pi + \operatorname{ld} e + 1) + \operatorname{ld} e\left(\frac{1}{12n} - \frac{1}{360n^3} + O\left(\frac{1}{n^5}\right)\right)$$

[1] $26! = 403\,291\,461\,126\,605\,635\,584\,000\,000 = 2^{23} \cdot 3^{10} \cdot 5^6 \cdot 7^3 \cdot 11^2 \cdot 13^2 \cdot 17 \cdot 19 \cdot 23$.

[2] Здесь $\operatorname{ld} x$ обозначает логарифм по основанию 2: $\operatorname{ld} x = \ln x / \ln 2 = \log_{10} x / \log_{10} 2$.

с числовыми значениями

$$\operatorname{ld} e = 1.442695041\ldots,$$
$$\frac{1}{2}(\operatorname{ld}\pi + \operatorname{ld} e + 1) = 2.047095586\ldots$$

Будем обозначать через N мощность $|V|$ алфавита V, и через $Z = |S|$ мощность класса методов S.

Ниже мы получим значения для комбинаторной сложности Z некоторых классов методов S. При этом, информация $\operatorname{ld} Z$ о классе методов S измеряется в [бит] (битах), а $\lg Z$ — в [бан] (банах) — единицах, введенных Тьюрингом, с практической единицей 1 [децибан] $= 0.1/\log_{10} 2 \approx 0.332$ [бит].

12.1. Одноалфавитные простые шифры

Простые подстановки односимвольны. Оставим пока вне рассмотрения омофоны и пустышки и будем рассматривать только перестановки.

12.1.1. Простые подстановки в целом. (Частный случай $n = 1$ см. в разд. 12.2.1.)

12.1.1.1. (Простые подстановки см. в разд. 3.2.) Множество перестановок $V \overset{N}{\leftrightarrow} V$ имеет ту же мощность, что и множество взаимнооднозначных отображений (без омофонов) из V в $W^{(m)}$, не зависящих от W и m, так что

$$Z = N! \asymp \sqrt{2\pi e}\left(\frac{N}{e}\right)^{N+1/2} \asymp 4.13\cdot\left(\frac{N}{e}\right)^{N+1/2},$$
$$\operatorname{ld} Z \asymp \left(N+\frac{1}{2}\right)\cdot(\operatorname{ld} N - 1.44) + 2.05.$$

Для $N = 26$ имеем: $Z \approx 4.03\cdot 10^{26}$, $\operatorname{ld} Z \approx 88.382$ [бит], $\log Z \approx 266.06$ [децибан].

12.1.1.2. (Одноцикловые простые подстановки, разд. 3.2.3.) Рассматриваются перестановки $V \overset{N}{\leftrightarrow} V$, состоящие из одного цикла порядка N. Тогда

$$Z = (N-1)! \asymp \sqrt{2\pi e}\left(\frac{N-1}{e}\right)^{N-1/2} \asymp 4.13\cdot\left(\frac{N-1}{e}\right)^{N-1/2},$$
$$\operatorname{ld} Z \asymp \left(N-\frac{1}{2}\right)\cdot\big(\operatorname{ld}(N-1) - 1.44\big) + 2.05.$$

Для $N = 26$ имеем: $Z \approx 1.55\cdot 10^{25}$, $\operatorname{ld} Z \approx 83.682$ [бит], $\log Z \approx 251.91$ [децибан].

12.1.1.3. (Собственные взаимообратные простые подстановки, разд. 3.2.1.) Для собственно взаимообратной перестановки $V \overset{2}{\leftrightarrow} V$ имеем (N — четное):

$$Z = (N-1)!! \overset{\text{def}}{=} (N-1)(N-3)(N-5)\ldots\cdot 5\cdot 3\cdot 1 \asymp \sqrt{2}\left(\frac{N}{e}\right)^{N/2},$$
$$\operatorname{ld} Z \asymp \frac{N}{2}\cdot(\operatorname{ld} N - 1.44) + \frac{1}{2}.$$

Для $N = 26$ имеем: $Z \approx 7.91\cdot 10^{12}$, $\operatorname{ld} Z \approx 42.846$ [бит], $\log Z \approx 128.98$ [децибан].

12.1.2. Числовые алфавиты.[3] (Частный случай см. в разд. 12.2.2.) Предположим, что алфавит линейно циклически квазиупорядочен (разд. 5.6). Тогда $Z = \varphi(N)$, где φ — функция Эйлера (разд. 5.6),

$$\mathrm{ld} Z = \mathrm{ld}\ N + \sum_{\mu=1}^{k} \mathrm{ld}\ \rho(p_\mu, 1)$$

(см. разд. 12.2.2). Для $N = 26$ имеем: $Z = 12$, $\mathrm{ld}\ Z \approx 3.58$ [бит], $\log Z \approx 10.79$ [децибан] (см. разд. 5.5, табл. 1 b).

	Z	$\mathrm{ld}\ Z$
12.2.1 Подстановка в целом	$(26^n)!$	$(26^n + 1/2)(4.70n - 1.44) + 2.05$
12.2.2 Преобразование ХИЛЛА	$0.265 \cdot 26^{n^2}$	$4.70n^2 - 1.916$
12.2.3 Сложение ЦЕЗАРЯ	26^n	$4.70n$
12.2.4 Перестановка	$n!$	$(n + 1/2)(\mathrm{ld} n - 1.44) + 2.05$

$\mathrm{ld}\ Z$				
$n = 1$	$n = 4$	$n = 16$	$n = 64$	$n = 256$
$8.84 \cdot 10^1$	$7.93 \cdot 10^6$	$3.22 \cdot 10^{24}$	$1.08 \cdot 10^{93}$	$2.07 \cdot 10^{365}$
3.58	73.29	$1.20 \cdot 10^3$	$1.93 \cdot 10^4$	$3.08 \cdot 10^5$
4.70	18.80	75.21	300.83	1 203.31
	4.58	44.25	296.00	1 684.00

Таблица 3. Сложность одноалфавитных (многосимвольных) схем шифрования для $N = 26$ в зависимости от ширины n шифрования

12.1.3. Сложение ЦЕЗАРЯ. (Частный случай $n = 1$, разд. 12.2.3.) Сложение ЦЕЗАРЯ $V \overset{+}{\leftrightarrow} V$, или сдвиг, является одноалфавитным частным случаем подстановки ВИЖЕНЕРА (разд. 7.4.1). Имеем:

$$Z = N,$$
$$\mathrm{ld}\ Z = \mathrm{ld}\ N.$$

Для $N = 26$ имеем: $Z = 26$, $\mathrm{ld}\ Z \approx 4.70$ [бит], $\log Z \approx 14.15$ [децибан].

12.2. Одноалфавитные многосимвольные шифры

Комбинаторная сложность многосимвольных подстановок зависит от ширины n шифрования открытого текста.

[3] Decimated alphabets.

12.2.1. Многосимвольные подстановки в целом. Множество перестановок $V^n \leftrightarrow V^n$ имеет ту же мощность, что и множество взаимно однозначных отображений V^n в W^m, независимых от W и m. Поэтому

$$Z = (N^n)!,$$
$$\text{ld } Z \asymp \left(N^n + \frac{1}{2}\right)(n \cdot \text{ld } N - 1.44) + 2.05.$$

Для $N = 26$ имеем: $Z = (26^n)!$, $\text{ld } Z \approx (26^n + 1/2)(4.70n - 1.44) + 2.05$.

Для диграфических подстановок:

$$Z \approx 1.88 \cdot 10^{1621}, \quad \text{ld } Z \approx 5387 \text{ [бит]};$$

для триграфических подстановок:

$$Z \approx 1.19 \cdot 10^{66978}, \quad \text{ld } Z \approx 222500 \text{ [бит]};$$

для тетраграфических подстановок:

$$Z \approx 4.82 \cdot 10^{2388104}, \quad \text{ld } Z \approx 7933000 \text{ [бит]}.$$

Множество подстановок ПЛЕЙФЕЙРА имеет ту же мощность, что и множество одноцикловых простых подстановок порядка $N = 25$. Поэтому

$$Z = 25!/(5 \cdot 5) \approx 6.20 \cdot 10^{23}, \quad \text{ld } Z \approx 79.038 \text{ [бит]}, \quad \log Z \approx 237.93 \text{ [децибан]}.$$

12.2.2. Многосимвольные однородные линейные подстановки (Преобразование ХИЛЛА). Предполагается, что алфавит линейно циклически квазиупорядочен (разд. 5.2.3). В таком случае

$$Z = N^{n^2} \cdot \rho(N, n), \text{ где при } N = p_1^{s_1} \cdot p_2^{s_2} \ldots p_k^{s_k} \text{ получаем}$$
$$\rho(N, n) = \rho(p_1, n) \cdot \rho(p_2, n) \ldots \rho(p_k, n),$$
$$\text{ld } Z = n^2 \text{ ld } N + \sum_{\mu=1}^{k} \text{ld}\rho(p_\mu, n).$$

Для большого n приближенное значение $\rho(p, n)$ приводится в разд. 5.2.3; для большого n и не слишком малого p, взяв $\text{ld } e \approx 1.44$, получим

$$\text{ld } \rho(p, n) \approx 1.44/(3/2 - p).$$

Для $N = 26$ и большого n получаем $\rho(2, n) \approx 0.289$ и $\rho(13, n) \approx 0.917$, так что $\rho(26, n) \approx 0.289 \cdot 0.917 = 0.265$, поэтому

$$Z \approx 0.265 \cdot 26^{n^2}, \text{ ld } Z \approx 4.70 \cdot n^2 - 1.92 \text{ [бит]}, \log Z \approx 14.15n^2 - 5.78 \text{ [децибан]}.$$

12.2.3. Многосимвольный сдвиг (многосимвольное сложение ЦЕЗАРЯ). Многосимвольное сложение ЦЕЗАРЯ $V^n \leftrightarrow V^n$ с шириной шифрования n, или сдвиг, является частным случаем линейной подстановки (см. начало гл. 3), когда T является единичной матрицей (порядка n). Поэтому

$$Z = N^n, \quad \text{ld } Z = n \text{ ld } N.$$

Для $N = 26$ имеем: $Z = 26^n$, ld $Z \approx 4.70 \cdot n$ [бит], lg $Z \approx 14.15 \cdot n$ [децибан].

12.2.4. Перестановки. Перестановки ширины n (см. начало гл. 6) являются линейными подстановками, матрицы которых (порядка n) перестановочны. Поэтому сложность Z не зависит от N и

$$Z = n!$$

$$\text{ld } Z \asymp \left(n + \frac{1}{2}\right)(\text{ld } n - 1.44) + 2.05.$$

12.2.5. Краткий обзор по одноалфавитным подстановкам. Комбинаторные сложности одноалфавитных подстановок приведены в табл. 3 и изображена графически на рис. 80.

Заметим, что сложность перестановки превосходит сложность многосимвольного сложения ЦЕЗАРЯ примерно при $n = N \cdot e$ (для $N = 26$ это происходит при $n = 68$ с ld $Z \approx 320$).

Для $N = 26$ перестановка при $n = 26$ достигает сложности простой (монограммной) подстановки; однородная линейная подстановка с ld $Z \approx 4.70n^2 - 1.92$ превышает по сложности при $n = 5$ простую (односимвольную) подстановку с ld $Z \approx 88.38$, а при $n = 34$ — диграфическую подстановку с ld $Z \approx 5.387 \cdot 10^3$. Многосимвольное сложение ЦЕЗАРЯ при $n = 2$ уступает по сложности однородной линейной подстановке.

	Z	ld Z
12.3.1 ПЕРЕСТАНОВКИ	$(26!)^d$	$88.38 \cdot d$
12.3.2 МУЛЬПЛЕКС	$(25!)^d$	$83.68 \cdot d$
12.3.3 АЛЬБЕРТИ	$26! \cdot 26^{d-1}$	$4.70 \cdot d + 83.68$
12.3.4 ВИЖЕНЕР	26^d	$4.70 \cdot d$

ld Z				
$d = 1$	$d = 10$	$d = 100$	$d = 1000$	$d = 10000$
88.38	883.82	8 838.20	88381.95	883 819.53
83.68	836.82	8368.15	83681.51	836815.36
88.38	130.69	553.73	4784.12	47088.08
4.70	47.00	470.04	4700.44	47004.40

Таблица 4. Сложность многоалфавитных (односимвольных) криптосистем для $N = 26$ в зависимости от числа d используемых алфавитов

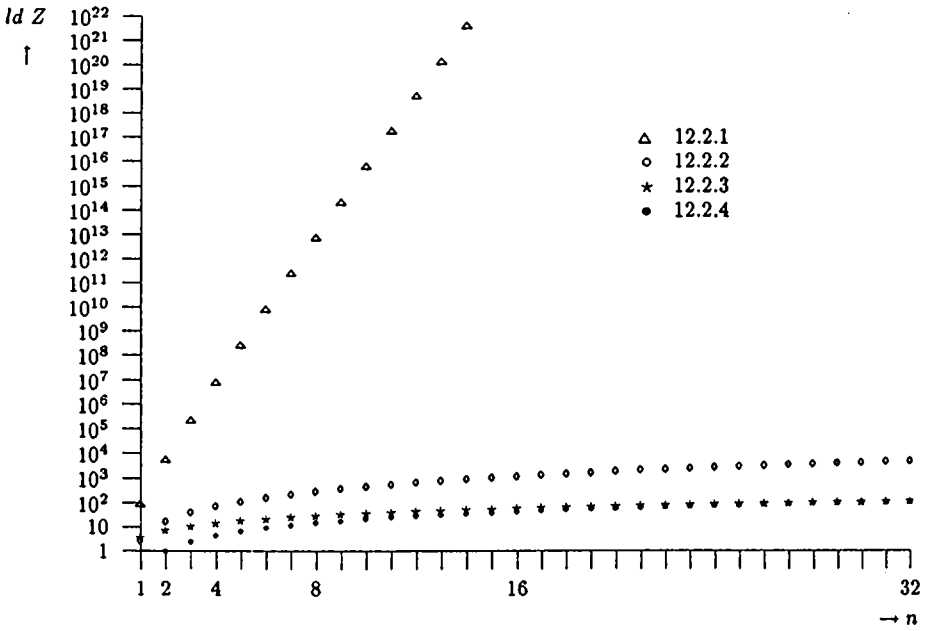


Рис. 81. Комбинаторная сложность одноалфавитных многосимвольных подстановок ширины n для $N = 26$

Заметим, что блочная перестановка ширины n полиграфична, но также одноалфавитна — и немного странно назвать n «периодом».

12.3. Многоалфавитные шифры

Комбинаторная сложность наиболее общих многоалфавитных (односимвольных) шифров с d различными алфавитами является произведением сложностей для разных алфавитов. Для случая родственных алфавитов сложность соответственно меньше.

12.3.1. Перестановочный шифр с d неродственными алфавитами. В этом случае имеем

$$Z = (N!)^d,$$

$$\operatorname{ld} Z = d \cdot \left(\left(N + \frac{1}{2}\right)(\operatorname{ld} N - 1.44) + 2.05\right).$$

Для $N = 26$ имеем: $Z \approx (4.03 \cdot 10^{26})^d$, $\operatorname{ld} Z = 88.382 \cdot d$ [бит].

12.3.2. Шифр МУЛЬТИПЛЕКС с d алфавитами. Имеем:

$$Z = ((N-1)!)^d,$$

$$\operatorname{ld} Z \asymp d \cdot \left(\left(N - \frac{1}{2}\right)(\operatorname{ld} N - 1.44) + 2.05\right).$$

Для $N = 26$ имеем: $Z \approx (1.55 \cdot 10^{25})^d$, $\operatorname{ld} Z = 83.682 \cdot d$ [бит].

12.3.3. Шифр АЛЬБЕРТИ с d алфавитами. Имеем:

$$Z = N!N^{d-1},$$

$$\mathrm{ld}\ Z \asymp d \cdot \mathrm{ld}\ N + \left(N - \frac{1}{2}\right)(\mathrm{ld}(N-1) - 1.44) + 2.05.$$

Для $N = 26$ имеем: $Z \approx 1.55 \cdot 10^{25} \cdot 26^d$, $\mathrm{ld}\ Z = 4.70 \cdot d + 83.682$ [бит].

12.3.4. Шифры ВИЖЕНЕРА или БОФОРТА с d алфавитами. Имеем:

$$Z = N^d,$$

$$\mathrm{ld}\ Z = d \cdot \mathrm{ld}\ N.$$

Для $N = 26$ имеем: $Z = 26^d$, $\mathrm{ld}\ Z \approx 4.70 \cdot d$ [бит].

12.3.5. Краткий обзор по многоалфавитным шифрам. Комбинаторные сложности многоалфавитных подстановок для односимвольного случая $n = 1$ приведены в табл. 4 и изображены графически на рис. 81.

Заметим, что шифр ВИЖЕНЕРА или БОФОРТА и одноалфавитная простая подстановка имеют одну и ту же сложность при

$$d \approx N + \frac{1}{2} - \frac{N - \ln 2\pi}{\ln N}$$

(т. е. для $N = 26$ при $d = 19$, а для $N = 26^2$ — при $d = 573$).

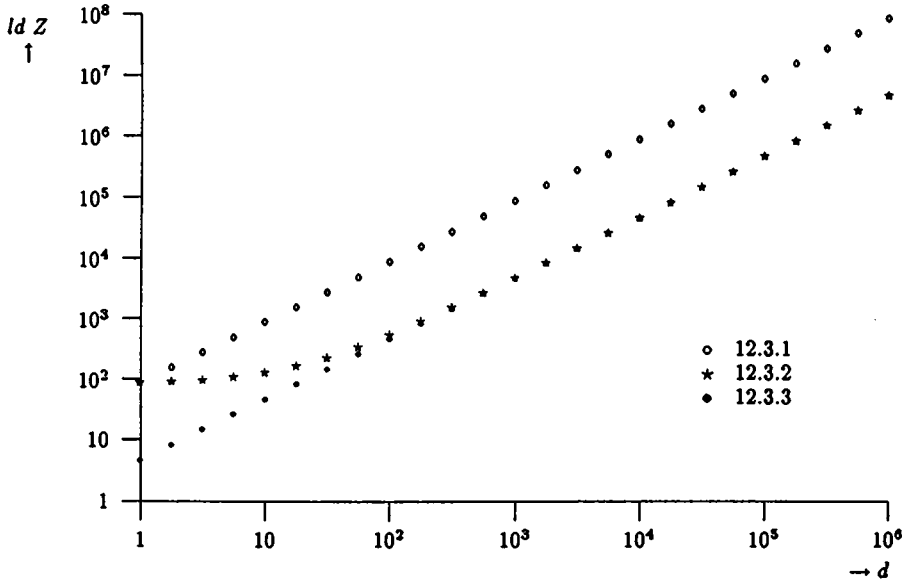


Рис. 82. Комбинаторная сложность многоалфавитного шифра с d различными алфавитами

МУЛЬТИПЛЕКСНЫЙ шифр с d алфавитами и (одноалфавитная) многосимвольная подстановка ширины n имеют примерно одну сложность при $d \approx n \cdot N^{n-1}$ (более точно, в биграммном случае, т. е. для $n=2$ и для $N=26$ — при $d=55$).

Комбинаторные сложности шифра ВИЖЕНЕРА или БОФОРТА с периодом $d=h$ и полиграфического сложения ЦЕЗАРЯ с шириной $n=h$ совпадают. Для $n=10$ можно использовать сумматор с h позициями, механизм переноса которого должен быть демонтирован в первом случае, а во втором — нет (ср. разд. 5.7.1. и разд. 8.3.3).

12.4. Общие замечания о комбинаторной сложности

Для теоретического изучения комбинаторной сложности существует целый класс методов. Однако незаконный дешифровальщик часто сталкивается с ограничениями, причиной которых являются привычки или глупость шифровальщиков, а также материальные условия.

12.4.1. Например, сложность цилиндра Джефферсона и Базерье (без информации о дисках) равна сложности шифра МУЛЬТИПЛЕКС[4]. Если информация о дисках доступна, то его сложность равна сложности перестановки. Для $d=25$ (М-94) это означает сокращение от $Z=(26!)^{25} \approx 1.38 \cdot 10^{665}$ до $Z=25! \approx 1.55 \cdot 10^{25}$ или от $\mathrm{ld}\ Z \approx 2209$ до $\mathrm{ld}\ Z \approx 83.68$.

Аналогичная ситуация имеет место с диском Альберти: если информация о нем попадает в руки врага, то шифр АЛЬБЕРТИ «проваливается» до шифра ВИЖЕНЕРА; соответственно Z уменьшается от $(N!) \cdot N^{d-1}$ до N^d, или для $N=26$ значение $\mathrm{ld}\ Z$ уменьшается от $4.70 \cdot d + 83.68$ до $4.70 \cdot d$.

12.4.2. Заметим также, что сложность двойной перестановки ширины n, равная $Z=(n!)^2$, меньше, чем сложность перестановки с удвоенной шириной $Z=(2n)!=(n!)^2 \cdot \binom{2n}{n}$ в

$$\binom{2n}{n} \asymp \frac{4^n}{\sqrt{\pi \cdot \left(n + 1/4 + 1/(32 \cdot n)\right)}}$$

раз. Но это означает только, что перебор, задаваемый фиксированным временным пределом, проходит быстрее для двойной перестановки, что не противоречит тому эмпирическому факту, что в области сложности, где перебор не поддается обработке, криптоанализ двойной перестановки гораздо сложнее, чем криптоанализ колонной перестановки с удвоенной шириной.

12.4.3. Наконец, замечательно, что для шифра ВИЖЕНЕРА сложность Z и $\mathrm{ld}\ Z$ зависят только от N^d, и таким образом, (для $N=2^k$) они инвариантны относительно перехода к бинарному кодированию: $(2^k)^d = 2^{(k \cdot d)}$. В противоположность этому, переход к бинарному кодированию резко уменьшает сложность шифра перестановки: $((2^k)!)^d > (2!)^{k \cdot d}$ для $k \geqslant 2$.

[4] «Криптоустройство может попасть в руки врага», — гласит Принцип №3 (разд. 11.2.3). Базерье изобрел свое устройство в 1891 г., через 8 лет после того, как Керкхоффс опубликовал свой совет.

12.5. Криптоанализ путем перебора

Должно быть ясно, что комбинаторная сложность есть мера надежности шифра только в том смысле, что это мера усилия, требуемого для особого рода «незаконного» дешифрования, очень простого и очень общего, которое мы будем называть атакой перебора.

После того, как мы угадали класс методов шифрования, мы строим все открытые тексты (все «варианты»), которые в процессе шифрования одного из этих методов приводят к данному криптотексту, и тогда читаем «правильное» сообщение или «собираем» его в правильном значении слова. Эта атака может привести к более чем одному собранному сообщению, и это показывает, что дешифрование неоднозначно, точнее — что шифрование не инъективно для того метода шифрования, который мы угадали. Это может означать, что следует найти более узкий класс методов шифрования. Мы вернемся к этому вопросу в разд. 12.6 «Расстояние единственности».

Если же не удается собрать ни одного сообщения, то это значит, что угаданный класс методов шифрования был ошибочным — или же была допущена ошибка в процессе шифрования заданного криптотекста.

H V Z D U V F K R Q G X Q N H O D O V L F K L Q E R Q Q D Q N D P L F K C
I W A E V W G L S R H Y R O I P E P W M G L M R F S R R E R O E Q M G L D
J X B F W X H M T S I Z S P J Q F Q X N H M N S G T S S F S P F R N H M E
K Y C G X Y I N U T J A T Q K R G R Y O I N O T H U T T G T Q G S O I N F
L Z D H Y Z J O V U K B U R L S H S Z P J O P U I V U U H U R H T P J O G
M A E I Z A K P W V L C V S M T I T A Q K P Q V J W V V I V S I U Q K P H
N B F J A B L Q X W M D W T N U J U B R L Q R W K X W W J W T J V R L Q I
O C G K B C M R Y X N E X U O V K V C S M R S X L Y X X K X U K W S M R J
P D H L C D N S Z Y O F Y V P W L W D T N S T Y M Z Y Y L Y V L X T N S K
Q E I M D E O T A Z P G Z W Q X M X E U O T U Z N A Z Z M Z W M Y U O T L
R F J N E F P U B A Q H A X R Y N Y F V P U V A O B A A N A X N Z V P U M
S G K O F G Q V C B R I B Y S Z O Z G W Q V W B P C B B O B Y O A W Q V N
T H L P G H R W D C S J C Z T A P A H X R W X C Q D C C P C Z P B X R W O
U I M Q H I S X E D T K D A U B Q B I Y S X Y D R E D D Q D A Q C Y S X P
V J N R I J T Y F E U L E B V C R C J Z T Y Z E S F E E R E B R D Z T Y Q
W K O S J K U Z G F V M F C W D S D K A U Z A F T G F F S F C S E A U Z R
X L P T K L V A H G W N G D X E T E L B V A B G U H G G T G D T F B V A S
Y M Q U L M W B I H X O H E Y F U F M C W B C H V I H H U H E U G C W B T
Z N R V M N X C J I Y P I F Z G V G N D X C D I W J I I V I F V H D X C U
A O S W N O Y D K J Z Q J G A H W H O E Y D E J X K J J W J G W I E Y D V
B P T X O P Z E L K A R K H B I X I P F Z E F K Y L K K X K H X J F Z E W
C Q U Y P Q A F M L B S L I C J Y J Q G A F G L Z M L L Y L I Y K G A F X
D R V Z Q R B G N M C T M J D K Z K R H B G H M A N M M Z M J Z L H B G Y
E S W A R S C H O N D U N K E L A L S I C H I N B O N N A N K A M I C H Z
F T X B S T D I P O E V O L F M B M T J D I J O C P O O B O L B N J D I A
G U Y C T U E J Q P F W P M G N C N U K E J K P D Q P P C P M C O K E J B

Таблица 5. Двадцать шесть вариантов шифра ЦЕЗАРЯ: HVZDU VFKRQ ...

Работа по перебору возможна, разумеется, лишь в случае, когда число вариантов, подлежащих тщательному изучению, не является слишком большим. Однако вовсе не обязательно для каждого тщательно проверяемого варианта собирать полный приводимый в доказательство открытый текст; отказ от дальнейшего собирания становится возможным, как только собранная часть текста оказывается абсурдной.

Мы проиллюстрируем метод перебора двумя небольшими примерами, где число построенных для тщательного изучения открытых текстов имеет порядок двух дюжин:

a) сложение ЦЕЗАРЯ с алфавитом $\mathbb{Z}_{26}$: 26 вариантов (табл. 5);

b) перестановка с шириной 4: 24 варианта (табл. 6).

S A E W S H R C N U O D K L N E L I A S H N C I O N B N N A A K I H M C W
A S E W H S R C U N O D L K N E I L A S N H C I N O B N A N A K H I M C N
A E S W H R S C U O N D L N K E I A L S N C H I N B O N A A N K H M I C N
E A S W R H S C O U N D N L K E A I L S C N H I B N O N A A N K M H I C Z
S E A W S R H C N O U D K L N E L A I S H C N I O B N N N A A K I M H C W
E S A W R S H C O N U D N K L E A L I S C H N I B O N N A N A K M I H C Z
S W E A S C R H N D O U K E N L L S A I H I C N O N B N N K A A I C M H W
W S E A C S R H D N O U E K N L S L A I I H C N N O B N K N A A C I M H A
W E S A C R S H D O N U E N K L S A L I I C H N N B O N K A N A C M I H A
E W S A R C S H O D N U N E K L A S L I C I H N B N O N A K N A M C I H Z
E S W A R S C H O N D U N K E L A L S I C H I N B O N N A N K A M I C H Z
S W A E S C H R N D U O K E L N L S I A H I N C O N N B N K A A I C H N W
W S A E C S H R D N U O E K L N S L I A I H N C N O N B K N A A C I H M A
W A S E C H S R U U N D E L K N S I L A I N H C N N D B K A N A C H I M A
A W S E H C S R U D N D L F K N I S L A N I H C N N O B A K M A H C I M N
S A W E S H C R N U D O K L E N L I S A H N I C O N N B N A K A I H C M W
A S W E H S C R U N D D L K E N I L S A N H I C N O N B A N K A H I C M N
A W E S H C R S U D O N L E N K I S A L N I C H N N B O A K A N H C M I N
W A E S C H R S D U O N E L N K S I A L T N C H N N B O K A A N C H M I A
W E A S C R H S D O U N E N L K S A I L I C N H N B N O K A A N C M H I A
E W A S R C H S D D U N N E L K A S I L O I N H A N N D A K A N M C H I Z
A E W S H R C S U D O N L N E K I A S L N C I H N B N O A A K N H M C I N
E A W S R H C S O U D N N L E K A I S L O N I H B N N O A A K N M H C I Z
S E W A S R C H N D D U K N E L L A S I H C I N O B N N N A K A I M C H W

Таблица 6. Двадцать четыре варианта перестановки ширины 4: SAEWS HRCNU ...

Метод перебора годится также в том случае, если число ключевых «вероятных слов», которые заданы или угаданы, не слишком велико. В эпоху Возрождения репертуар известных цитат не был чрезмерно велик, встречаются, например, в криптографической литературе

OMNIA VINCIT AMOR [любовь побеждает все (*лат.*, Вергилий)],
VIRTUTI OMNIA PARENT,
SIC ERGO ELEMENTIS,
IN PRINCIPIO ERAT VERBUM.

На самом деле, даже сегодня находятся приверженцы таких программных ключевых слов — от любителей до государственных мужей.

12.6. Расстояние единственности

Продолжение постепенного, буква за буквой, наращивания фрагментов возможного открытого текста приводит к наблюдению, что после некоторой (довольно отчетливо определяемой) длины можно с уверенностью вынести решение о единственности открытого текста. Эта длина называется эмпирическим расстоянием единственности U для рассматриваемого класса методов. Замечательно, что в двух примерах, представленных табл. 5 и 6 с почти равной сложностью ($Z \approx 25$ и $\mathrm{ld}Z \approx 4.64$) расстояния единственности примерно равны четырем символам. Существует очень мало четырехбуквенных слов, допускающих двусмысленное шифрование ЦЕЗАРЯ для английского языка (алфавит Z_{26}): mpqy: ADEN, KNOW; aliip: DOLLS, WHEEL; afccq: JOLLY, CHEER; для немецкого языка (Z_{26}): zydd: BAFF, POTT; qfzg: LAUB, TICK; qunq: EIBE, OSLO; himy: ABER, NORD, KLOA(KE), (ST)OPSE(L) (Z_{25}!). Здесь существенны лишь слова с разными буквами.

Расстояние единственности может быть оценено опытными криптоаналитиками и для шифров с гораздо большей сложностью Z, например, для одноалфавитной простой подстановки (простой замены) ($Z = 26!$, $\mathrm{ld}\ Z = 88.38$): для очень коротких криптотекстов существует двусмысленность, для очень длинных криптотекстов существует единственное решение. В случае одноалфавитной простой подстановки сообщалось, что эмпирическое расстояние единственности находится между 25 и 30: «...точка единственности равна примерно 27 буквам... При 30 буквах всегда существует единственное решение для криптограммы этого типа, а при 20 обычно легко найти несколько решений» (Шеннон, 1945 г.); «Практически каждый пример с 25 или более символами, представляющий одноалфавитную шифрограмму «осмысленного сообщения» на английском языке может быть без труда решен» (Фридман, 1973 г.). Экспериментальные проверки шифровальными методами очень большой сложности подтверждают следующий эмпирический закон:

Расстояние единственности зависит (при одном и том же естественном языке) только от комбинаторной сложности Z класса рассматриваемых методов. Более того, оно (для не слишком малых Z) пропорционально $ld\ Z$.

Этот количественный результат в 1935 г. еще не был опубликован. После Касиски (1863 г., разд. 17.4) стали появляться лишь интуитивные количественные соображения, вроде «длины ключей должны быть сравнимы с длинами самих сообщений» (Хитт, 1914 г., разд. 8.8.2), хотя, вероятно, Фридман уже имел смутные подозрения по этому поводу. Рассматриваемый закон означает, что все влияние избыточности языка, содержащейся в тексте, можно выразить всего лишь коэффициентом пропорциональности. Эта идея была отправной точкой для создания Клодом Шенноном теории информации, основы

которой были заложены в написанном им в 1945 г. секретном докладе. Этот доклад был опубликован в открытой печати в 1949 г.

Предполагая, что фридмановское значение расстояния единственности U для одноалфавитной простой подстановки равно 25 и $\mathrm{ld}\ Z \approx 88.38$, получаем эмпирическую формулу пропорциональности:

$$U \approx \frac{1}{3.54}\ \mathrm{ld}\ Z \approx \frac{1}{1.06}\ \log_{10} Z. \qquad (*)$$

Таблица 7, рассчитанная в соответствии с разд. 12.2 для разных значений ширины n, дает расстояние единственности для одноалфавитной многосимвольной подстановки.

	$n=1$	$n=4$	$n=16$	$n=64$	$n=256$
Подстановка в целом	**25**	2240000	10^{24}	10^{93}	10^{365}
Однородная линейная подстановка	(1.02)	22	340	5500	88000
Сложение Цезаря	(1.34)	6	22	86	340
Перестановка		(1.30)	13	85	480

Таблица 7. Эмпирическое расстояние единственности U, соответствующее формуле (*) и округленное для $N=26$

Величины в круглых скобках табл. 7 слишком малы, чтобы иметь серьезное значение[5)].

В частности, получаем три результата:

для диграфической подстановки $U \approx 1530$,
для триграфической подстановки $U \approx 63000$,
для тетраграфической подстановки $U \approx 2240000$.

Для получения расстояния единственности многоалфавитного шифра надо эмпирическое расстояние единственности основного одноалфавитного шифра умножить на длину d периода. Таким образом, для шифра ВИЖЕНЕРА, основанного на сложении ЦЕЗАРЯ, получаем

при $d=10^2$ $\quad U \approx 134$,
при $d=10^4$ $\quad U \approx 13\ 400$,
при $d=10^6$ $\quad U \approx 1\ 340\ 000$.

Если для метода шифрования эмпирическое расстояние единственности существует, то можно ожидать, что при подходящих атаках (отличных от

[5)] Это правило имеет следующую подоплеку в теории информации:
Значение $4.7 = \mathrm{ld}\ 26$ [бит/символ] можно разложить на сумму 3.5 [бит/символ] — долю избыточности — 74.5% и 1.2 [бит/символ] — долю информации (для Z_{26} и для английского языка) — 24.5%.

перебора) взлом шифра облегчается и становится более вероятным с увеличением длины криптотекста, поскольку когда эта длина превзойдет расстояние единственности, решение достигается без проблем, если приложить достаточные усилия. Для недешифруемых (невзламываемых) шифров (разд. 8.8.4) никакого расстояния единственности не существует.

12.7. Практическое выполнение перебора

Практическая работа по перебору проводится поэтапно, на каждом этапе увеличивается длина исследуемых фрагментов текстов, выбрасываются «невозможные» варианты и оставляются «возможно правильные».

Таблицы биграмм и триграмм, имеющиеся в литературе, показывают, что в английском, французском и немецком языках среди 676 биграмм около половины «возможных», но среди 17 576 триграмм «возможных» всего лишь около тысячи.

Такую работу можно без труда провести при помощи компьютера, если исходное число вариантов ненамного больше десяти тысяч. В таком случае на экране монитора группы от пяти до восьми символов легко выбираются взглядом, и по крайней мере 100 таких выборов можно сделать за одну ми-

```
V VF  VFK?
W WG  WGL   WGLS?
X XH  XHM?
Y YI  YIN?
Z ZJ  ZJO   ZJOV   ZJOVU?
A AK  AKP?
B BL  BLQ?
C CM  CMR?
D DN  DNS?
E EO  EOT   EOTA   EOTAZ?
F FP  FPU   FPUB   FPUBA?
G GQ  GQV?
H HR  HRW?
I IS  ISX   ISXE   ISXED?
J JT?
K KU  KUZ   KUZG?
L LV  LVA   LVAH   LVAHG?
M MW MWB?
N NX  NXC?
O OY  OYD   OYDK?
P PZ  PZE   PZEL   PZELK   PZELKA PZELKAR?
Q QA?
R RB  RBG   RBGN   RBGNM?
S SC  SCH   SCHO   SCHON   SCHOND SCHONDU SCHONDU•
T TD  TDI   TDIP   TDIPO   TDIPOE TDIPOEV?
U UE  UEJ?
```

Рис. 83. Перебор для 26 вариантов сложения ЦЕЗАРЯ

нуту, а это значит, что за час можно просмотреть 6 000 исходных вариантов. Позднее число оставшихся вариантов радикально сокращается, так что менее чем за два часа должно быть найдено «правильное» решение или показано, что его не существует. Для примеров в табл. 5 и 6 это можно видеть на рис. 82 и 83. Даже читатель, который лишь смутно знаком с языком, увидит, что выполоть бессмысленные сочетания совсем нетрудно. Для исключения влияния концов, начинать нужно с 6-го столбца.

Для теоретического обоснования см. Приложение «Аксиоматическая теория информации».

Заметим, что в соответствии с разд. 12.3.3 и разд. 12.2.4 для схемы шифрования ВИЖЕНЕРА (многоалфавитного сложения ЦЕЗАРЯ)

$$Z = 17\,576 \quad \text{для периода} \quad d = 3,$$
$$Z = 456\,976 \quad \text{для периода} \quad d = 4,$$

а для перестановки (особого многосимвольного шифра)

$$Z = 40\,320 \quad \text{для ширины} \quad n = 8,$$
$$Z = 362\,880 \quad \text{для ширины} \quad n = 9.$$

Эти цифры показывают пределы атаки перебора. Для общей одноалфавитной подстановки ПЛЕЙФЕЙРА, имеющей сложность Z порядка 10^{25}, такая

```
H HR HRC  HRCN?
S SR SRC?
R RS RSC  RSCU  RSCUO?
H HS HSC  HSCO  HSCOU  HSCOUN HSCOUND  HSCOUNDN
R RH RHC?
S SH SHC?
C CR CRH?
S SR SRH  SRHD?
R RS RSH  RSHD  RSHDO  RSHDON RSHDONU ?
C CS?
S SC SCH  SCHO  SCHON  SCHOND SCHONDU  SCHONDUN•
C CH CHR  CHRN?
S SH SHR  SHRD?
H HS HSR?
C CS?
H HC HCR?
S SC SCR?
C CR CRS?
H HR HRS  HRSD?
R RH RHS?
C CH CHS  CHSD  CHSDD?
R RC RCS?
H HC HCS?
R RC RCH  RCHN  RCHND?
```

Рис. 84. Перебор для 24 вариантов перестановки ширины $d = 4$

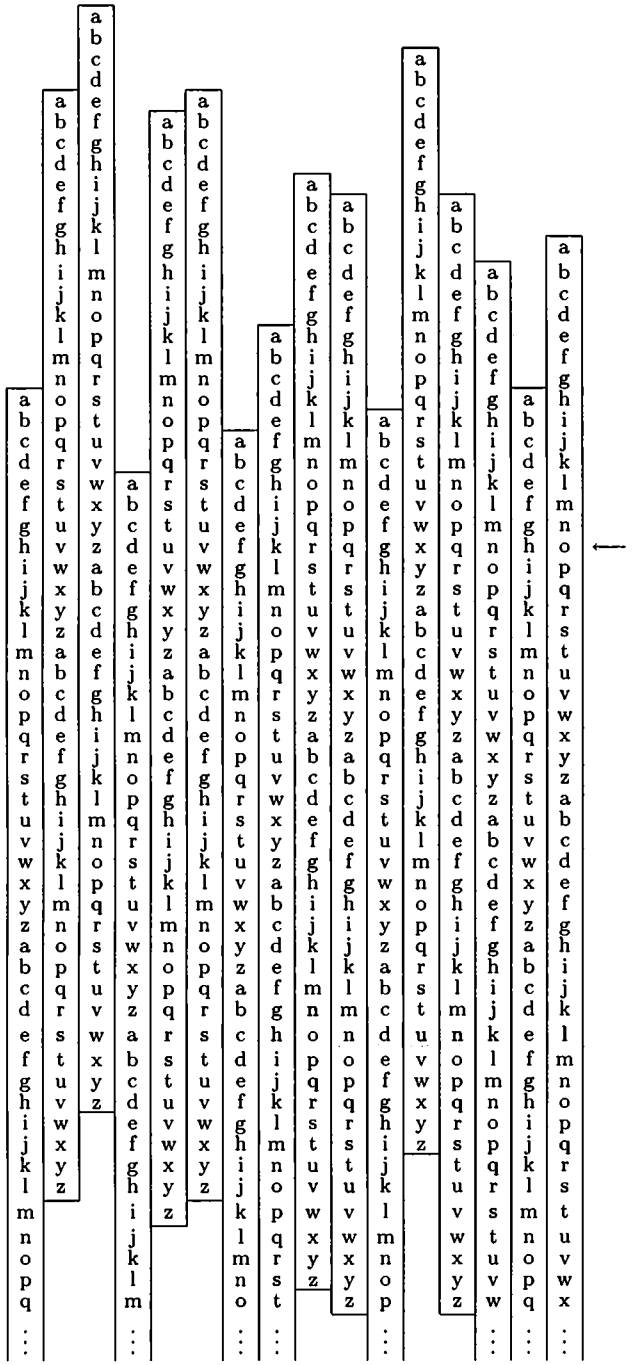


Рис. 85. Метод полос для дешифрования шифра ЦЕЗАРЯ

атака бесполезна, по крайней мере, в ее чистом виде (т. е. без применения вычислительной техники). Компьютеры могут помочь удалить бессмысленные варианты гораздо быстрее, но и этого еще может быть недостаточно. Однако, если высокая комбинаторная сложность Z может быть радикально понижена другими подходящими средствами, то перебор можно выполнить. Другими словами,

Несмотря на то, что атака перебора сама по себе неэффективна, в комбинации с другими также автоматизируемыми атаками она является основным методом разумного криптоанализа.

Перебор используется также и «законным» дешифровальщиком — в случае полифонических шифров. Наиболее известным примером является многоалфавитный шифр с неродственными одноцикловыми подстановочными алфавитами, используемыми в цилиндрах Джефферсона и Базерье, где открытый текст требуется найти среди двух дюжин вариантов.

12.8. Механизированный перебор

12.8.1. Перебор подстановки. Перебор простого сложения ЦЕЗАРЯ можно механизировать при помощи метода полос (лент). Подготовленные полосы, содержащие дубликаты стандартного алфавита, используются для демонстрации криптотекста, а также всех вариантов открытого текста (рис. 84). С той же целью могут быть использованы цилиндры, которые содержат стандартные алфавиты на своих ободах. Здесь механическая дешифровальная помощь заключается в применении шифровального устройства «в обратную сторону». Эту идею можно применить также и к другим механическим средствам. Например, имитация шифромашины ЭНИГМА может быть использована для нахождения одной из $26^3 = 17\,576$ роторных позиций, которая задается фрагментом криптотекста для вероятного слова — в случае, если роторы попали в руки криптоаналитика.

12.8.2. Перебор перестановок. Для перебора перестановок известной ширины n криптотекст выписывается горизонтально в последовательные строки таблицы, имеющей n столбцов, затем таблица разрезается на n вертикальных полос, которые требуется переставить подходящим образом (рис. 85).

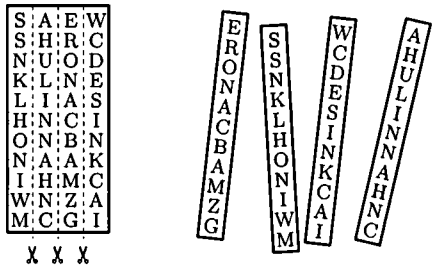


Рис. 86. Метод «разрежь и склей» для дешифрования перестановки

12.8.3. Грубая сила против инвариантности. Криптоанализ путем перебора является методом «грубой силы» и поэтому ограничен предельным значением этой «силы». В последующих главах обсуждаются менее трудоемкие методы криптоанализа, которые основаны на «инвариантных характеристиках используемой криптографической системы» (Кульбак «Statistical Methods in Cryptanalysis», 1935 г.).

Глава 13

Анатомия языка: шаблоны

Не имеет значения, насколько сопротивляется криптограмма; все, что действительно нужно, это «вход», т. е. идентификация одного слова, или трех-четырех букв.

Элен Фуше Гейнс, 1939 *г.*

Каждый язык имеет различные внутренние регулярные структуры, которые трудно подавить. Одну из таких регулярных структур образуют шаблоны.

13.1. Инвариантность повторяющихся шаблонов

Теорема инвариантности 1. *Для всех одноалфавитных простых подстановок, в частности, для всех одноалфавитных линейных простых подстановок* (включая сложение ЦЕЗАРЯ и реверсии), *повторяющиеся шаблоны индивидуальных символов в тексте инвариантны.*

Например, открытый текст w i n t e r s e m e s t e r

и его зашифрования

сложением ЦЕЗАРЯ: Z L Q W H U V H P H V W H U,
обращенным алфавитом: D R M G V I H V N V H G V I,
переставленным алфавитом: V A H O R M N R G R N O R M,

имеют инвариантные шаблоны повторяющихся символов. Согласно Шеннону, шаблоны это как раз «классы вычетов» простых подстановок. Отрезки текста с одним и тем же повторяющимся шаблоном называются идиоморфами.

Одноалфавитные и функциональные *многосимвольные* подстановки $V^{(n)} \to W^{(m)}$ оставляют инвариантными шаблоны повторяющихся полиграмм (с учетом пробелов). Наоборот, омофонные и особенно многоалфавитные подстановки разрушают повторяющиеся шаблоны.

Шаблоны обычно обозначаются конечными числовыми последовательностями в нормальной форме, т. е. каждое число (буква) при его (ее) первом появлении (слева направо) получает наименьший из остающихся номеров. Например, шаблон 1 2 3 3 2 4 1 5 6 записан в нормальной форме.

В приведенном выше криптотексте

V A H O R M N R G R N O R M

группа N R G R N имеет шаблон 12321, N R G R N O R имеет шаблон 1232142, O R M N R G R N O R M — 12342524123. Шаблон 12321 группы N R G R N особенно заметен. Кроме того, слово O R M встречается дважды. Таким образом, шаблон 12345675857456 описывает текст с пересекающимися группами из 6 и 8 букв. Фактически мы знаем, что решением является wintersemester, но трудность состоит в том, что это немецкий идиоморф, а английского, скорее всего, не существует.

Короткие шаблоны допускают много идиоморфов или реализаций осмысленными словами или их фрагментами; так, шаблон 1221 допускает в английском языке идиоморфы, собранные на рис. 86. Предполагается, что этот список (исключающий собственные имена) содержит все слова или их фрагменты (без грамматических вариантов), входящие в «Словарь английского языка» Касселла.

Заметим, что этот список содержит всего несколько слов военного жанра, вроде assa(ult) [нападение], atta(ck), (b)atta(lion), (b)arra(ck), (z)eppe(lin), (sh)ippi(ng) [флот], (m)issi(le) [ракета], (c)ommo(dore) [капитан 1 ранга], и совсем мало слов дипломатического жанра, вроде affa(ir) [МИД], (amb)assa(dor) [посол], assa(sin) [наемный убийца], (chanc)elle(ry) [канцлер], (sh)illi(ng), immi(grant), (comm)issi(on) [комитет]. Пространство поиска реализаций шаблона значительно сужается знанием обстоятельств.

Кроме 1221 другими интересными шаблонами из четырех символов, являются 1211, 1212, 1231, 1122, 1112, 1111. Если для первых трех из них реализация существует, например, lull(aby) [колыбельная], (r)emem(ber) [помнить], (b)eave(r) [бобер], то для остальных трудно найти естественные реализации. Заметим, что шаблон типа 123245678, означающий, что в него входит восемь различных символов, удобнее записать в виде *232*****. Он не выражает ничего иного, кроме шаблона 121. Длинные шаблоны с более чем одним повторяющимся знаком, вроде 12134253, обычно имеют очень мало реализаций или не имеют их вообще, — для данного примера papeline, paperine, papelike.

Вывод ясен: слова и фразы, которые образуют заметные шаблоны, должны быть исключены шифровальщиком, обычно перефразированием, как это регулярно делалось при передаче сообщений Британского Адмиралтейства. Ярким примером является шаблон 1234135426, который по-немецки не допускает иной реализации кроме *heilhitler*. Кто бы отважился в Рейхе исключить это стереотипное окончание? Керкхоффс даже подчеркивает, что повторений вроде французского *pouvez-vous vous défendre* следует избегать. Но в явном противоречии с этим существовала общая практика связи воинских частей — подчеркнуть какую-либо группу, повторив ее, например, OKMMMANAN (разд. 9.2.5) в немецкой армии. Союзники делали то же самое: в союзном конвое SC.48 (Beesley) был сигнал SC48SC48, или CHICKEN-WIRE£CHICKEN-WIRE в сообщении из Блетчли Парк, содержащем дешифровку немецкого сообщения про американские пароли и ответы (Левин).

abbacy cabbage cabbala sabbath scabbard baccalaureate maccabee staccato affable affair baggage braggart haggard laggard allah allay ballad ballast fallacy gallant installation mallard palladium parallax wallaby diagrammatic flammable gamma grammar mamma programmatic annalist annals bandanna cannabis hosanna manna savannah appal apparatus apparel apparent kappa arrack arraign arrange arrant arras array barrack barracuda barrage carragheen embarrass narrate tarragon warrant ambassador assail assassin assault assay cassandra massacre massage passage vassal wassail attach attack attain rattan attar battalion coattail rattan regatta wattage piazza beebread boob booby deed deedless indeed doodle ebbed eccentric bedded reddear redden shredder wedded effect effeminate effendi efferent effervesce effete begged bootlegger egged legged pegged trekked aquarelle bagatelle belle chancellery chanterelle driveller dweller excellent feller fontanelle gazelle groveller hellebor hellenic impellent intellect jeweller libeller mademoiselle nacelle pellet propeller repellent seller teller traveller emmet barrenness comedienne fennec fennel jennet kennel rennet tenner pepper stepper zeppelin deterrent ferret interregnum interrelation overreact parterre terrestrial adressee dessert dresser essence essential finesse largesse lessen messenger noblesse quintessence tesselate vessel begetter better burette corvette curette fetter gazette getter letter marionette pirouette rosette roulette setter silhouette geegee googol heehaw capriccio pasticcio forbidding yiddish difficile difficult griffin tiffin biggish bacilli billiard billion brilliant chilli cyrillic fillip illicit illinois illiquid illiberal illiterate illimitable lilliput milliard millibar milligram milliliter millimeter milliner millionaire millivolt penicillin postillion shilling silliness tranquillize trillion trillium vanillin gimmick immigrant imminent immiscible immitigable finnish innings pinniped zinnia pippin irrigate irritate admission commission dissident dissimilar dissipate emission fissile fission fortissimo missile mission missive omission permission permissive acquitting fitting kittiwake civvies noon broccoli sirocco apollo collocate colloid colloquial colloquium follow hollow rollout common accommodate commode commodore commotion connote opponent opportune oppose opposite borrow corroborate corrode horror morrow sorrow blossom crossover blotto bottom cotton grotto lotto motto ottoman risotto glowworm powwow peep poop career seesaw teeter teethe teetotal teetotum toot toothache tootle hubbub succulent succumb succuss pullup nummulite unnumbered chaussure guttural

Рис. 87. Реализации шаблона 1221 в английском языке по Кейсменту

Число шаблонов из n элементов равно числу разбиений n в сумму натуральных чисел, или числу Белла $B(n)$, которое с ростом n растет довольно быстро, как показывает следующая таблица:

n	0	1	2	3	4	5	6	7	8	9	10	11	12
$B(n)$	1	1	2	5	15	52	203	877	4 140	21 147	115 975	678 570	4 213 597

13.2. Исключение из шифровальных методов

Теорему 1 можно применить в отрицательном смысле для исключения из рассмотрения одноалфавитных функциональных простых подстановок — в том случае, когда криптотекст содержит шаблонов не больше, чем случайный текст. Однако здесь рекомендуется осторожность. Например, недостаток удвоенных символов не означает ничего особенного. Еще начиная с работы Дж. Б. и М. Ардженти профессиональные криптографы знали правило создания препятствий в нахождении шаблонов путем запрещения удвоения символов даже в открытом тексте, например, следовало писать *sigilo* вместо *sigillo*. К тому же классическое подавление пробелов между словами приводит к сокращению образования шаблонов (разд. 13.6.1). Подавление является полифоническим шагом; в довольно редких случаях подавление пробелов между словами приводит к нарушению инъективности: *the messages that were translated* [сообщения, которые были переведены] — *the messages that we retranslated* [сообщения, которые мы перевели (*англ.*)], или *we came together* [мы приходили вместе] — *we came to get her* [мы зашли за ней].

13.3. Нахождение шаблонов

Теорему 1 можно использовать в положительном смысле, если существуют основания считать, что мы имеем дело с одноалфавитной функциональной простой подстановкой.

13.3.1. Один пример. В следующем примере (Элен Фуше Гэйнс), зашифрованном простой подстановкой, пробелы, обозначаемые ␣, не подавлены).

F D R J N U ␣ H V X X U ␣ R D ␣ M D ␣ S K V S O ␣ P J R K ␣ Z D
Y F Z J X ␣ G S R R V T ␣ Q Y R ␣ W D A R W D F V ␣ R K V ␣ D R
K V T ␣ D F ␣ S Z Z D Y F R ␣ D N ␣ N V O V T S X ␣ S A W V Z R

Гэйнс начинает со слов с шаблоном 1231, где 1 обозначает пробел ␣, т. е. с двухбуквенных слов ␣RD␣, ␣MD␣, ␣DF␣, ␣DN␣. Встречаемость D в каждом из них наводит на мысль проверить реализации /of/, /оп/, /or/, /do/, /go/, /по/, /to/. Таким образом, имеем вход D $\hat{=}$ o. Кроме того, дважды встречается шаблон 12341, точнее, трехбуквенные слова ␣RKV␣ и ␣QYR␣ с общей буквой R, которая также встречается в группе ␣RD␣. Среди трехбуквенных слов открытого текста, которые начинаются с /d/, /g/, /п/ или /t/, разумно предположить /the/. Предполагая RKV $\hat{=}$ the, из слова ␣DRKVT␣ получаем ␣othcT␣, и почти определенно должно быть $T \hat{=}$ r. Таким образом, пять букв известны, и частично дешифрованное сообщение читается так:

F o t J N U ␣ H e X X U ␣ t o ␣ M o ␣ S h e S O ␣ P J t h ␣ Z o
Y F Z J X ␣ G S t t e r ␣ Q Y t ␣ W o A t W o F e ␣ t h e ␣ o t
h e r ␣ o F ␣ S Z Z o Y F t ␣ o N ␣ N e O e r S X ␣ S A W e Z t

Подтверждение получаем из GSRRVT, превращенной в GStter. Для другого трехбуквенного слова QYt реализации /not/, /got/, /out/, /yet/ исключаются, так как /o/ и /e/ уже определены, остается /but/. Кроме того, остаются

лишь возможности DF ≙ on и DN ≙ of (или наоборот), так как /r/ уже определено. В первом (счастливом) случае получаем теперь следующий фрагмент:

```
n o t J f U␣H e X X U␣t o␣M o ␣S h e S O␣P J t h ␣Z o
u n Z J X␣G S t t e r␣b u t ␣W o A t W o n e␣t h e ␣o t
h e r␣o n␣S Z Z o u n t␣o f ␣f e O e r S X␣S A W e Z t
```

Теперь SZZount читается как /account/, так что из ZounZJX получаем councJX, что означает /council/. Получаем следующий фрагмент:

```
n o t i f U␣H e l l U␣t o␣M o ␣a h e a O␣P i t h ␣c o
u n c i l␣G a t t e r␣b u t ␣W o A t W o n e␣t h e ␣o t
h e r␣o n␣a c c o u n t␣o f ␣f e O e r a l␣S A W e c t
```

А это уже можно прочитать как открытый текст (может быть, кроме /Helly/, что может быть собственным именем).

Фазу проведенного здесь дешифрования от входа и до трех-пяти символов после него, находимых в порядке рабочей гипотезы, можно назвать «шагом», следующую фазу — до нахождения восьми-десяти символов, когда уже больше нет сомнений — «рысью», а остальную работу — «галопом». Это отражено в строении таблицы дешифрования:

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
      o           h               t r e
        n               f   b             u
              i                   a       l c
s           m         g   d w         y   p
```

Неясно лишь, что означает H, тогда как B, C, E, I и L не встречались в криптотексте. В этой последней фазе надо попытаться восстановить полную таблицу. Читатель может заметить, что N и F переходят друг в друга; поэтому oF и oN становятся /on/ и /of/. То же самое, повидимому, верно и для пар A и S, D и O, P и W, R и T, U и Y. Если шифрование было взаимообратным, то из K≙h получаем K≙k и Helly означает /kelly/. Полная таблица шифрования принимает вид

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
s q z o v n m k j i h x g f d w b t a r y e p l u c
- -   -   - - - - - -   - - -   -   -
```

Поскольку подчеркнутые буквы нижнего ряда пробегаются в обратном порядке относительно верхнего, то предположительно существует конструкция подстановочного алфавита, полученного при помощи известного пароля. На самом деле это переупорядочение производит инволютивная подстановка[1)]

[1)] Вудхэл и Таунсенд в 1779 г. обеспечивали генерала Вашингтона ценной информацией из Нью-Йорка, который был оккупирован английскими войсками; они использовали для прикрытия псевдонимы Culper Sr. и Culper Jr. (разд. 4.4.1).Было ли это сокращением от Culperer, которое иногда использовалось в криптографической литературе для построения ключей? Калперер (1660–1738 гг.) был известным английским создателем музыкальных инструментов.

```
c u l p e r a b d f g h i
↕
z y x w v t s q o n m k j
```

Прокурор округа мог обосновать обвинительный приговор на таком абсолютно правдоподобном дешифровании, полностью раскрывающем систему. «Систематическое и точное восстановление метода шифрования и используемых паролей и ключей» (Рорбах, 1946 г.) требуется, если криптоаналитики выступают свидетелями обвинения, подобно Базерье в 1898 г. в судебном процессе против герцога Орлеанского или Элизабет Фридман, жене У. Ф. Фридмана в судебном разбирательстве против компании Consolidated Exporters Company, за организацию контрабанды во время сухого закона.

13.3.2. Аристократы. Должно быть ясно, что дешифрование в приведенном выше примере было таким легким потому, что вопреки профессиональной традиции пробелы между словами не были запрещены. Не требовалось запрета на пробелы и в правилах игры «Тайны» («Cryhtos»), проводимой в 1977 г. американскими газетами (рис. 87), по крайней мере для тех, кто выступал под именем «аристократы». Пробелы и знаки пунктуации оставались в неприкосновенности, только буквы разрешалось менять в словаре символов криптотекста, причем ни одна буква не могла представлять себя. Длина криптотекста у настоящих аристократов (т. е. не пользующихся «подсказками») составляла от 75 до 100 символов, т. е. была довольно большой, учитывая что для простой подстановки с переставленным алфавитом расстояние единственности равно примерно 25. Зато криптотекст мог содержать наиболее экстраординарные и необычные американские (но не иностранные) слова и, если не считать требования формальной грамматической правильности, не нуждался в придании смысла; понять его возможно было так же трудно, как

Cryptoquip

K I O S P X F I E V B O S F E F P M H Y I M K K J X F I E V B J F K Y F I - K O M H

Yesterday's Cryptoquip— GLUM GOLFER TODAY STUDIES SNOWMEN ON FAIRWAY.

Today's Cryptoquip clue: S equals C

The Cryptoquip is a simple substitution cypher in which each letter used stands for another. If you think that X equals O, it will equal O throughout the puzzle. Single letters, short words, and words using an apostrophe can give you clues to locating vowels. Solution is accomplished by trial and error.

Cryptoquip

K R K K R L H P L R U I O Z G K A Y M - M G O R A U L Y P Q , Q R U A H U L Z I U

Yesterday's Cryptoquip— TRICK HARMONICA MAKES PRETTY HARMONY AT PARTIES.

Today's Cryptoquip clue: I equals M

The Cryptoquip is a simple substitution cypher in which each letter used stands for another. If you think that X equals O, it will equal O throughout the puzzle. Single letters, short words, and words using an apostrophe can give you clues to locating vowels. Solution is accomplished by trial and error.

Рис. 88. Криптографические головоломки из Лос-Анжелес Таймс, 1977 г.

и решить саму криптограмму. Могли встречаться слова из биологии, вроде *pterodactyl*, *ichtyomancy* и *syzygy*, но также могли быть найдены *yclept*, *crwth* и *cwm*. Текст мог быть выбран так, чтобы нормальная частота букв и фраз была фальсифицирована, а это значит, что методы, основанные на частотном анализе (которые обсуждаются в следующей главе), бесполезны («все внимание шифровальщика должно быть направлено на подтасовку характеристик букв», Фуше Гэйнс).

Кан дает решение одной криптограммы типа «аристократ», которую сам описывает так: «Сложная криптограмма, содержит ловушки (в которые попадают неосторожные решатели), ненормальные частоты, невообразимые комбинации согласных, причудливые окончания, замысловатые вопросы, ставящие в тупик, вроде «myrrh».

13.3.3. Липограммы. Существуют тексты (липограммы), написанные полностью без какой-нибудь буквы. Наиболее известен мастерски написанный, роман Райта «Gadsby» (Гэдсби), (Wetzel Publishing Co., Los Angeles, 1939, 287 pp.) (рис. 88). В предисловии Райт написал, что у него в машинке западает клавиша /e/ и потому он намерен избегать буквы /e/ в своем манускрипте, что он блестяще выполнил, несмотря на то, что /e/ — самая частая буква английского языка.

XXIX

GADSBY WAS WALKING back from a visit down in Branton Hills' manufacturing district on a Saturday night. A busy day's traffic had had its noisy run; and with not many folks in sight, His Honor got along without having to stop to grasp a hand, or talk; for a Mayor out of City Hall is a shining mark for any politician. And so, coming to Broadway, a booming bass drum and sounds of singing, told of a small Salvation Army unit carrying on amidst Broadway's night shopping crowds. Gadsby, walking toward that group, saw a young girl, back towards him, just finishing a long, soulful oration, saying:—

". . .and I can say this to you, for I know what I am talking about; for I was brought up *in a pool of liquor!!*"

As that army group was starting to march on, with this girl turning towards Gadsby, His Honor had to gasp, astonishingly:—

"Why! Mary Antor!!"

"Oh! If it isn't Mayor Gadsby! I don't run across you much, now-a-days. How is Lady Gadsby holding up during this awful war?"

[201]

Рис. 89. Страница из «Гэдсби» Эрнеста Винсента Райта

По тому же пути, но с гораздо большей претензией пошел также Перес (1936–1982 гг.) со своим романом «La disparation», вышедшим в 1969 г. (Английский перевод «A Void» [Вакуум (*англ.*)] Эдера, (Harper Collins, 1995, 285 pp.). Перес, который также любил играть с акронимами, акростихами, анаграммами, палиндромами, был в своей лингвистической деятельности активен (он использовал компьютерные программы и представил в 1969 г. палиндром из 5000 букв), опубликовал в 1973 г. историю липограмм. Задолго до этого, в 1820 г. Риттлер из Вены опубликовал роман «Die Zwillinge» [Близнецы (*нем.*)], написанный целиком без буквы /r/. Но даже еще раньше, в 1800 г. великий русский поэт Гаврила Романович Державин (1743–1816 гг.) написал новеллу «Шутливое желание» ни разу не использовав букву /р/ и использовав всего несколько раз букву /о/[2)].

Джеймс Джойс тоже писал криптологическую прозу. Вот последние фразы из «Finnegans Wake» [Поминки по Финнегану]:

> End here. Us then. Finn, again!
> Take. Bussoftlee, mememormee!
> Till thousendsthee. The keys to. Given!
> Lps. A way a lone a last a loved a long the.

Если дать эти стихи криптоаналитику, они создадут ему значительные трудности. Ранний роман Джойса «Улисс» также содержит множество криптологических головоломок.

Криптологически эти любопытные вещи, конечно, не слишком важны, впрочем, как и криптологические изыски Владимира Набокова, включенные в его работы. Вацлав Гавел сделал предметом шуток (секретный) язык марксистско-ленинской партии («новояз» Оруэлла).

13.4. Нахождение многосимвольных шаблонов

Одной из причин использования кодов является подавление ими заметных шаблонов. Но плохо спроектированные коды, которые не обращают внимания на часто используемые длинные фразы, а также плохое использование кодов, приводят к их взлому.

13.4.1. Луиджи Сакко, который в 1916 г. в возрасте 32 лет стал главой Департамента криптографии Итальянской ставки главного командования, на

[2)]Эта «новелла» оказалась известным стихотворением, написанным в 1802 г. и состоящим всего из 12 строк. В нем действительно отсутствует буква /р/, но буква /о/, кроме одного раза в заглавии, встречается еще 12 раз. Стихотворение прославилось куплетами Томского из оперы Чайковского «Пиковая дама»:

> Если б милые девицы
> так могли летать, как птицы,
> и садились на сучках, —
> я желал бы быть сучочком,
> чтобы тысячам девочкам
> на моих сидеть ветвях.
>
> Пусть сидели бы и пели,
> вили гнезда и свистели,
> выводили и птенцов;
> никогда б я не сгибался,
> вечно ими любовался,
> был счастливей всех сучков!

— *Прим. перев.*

фронте под Изонцо и Пиаве в северной Италии получил 30 июня 1918 г. два радиоперехвата с одним и тем же окончанием

... 4 92073 06583 47295 89255 07325 58347 29264.

Это было серьезной ошибкой австрийцев, но хуже всего было то, что фрагмент 073 * * * 5834729 был повторен на коротком расстоянии в 18 разрядов. Это дало ясное указание об использовании трехзначного кода:

... 492 073 065 834 729 589 255 073 255 834 729 264.

Сакко был немного знаком с австрийскими обычаями и предположил, что одно длинное слово было неосторожно зашифровано побуквенно. Шаблон кодовой группы был таким: 12345672745 8, и у Сакко, который был инженером, возникла великолепная идея прочитать это как *radiostation*. Ленивые австрийские кодировщики не видели необходимости искать кодовые группы для *radio* и *station*. И если в исключительных случаях — например, для собственных имен — шифрование буквы за буквой является неизбежным, то такое разоблачение кода для отдельных букв не должно было случаться в начале или конце текста.

Как бы то ни было, это дало Сакко вход, чтобы дешифровать другие слова, зашифрованные побуквенно, — и таким образом в итоге взломать весь код. Будьте уверены, криптоаналитически австрийцы были подготовлены не хуже итальянцев. В военной шифрогруппе под командой полковника Ронге был Итальянский отдел, в котором майор (позднее полковник) Фигль отлично справлялся со своей работой, подобно своему коллеге Покорны из Русского отдела.

13.4.2. Но даже без воображения Сакко можно было бы сразу получить вход, используя изготовленный заводским способом список шаблонов и их реализаций. Для шаблона 123456727458 реализация $radiostation$, очень возможно, является единственной, а если нет, то метод проб и ошибок с немногими реализациями дал бы немедленные результаты.

13.5. Метод вероятного слова

До сих пор использовались лишь формальные соображения, и предполагалась не более чем догадка в отношении естественного языка, лежащего в основе криптограммы. Теперь мы примем в расчет другую информацию. Самым излюбленным методом для входа в одноалфавитную простую подстановку является метод вероятного слова (по-французски — *mot probable*, по-немецки — *wahrscheinliches Wort*). Это не какие-либо заметные шаблоны, которые мы разыскиваем в тексте (а позднее ищем их реализации); вместо этого в криптотексте ищется шаблон вероятного слова — имеется ли он и если да, то где он может встретиться. Каждый найденный фрагмент криптотекста вместе с вероятным словом образует «шпаргалку».

13.5.1. Шпаргалки. Этот метод был описан еще Джиованни Баттистой Порта (1535–1615 гг.) в трактате *De furtivis* [Тайное, (*лат.*)], 1563 г. Например, если (в соответствии с обстоятельствами) вероятным словом является *division*, то требуется поиск шаблона 12131 слова (d)ivisi(on). Из рис. 95 видно, что в военной переписке вероятность нахождения ложного слова с таким шаблоном довольно мала, хоть шаблон и короткий.

Вместо одного слова можно использовать целые фразы, вроде «Oberkommando der Wehrmacht», или «Combined Chiefs of Staff». В частности, подходящими являются стереотипные выражения, часто используемые в начале и конце открытого текста как в коммерческом, так и в военном мире. Например,

> *Reference to you letter,*
>
> *Hochachtungsvoll Ihr,*
>
> *An SS-Gruppenführer Generalleitenant der Waffen-SS Berger, Berlin W.*35, *SS-Hauptamt, mit der Bitte um absprachegemässe Weitergabe,*
>
> *From Algeria to Washington,* 21.7. *To the State Department in Washington. Strictly confidential. Most urgent and personal for Deputy Under State Secretary. From Murphy.*

Эти примеры показывают, что обычно существует достаточно много шпаргалок. Не помогает даже «Русское соединение» (когда криптотекст произвольно разрезается в одном месте и куски соединяются в обратном порядке), так как это не устраняет шаблонов. Кроме того, способность проникнуть в ситуацию противника и умение поставить себя на его место может вызвать цепную реакцию, которую хорошо описал Гуд как «успех, приводящий к большему успеху». И даже если не оказывается ни одной шпаргалки, они могут быть спровоцированы: например, в результате определенных военных действий вполне ожидаемы слова вроде *attack* или *bombardment*.

13.5.2. Мэрфи и Ягер. Бессмертную славу за заслуги перед Германией во Второй мировой войне добыл американский дипломат, а позднее заместитель государственного секретаря Мэрфи (1894–1978 гг.), который подчеркивал свою значительность постоянным использованием выражений «От Мэрфи», «Для Мэрфи» в своих телеграммах. Тем не менее лейтенант Ягер, упомянутый в разд. 4.4, помог союзникам еще больше. К сожалению, повиновение не заменяет дисциплины, которая требует мысли и потому редка. Постоянно сообщать «Сообщений нет» — противоречиво.

13.5.3. Führerbefehl [Приказ фюрера (нем.)]. Следующий пример, придуманный Кратцером, основан на неизвестном Führerbefehl [приказе фюрера] 1939 г. (рис. 89). Пример иллюстрирует тезис: как много может дать единственное вероятное слово.

Учитывая обстоятельства, легко можно было предположить, что в открытом тексте сообщения (приказа Гитлера) встретится год «1939», а ввиду напыщенного стиля, усвоенного гитлеровскими генералами, не следует даже

J H K H L P H N R P P D Q G R V D F K H Z H L V X Q J Q U H
L Q V I X H U G L H N U L H J V I X H K U X Q J Q D F K G H
P D O O H S R O L W L V F K H Q P R H J O L F K N H L W H Q
H U V F K R H S I W V L Q G X P D X I I U L H G O L F K H P
Z H J H H L Q H I X H U G H X W V F K O D Q G X Q H U W U D
H J O L F K H O D J H D Q V H L Q H U R V W J U H Q C H C X
E H V H L W L J H Q K D E H L F K P L F K C X U J H Z D O W
V D P H Q O R H V X Q J H Q W V F K O R V V H Q G H U D Q J
U L I I D X I S R O H Q L V W Q D F K G H Q I X H U G H Q I
D O O Z H L V V J H W U R I I H Q H Q Y R U E H U H L W X Q
J H Q C X I X H K U H Q P L W G H Q D E D H Q G H U X Q J H
Q G L H V L F K E H L P K H H U G X U F K G H Q L Q C Z L V
F K H Q I D V W Y R O O H Q G H W H Q D X I P D U V F K H U
J H E H Q D X I J D E H Q Y H U W H L O X Q J X Q G R S H U
D W L R Q V C L H O E O H L E H Q X Q Y H U D H Q G H U W D
Q J U L I I V W D J H U V W H U Q H X Q W H U Q H X Q C H K
Q K X Q G H U W Q H X Q X Q G G U H L C L J D Q J U L I I V
C H L W Y L H U X K U I X H Q I X Q G Y L H U C L J

Рис. 90. Фиктивная зашифровка приказа Гитлера, 1939 г.

J H K **H L P H** N R P P D Q G R V D F K H Z H L V X Q J Q U H
L Q V I X H U G L H N U L H J V I X H K U X Q J Q D F K G H
P D O O H S R O L W L V F K H Q P R H J O L F K N **H L W H** Q
H U V F K R H S I W V L Q G **X P D X** I I U L H G O L F K **H P**
Z H J H **H L Q H I X H U G H** X W V F K O D **Q G X Q** H U W U D
H J O L F K H O D **J H** D Q V **H L Q H** U R V W J U **H Q C H** C X
E H V H L W L J H Q K D E H L F K P L F K C X U J H Z D O W
V D P H Q O R H V X **Q J H Q** W V F K O R V V **H Q G H** U D Q J
U L I **I** D **X I** S R O H Q L V W Q D F K G H Q I X **H U G H** Q I
D O O Z H L V V J H W U R I I H Q H Q Y R **U E H U** H L W X **Q**
J H Q C X I X **H K U H** Q P L W G H Q D E D **H Q G H** U X Q J H
Q G **L H V L** F K E H L P K H H **U G X U** F K G H Q L Q C Z L V
F K H Q I D V W Y R O O **H Q G H** W H Q D X I P D U V F K **H U**
J H E H Q D X I J D E **H Q** Y **H U W H** L O X **Q J X Q** G R S H U
D W L R Q V C L H O E O **H L E H** Q X Q Y **H U D H Q G H** U W D
Q J U L I I V W D J H U V W H U **Q H X Q** W H U **Q H X Q** C H K
Q K X Q G H U W **Q H X Q** X Q G G U H L C L J D Q J U L I I V
C H L W Y L H U X K U I X H Q I X Q G Y L H U C L J

Рис. 91. Встречаемость шаблона 1231

исключать, что, вопреки всем предосторожностям, выражение «neunzehnhundertneununddreissig» (в 1939, нем) встретится побуквенно. Это наводит на

```
J e h e L P e N R P P D n d R V D F h e Z e L V u n J n r e
L n V I u e r d L e N r L e J V I u e h r u n J n D F h d e
P D O O e S R O L t L V F h e n P R e J O L F h N e L t e n
e r V F h R e S I t V L Q d u P D u I I r L e d O L F h e P
Z e J e e L n e I u e r d e u t V F h O D n d u n e r t r D
e J O L F h e O D J e D n V e L n e r R V t J r e n z e z u
E e V e L t L J e n h D E e L F h P L F h z u r J e Z D O t
V D P e n O R e V u n J e n t V F h O R V V e n d e r D n J
r L I I D u I S R O e n L V t n D F h d e n I u e r d e n I
D O O Z e L V V J e t r R I I e n e n Y R r E e r e L t u n
J e n z u I u e h r e n P L t d e n D E D e n d e r u n J e
n d L e V L F h E e L P h e e r d u r F h d e n L n z Z L V
F h e n I D V t Y R O O e n d e t e n D u I P D r V F h e r
J e E e n D u I J D E e n Y e r t e L O u n J u n d R S e r
D t L R n V z L e O E O e L E e n u n Y e r D e n d e r t D
n J r L I I V t D J e r V t e r n e u n t e r n e u n z e h
n h u n d e r t n e u n u n d d r e L z L J D n J r L I I V
z e L t Y L e r u h r I u e n I u n d Y L e r z L J
```

Рис. 92. Фрагментарное дешифрование с помощью «neunzehnhundertneun»

```
g e h e i P e N R P P D n d R V D F h e Z e i V u n g n r e
i n V f u e r d i e N r i e g V f u e h r u n g n D F h d e
P D O O e S R O i t i V F h e n P R e g O i F h N e i t e n
e r V F h R e S f t V i Q d u P D u f f r i e d O i F h e P
Z e g e e i n e f u e r d e u t V F h O D n d u n e r t r D
e g O i F h e O D g e D n V e i n e r R V t g r e n z e z u
E e V e i t i g e n h D E e i F h P i F h z u r g e Z D O t
V D P e n O R e V u n g e n t V F h O R V V e n d e r D n g
r i f f D u f S R O e n i V t n D F h d e n f u e r d e n f
D O O Z e i V V g e t r R f f e n e n v R r E e r e i t u n
g e n z u f u e h r e n P i t d e n D E D e n d e r u n g e
n d i e V i F h E e i P h e e r d u r F h d e n i n z Z i s
F h e n f D V t v R O O e n d e t e n D u f P D r V F h e r
g e E e n D u f g D E e n v e r t e i O u n g u n d R S e r
D t i R n V z i e O E O e i E e n u n v e r D e n d e r t D
n g r i f f V t D g e r V t e r n e u n t e r n e u n z e h
n h u n d e r t n e u n u n d d r e i z i g D n g r i f f V
z e i t v i e r u h r f u e n f u n d v i e r z i g
```

Рис. 93. Дальнейшее фрагментарное дешифрование с помощью «vieruhrfuenfundvierzig»

```
g e h e i m e N o m m a n d o s a c h e Z e i s u n g n o e
i n s f u e r d i e N r i e g s f u e h r u n g n a c h d e
m a O O e S o O i t i s c h e n m o e g O i c h N e i t e n
e r s c h o e S f t s i n d u m a u f f r i e d O i c h e m
Z e g e e i n e f u e r d e u t s c h O a n d u n e r t r a
e g O i c h e O a g e a n s e i n e r o s t g r e n z e z u
E e s e i t i g e n h a E e i c h m i c h z u r g e Z a O t
s a m e n O o e s u n g e n t s c h O o s s e n d e r a n g
r i f f a u f S o O e n i s t n a c h d e n f u e r d e n f
a O O Z e i s s g e t r o f f e n e n v o r E e r e i t u n
g e n z u f u e h r e n m i t d e n a E a e n d e r u n g e
n d i e s i c h E e i m h e e r d u r c h d e n i n z Z i s
c h e n f a s t v o O O e n d e t e n a u f m a r s c h e r
g e E e n a u f g a E e n v e r t e i O u n g u n d o S e r
a t i o n s z i e O E O e i E e n u n v e r a e n d e r t a
n g r i f f s t a g e r s t e r n e u n t e r n e u n z e h
n h u n d e r t n e u n u n d d r e i z i g a n g r i f f s
z e i t v i e r u h r f u e n f u n d v i e r z i g
```

Рис. 94. Последнее промежуточное дешифрование

```
g e h e i m e k o m m a n d o s a c h e w e i s u n g n o e
i n s f u e r d i e k r i e g s f u e h r u n g n a c h d e
m a l l e p o l i t i s c h e n m o e g l i c h k e i t e n
e r s c h o e p f t s i n d u m a u f f r i e d l i c h e m
w e g e e i n e f u e r d e u t s c h l a n d u n e r t r a
e g l i c h e l a g e a n s e i n e r o s t g r e n z e z u
b e s e i t i g e n h a b e i c h m i c h z u r g e w a l t
s a m e n l o e s u n g e n t s c h l o s s e n d e r a n g
r i f f a u f p o l e n i s t n a c h d e n f u e r d e n f
a l l w e i s s g e t r o f f e n e n v o r b e r e i t u n
g e n z u f u e h r e n m i t d e n a b a e n d e r u n g e
n d i e s i c h b e i m h e e r d u r c h d e n i n z w i s
c h e n f a s t v o l l e n d e t e n a u f m a r s c h e r
g e b e n a u f g a b e n v e r t e i l u n g u n d o p e r
a t i o n s z i e l b l e i b e n u n v e r a e n d e r t a
n g r i f f s t a g e r s t e r n e u n t e r n e u n z e h
n h u n d e r t n e u n u n d d r e i z i g a n g r i f f s
z e i t v i e r u h r f u e n f u n d v i e r z i g
```

Рис. 95. Окончательное дешифрование: Weisung № 1 für die Kriegführung

мысль о поиске шаблона 1231 слова «**neun**», несмотря на то, что оно очень коротко и можно ожидать много «промахов» («слепых попаданий»).

И действительно, такой шаблон встречается несколько раз (рис. 90), в частности, как **HOGH** — четыре раза, и как **QHXQ** — дважды в третьей снизу строке и один раз во второй снизу. Слову «**neun**zehnhundert**neun**» соответствует вторая группа **QHXQ** третьей снизу строки и та же группа второй снизу строки, так как расстояние между этими группами равно расстоянию между группами **neun** в рассматриваемом слове.

Итак, найден вход, спровоцированный повторением группы **QHXQ**, приводящий к первому предварительному дешифрованию, показанному на рис. 91.

Очевидно, что в конце текста имеется больше дат. Не нужно много воображения, чтобы прочитать самый конец как «**vier**uhr**f**uen**f**und**vierzig**». Это дает фрагментарное дешифрование, приведенное на рис. 92.

Таким образом, дешифрована уже дюжина символов:

```
...d e f g h i....n...r.t u v...z
...GHIJKL....Q...U.WXY...C
```

Текст на рис. 92 можно прочитать уже совершенно бегло. Результаты /m/ для P, /a/ для D, /s/ для V, /o/ для R и /c/ для F подтверждают, что мы на верном пути. Теперь восстановлены уже 17 символов:

```
a.c d e f g h i...m n o..r s t u v...z
D.FGHIJKL...PQR..UVWXY...C
```

Рис. 93 показывает этот последний промежуточный результат, а рис. 94 представляет окончательный результат, а именно Weisung № 1 für die Kriegsführung [Предписание № 1 для ведения войны (*нем.*)]

13.6. Автоматический перебор реализаций шаблона

Элен Фуше Гэйнс указала, что изготовленные фабричным способом списки слов с одним и тем же шаблоном могут помочь решить наиболее запутанные одноалфавитные подстановки.

13.6.1. Списки. Можно спокойно предположить, что профессиональные криптографические организации знают это и что они уже придумали, как практически использовать эту идею, по крайней мере, с тех пор, как память компьютера с большими магнитными лентами стала достаточно большой (около 1955 г.). Потом по частной инициативе специфицированные таблицы английских реализаций шаблонов были опубликованы (до 12 букв — в 1971 г. и 1972 г. Левиным, и до 15 букв — в 1978 г., 1982 г. и 1983 г. — Эндрю. Список делается способом KWIC («Key Word in Context») [Ключевое слово в контексте (*англ.*)], когда начало и конец слова, в которое входит реализация шаблона (левый и правый контекст) печатаются в круглых скобках. Рис. 95 дает многоязычный пример реализации шаблона 12131 слова (d)ivisi(on). Заметим, что слова anana(s) и (r)ococo принадлежат не к шаблону 12131, а к шаблону 12121.

(m) acada (m)	ebene	(fr) igidi (ty)	
(m) ahara (ni)	(l) edere (inband)	(r) igidi (ty)	(l) oboto (my)
alaba (ma)	(h) egeme (ister)	(n) ihili (sm)	(s) olomo (n)
(m) alaga	(v) eheme (nt)	(b) ikini	(d) oloro (sa)
(c) alama (ry)	(b) elebe (n)	(m) iliti (a)	(g) onoko (ccus)
(p) alata (l)	(b) elege (n)	imiti (eren)	(m) onolo (gue)
(m) alaya	(g) elege (n)	(l) imiti (eren)	(m) onopo (ly)
(t) amara	eleme (nt)	(d) irigi (eren)	(m) onoto (ny)
(p) anama	(t) eleme (try)	(v) isiti (eren)	(t) opolo (gy)
(s) araba (nd)	(h) elene	(c) ivili (an)	(d) oxolo (gy)
(f) arada (y)	(s) elene	(d) ividi (eren)	
(k) araja (n)	(g) elese (n)	(d) ivisi (on)	
(c) arapa (ce)	eleve (n)		(c) umulu (s)
(c) arava (n)	eleve		
(c) atama (ran)	(h) exere (i)		
(c) atara (ct)			
(c) atafa (lque)			(s) tatut

Рис. 96. Двуязычный KWIC — список слов с шаблоном 12131 слова (d)ivisi(on)

Такие коллекции шаблонов можно собирать механически, используя словари рассматриваемого языка или языков и оптическое сканирование. При таком способе, однако, пробел препятствует контактам между словами; шаблоны, возникающие в силу запрета на пробелы, не берутся в расчет. Также можно игнорировать и грамматические окончания. Поэтому лучше использовать тексты рассматриваемого жанра, объем которых доходит до миллиарда символов — например, годовая подшивка газеты, записанная на CD.

13.6.2. Поиск шаблонов. Компьютерная поддержка, т. е. согласованная с компьютером работа полезна, когда для данного вероятного слова в выводимом на экран криптотексте ищутся реализации шаблона этого слова. Помощь компьютера особенно необходима, если нет ни одного вероятного слова и не дано ни одного шаблона, но из криптотекста надо извлечь редкие и повторяющиеся шаблоны[3]. Если какой-нибудь шаблон существует (т. е. если текст достаточно длинный), то это почти обязательно приводит к некоторому входу. При этом мы не говорим о том, что последующее фрагментарное дешифрование с компьютерной поддержкой может быть сделано полуавтоматическим, с небольшим участием человека.

Если это делать систематически, то такой избавленный от интуиции метод нахождения шаблонов может быть полностью автоматизирован; и, действующий без всяких семантических предположений, он является первым примером атаки, основанной только на анализе криптотекста («чистый криптоанализ»). Проблема состоит в том, чтобы сохранять пространство поиска малым, а чис-

[3] К поиску повторяющихся шаблонов мы еще вернемся в разд. 17.4.

ло допустимых вариантов низким, и таким образом сократить в этом методе элемент перебора. Для этого можно использовать различные усовершенствования нахождения шаблонов. Одно из них использует нахождение двух или группы шаблонов в рассматриваемом далее смысле.

13.6.3. Стыковка шаблонов. Если исследуются два шаблона или больше, то часто случается, что некоторые их реализации являются взаимно исключающими. Это сокращает пространство поиска. Например, в криптотексте

S E N C E I S E J P A N O A I A O P A N C A H A O A J
1 2 3 2 1 4 2 1 2 1 3 1

встречаются шаблоны 1232142 и 12131; в разд. 13.1 упоминалась реализация s e m e s t e (r) для шаблона 1232142, шаблон же 12131 допускает очень мало реализаций, и именно те, которые совместимы с e H e s e. Из списка на рис. 95 таким является лишь (g) e l e s e (n). Но существует и другая реализация для шаблона 1232142, а именно g e r e g n e (t); с этой реализацией совместимы лишь реализации для 12131, которые совместимы с e H e g e, которыми в списке на рис. 95 являются (g) e l e g e (n) и (b) e l e g e (n). $P \hat{=} \text{п}$ в реализации g e r e g n e (t) сталкивается с $J \hat{=} \text{п}$ из (g) e l e g e (n) и (b) e l e g e (n). Попытка не удалась.

Это показывает, как можно стыковать два коротких шаблона. В основе этого рассмотрения лежит нахождение объединенного шаблона

S E N C E I S E J P A N O A I A O P A N C A H A O A J
1 2 3 2 1 4 2 5 6 2 7 2 1 2 8

и его реализаций.

Возвращаясь к дешифрованию s e m e s t e r для O A I O A O P A и g e l e s e n для C A H A O A J, мы теперь в порядке рабочей гипотезы находим семь букв и следующий фрагмент:

S e r Z E m S E n t e r s t m t s t e r g e l e s e n

Для S, E и Z выбор символов открытого текста имеет $19 \cdot 18 \cdot 17 = 5814$ возможностей. Пространство поиска может быть сокращено далее исследованием около дюжины реализаций для слова S E n t e r, и для каждого из них испытанием 17 реализаций символа Z. Эти $12 \cdot 17 \approx 200$ проб с помощью компьютера проводятся всего за несколько секунд. Дешифрованное сообщение получаем в виде

W i r d i m w i n t e r s e m e s t e r g e l e s e n

Читатель, который сомневается в этом дешифровании (кроме прочего, не каждый может легко обращаться с текстом на непонятном иностранном языке) или кто думает, что текст, намного более короткий, чем расстояние единственности для одноалфавитного простого шифра, может иметь более

одного дешифрования, почувствует себя более уверенным, когда последует совету Рорбаха, и проверит, что шифрованием было сложение ЦЕЗАРЯ с ключом $22 \equiv -4 \pmod{26}$. Следует заметить, что снова в методе дешифрования не были использованы особенности сложения ЦЕЗАРЯ.

13.6.4. Сокращение пространства поиска. Можно было бы ожидать, что в одноалфавитно зашифрованном криптотексте число шаблонов определенной длины (скажем, до 15) пропорционально длине текста. Число стыковок между шаблонами, однако, растет, по крайней мере, квадратично с длиной текста, так что ограничения, возникающие от быстрого роста стыковок, сокращают пространство поиска соответствующим образом, а это означает, что найдется такая длина текста, для которой регулярно срабатывает чистый метод нахождения шаблона.

13.7. Панграммы

Частным случаем шаблонов являются такие, которые не содержат повторяющихся символов, в частности, длинных шаблонов вида $1234 \ldots \mathcal{N}$. Необходимо, чтобы $\mathcal{N}$ не превосходило мощности N словаря открытого текста. Реализации таких шаблонов называются безшаблонными текстами[4]. Эндрю привел список из около 6000 безшаблонных слов длины 6 и длины 7, 4200 длины 8, 2400 длины 9 и 1050 длины 10. Там есть еще несколько сотен безшаблонных слов длины 11, вроде

abolishment, atmospheric, comradeship, exculpation, filamentous,
hypogastric, nightwalker, questionary, slotmachine, spaceflight

и несколько дюжин слов длины 12, таких, как

ambidextrous, bakingpowder, bodysnatcher, disreputable,
housewarming, hydrosulfite, springbeauty, talcumpowder.

В списке есть даже несколько безшаблонных слов длины 13:

bowstringhemp, doubleparking, doublespacing, groupdynamics, publicservant

и одно безшаблонное слово длины 14:

ambidextrously.

Заметим, что в этих примерах пробелы между словами запрещены. Существуют, конечно, также более длинные безшаблонные предложения. В открытом тексте следует избегать безшаблонных слов или предложений достаточно большой длины $\mathcal{N}$ или подавлять их, так как они сразу подвергаются опасности дешифрования путем перебора в довольно малом пространстве поиска.

Настоящими панграммами называются предложения, содержащие каждую букву только один раз (т. е. длины $\mathcal{N} = N$). В английском языке возможны настоящие панграммы в довольно свободном стиле, например,

[4] Richard V. Andree, *Nonpattern Words of* 3 *to* 14 *Letters*, Raja Press, Norman, Oklahoma, 1982.

cwim, fjord-bank glyphs rest quiz (Боргман),
squdgy fez, blank jimp, crwth vjx (Клод Шеннон),
Zing! Vext cwim fly jabs Kurd qoph (автор неизвестен).

Хорошим приближением являются

waltz, nymph, for quick jigsvex bud (28 символов),
jackdaws ljvt my big sphinx of quartz (31 символ),
pack my box with five dozen liquor jugs (32 символа).

На немецком и французском настоящие панграммы неизвестны. Приближениями являются

sylvia wagt quick den jux bei pforzheim (33 символа, *нем.*),
bayerische jagdwitze von maxl querkopf (34 символа, *нем.*),
zwei boxkaempfer jagen eva quer durch sylt (36 символов, *нем.*),
Qui, flamboyant, guida Zéphire sur ses eaux (35 символов, Guyot, 1772, *фр.*).

Много лет используются тестовые тексты для телетайпных линий:

kaufen sie jede woche vier gute bequeme pelze (38 символов, нет x, y, *нем.*),
the quick brown fox jumps over the lazy dog (35 символов, *англ.*)
voyez le brick geant que j'examine pres du wharf (39 символов, *фр.*).

Французский язык, в частности, богат на контакты гласных, например, ouïe, и поэтому особенно подходит для панграмм-гласных, — слов, которые содержат каждую гласную лишь один раз. Хорошими примерами являются:

ultraviolet, trouvaille, autrefois, ossuaire, oripeau, ouaille,

и только с шестью буквами

oiseau.

Глава 14

Многоалфавитный случай: вероятные слова

14.1. Несовпадающий перебор позиций вероятного слова

Нахождение шаблона, использующее совпадение двух шаблонов, по необходимости ограничено лишь одноалфавитными шифрами. Но для широкого класса многоалфавитных шифров, а именно для тех, схемы шифрования которых свободны от неподвижных точек, т. е. алфавиты которых подчинены условию: «ни одна буква не может представлять себя», ни в одной букве не может быть совпадения между открытым текстом и криптотекстом. Это позволяет нам исключать некоторые позиции вероятного слова и таким образом квалифицировать оставшиеся как возможные позиции. Вероятное слово атакуется перебором позиций. Перебор проводится лишь по длине текста, и осуществим.

Предварительное условие, что ни одна буква не может представлять себя, выполняется гораздо чаще, чем может показаться сначала. Часто случается, что сам шифровальщик интуитивно избегает неподвижных точек, руководствуясь лучшими намерениями. В одноалфавитных простых подстановках это встречается регулярно, а для «аристократов» (разд. 13.3.2) даже было обязательно. Если многоалфавитные подстановки используют набор таких алфавитов, то они и сами удовлетворяют этому условию, независимо от любой иллюзорной сложности.

Кроме того, все многоалфавитные подстановки с взаимообратными алфавитами удовлетворяют этому условию. Такими будут, в частности, методы со схемами шифрования ПОРТА (разд. 7.4.4) (требующие даже, чтобы $N = |V|$), но не методы со схемами шифрования БОФОРТ (разд. 7.4.3).

Атака «несовпадающего» перебора обычно допускает несколько возможных позиций пробного вероятного слова, которые требуют переборного исследования. Если для криптосистемы Шеннона (разд. 9.1.1) алфавиты известны и если вероятное слово действительно встречается, то оно дает вход, приводящий к восстановлению части ключа. В случае ключа с принципиально известной конструкцией это как раз то, что нужно, в случае же периодического ключа раскрываются большие части открытого текста. В одноалфавитном случае, разумеется, несовпадающий перебор тоже работает.

Для фразы «Erloschen ist Leuchttonne [погасла светящаяся бочка (*нем.*)]» (разд. 11.1.3), имеется следующий криптотекст при шифровании, свободном от неподвижных точек, например, схемами шифрования ЭНИГМА, при этом возможными оказываются лишь две позиции (все остальные, приводящие к краху, отмечены жирным шрифтом):

```
YOAQUTHNCHWSYTIWHTOJQMTCFKUSLZVSMFNGTDUQNYAVH
erloschenistleuchttonne
 erloschenistleuchttonne
  erloschenistleuchttonne
   erloschenistleuchttonne
    erloschenistleuchttonne
     erloschenistleuchttonne
      erloschenistleuchttonne
       erloschenistleuchttonne
        erloschenistleuchttonne
         erloschenistleuchttonne
       → erloschenistleuchttonne
           erloschenistleuchttonne
            erloschenistleuchttonne
             erloschenistleuchttonne
              erloschenistleuchttonne
               erloschenistleuchttonne
                erloschenistleuchttonne
                 erloschenistleuchttonne
                  erloschenistleuchttonne
                   erloschenistleuchttonne
                    erloschenistleuchttonne
                     erloschenistleuchttonne
                   → erloschenistleuchttonne
```

Многие механические криптосистемы чувствительны к этой атаке перебора. Прежде всего, это ленточные и цилиндрические мультиплексные шифры (разд. 7.5.3), чьи нетождественные алфавиты являются одноцикловыми подстановками (единственная подстановка, нарушающая это условие, т. е. тождественная, заведомо исключена).

Во вторых, существуют машины, работающие «для облегчения» взаимообратно. Тогда каждый алфавит является взаимообратным. А если 1-циклы (подстановочные циклы единичной длины) исключаются по техническим причинам, то предварительное условие «ни одна буква не может представлять себя» выполнено.

Если машина Хагелина, работая с механическими схемами шифрования БОФОРТ, допускала 1-циклы, то электрическая ЭНИГМА их не допускала, и страдала от этой слабости, которая была следствием введения отражающего ротора — считалось, что это особое усовершенствование. Трудно поверить, что Группа IV Шифровального отдела OKW (Хюттенхайн), следящая за надежностью своих собственных систем, не знала об этой возможности для взлома, но вероятно они недооценивали ее. Как бы то ни было, это оказалось большой удачей для работы Польского бюро шифров и английских дешифровальщиков из Блэтчли Парк. Но ленточный шифр CSP-642, используемый столь же беззаботно в американском флоте, тоже был многоалфавитным со многими смешанными алфавитами и тоже был под угрозой; японцы его взломали после того, как они захватили это ленточное устройство на островах Вейк и Киска.

Короткие вероятные слова допускают, конечно, много «попаданий в цель», но также много и «промахов». Полная вероятность попадания для вероятного слова длины n выражается формулой

$$P_n = (1 - N^{-1})^n \asymp e^{-n/N};$$

число n должно быть достаточно большим, чтобы сделать P_n хоть немного меньше, чем частота вероятного слова. Рисунок 96 дает некоторые значения P_n для обычного случая $N = 26$.

n	P_n[%]	n	P_n[%]
1	96.15	12	62.45
2	92.45	16	53.39
3	88.90	24	39.01
4	85.48	32	28.51
5	82.19	48	15.22
6	79.03	64	8.13
8	73.07	100	1.98
10	67.56	128	0.66

Рис. 97. Полная вероятность попаданий для несовпадающих переборов ($N = 26$)

Следующий открытый текст длины 60 (Конхейм) зашифрован машиной ЭНИГМА:

```
m a n y o   r g a n i   z a t i o   n s r e l   y o n c o   m p u t e   r s a
G R S U Z   T L D S Z   N K W N E   R D P F B   O V V Q N   O B K Y I   Q N J
```

Несовпадающий перебор для вероятного слова /computer/ из 8 символов дает по указанной формуле для первых 26 позиций около 19 возможных пози-

ций ($26 \cdot 0.7307$). Фактически же 21 из этих позиций не может быть исключена, и, таким образом, имеется 20 промахов.

```
G R S U Z T L D S Z N K W N E R D P F B O V V Q N O B K Y I Q N J
  c o m p u t e r
→ c o m p u t e r
  → c o m p u t e r
    → c o m p u t e r
      → c o m p u t e r
        → c o m p u t e r
          → c o m p u t e r
            → c o m p u t e r
              c o m p u t e r
            → c o m p u t e r
              → c o m p u t e r
                → c o m p u t e r
                  → c o m p u t e r
                    → c o m p u t e r
                      c o m p u t e r
                    → c o m p u t e r
                      → c o m p u t e r
                        → c o m p u t e r
                          → c o m p u t e r
                            → c o m p u t e r
                              → c o m p u t e r
                                c o m p u t e r
                              → c o m p u t e r
                                → c o m p u t e r
                                  → c o m p u t e r
                                    → c o m p u t e r
                                      c o m p u t e r
                                    → c o m p u t e r
```

Дело в том, что /computer/ — очень короткое слово. Для 24-символьного слова /oberkommandoderwehrmacht/ было бы исключено намного больше позиций. Подсчет дает для 36 позиций этого слова 14 возможных позиций (причем, все являются промахами); сравнивая с ожидаемыми $36 \cdot 0.3901 = 14.04$, получаем почти полное совпадение. Вероятное слово, длина которого заметно превосходит 100, имеет хороший шанс обойтись без промахов в тексте из 300 символов.

14.2. Бинарный несовпадающий перебор позиции вероятного слова

Бóльшая информация о шифрах многоалфавитной криптосистемы может улучшить ситуацию с промахами. Чтобы привести пример, возьмем многоалфавитную схему шифрования ПОРТА, отображающую одну половину алфа-

вита в другую половину и наоборот (разд. 7.4.4), скажем, с $V = Z_{26}$

$$\{\text{abcd} \ldots \text{lm}\} \overset{2}{\longleftrightarrow} \{\text{nopq} \ldots \text{yz}\}.$$

Бинарный шаблон текста получается, если каждый символ заменяется на 0 или 1 — в зависимости от того, в первую или вторую половину текста входит этот символ.

Теперь предположим, что тот же, что и выше, текст каким-нибудь образом зашифрован шифрованием ПОРТА:

```
m a n y o  r g a n i  z a t i o  n s r e l  y o n c o  m p u t e  r s a
P R G B F  I O Z G P  L Y C N E  E D G W S  A I K Q D  O B K J Q  C M P
```

Вероятное слово /computer/ имеет бинарный шаблон 01011101. Проводя несовпадающий бинарный перебор криптотекста с этим бинарным шаблоном и выделяя жирным шрифтом позиции несовпадения, получаем

```
P R G B F I O Z G P L Y C N E E D G W S A I K Q D O B K J Q C M P
1 1 0 0 0 0 1 1 0 1 0 1 0 1 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0 1 0 0 1
```

```
0 1 0 1 1 1 0 1
  0 1 0 1 1 1 0 1
    0 1 0 1 1 1 0 1
      0 1 0 1 1 1 0 1
        0 1 0 1 1 1 0 1
          0 1 0 1 1 1 0 1
            0 1 0 1 1 1 0 1
              0 1 0 1 1 1 0 1
                0 1 0 1 1 1 0 1
                  0 1 0 1 1 1 0 1
                    0 1 0 1 1 1 0 1
                      0 1 0 1 1 1 0 1
                        0 1 0 1 1 1 0 1
                          0 1 0 1 1 1 0 1
                            0 1 0 1 1 1 0 1
                              0 1 0 1 1 1 0 1
                                0 1 0 1 1 1 0 1
                                  0 1 0 1 1 1 0 1
                                    0 1 0 1 1 1 0 1
                                      0 1 0 1 1 1 0 1
                                        0 1 0 1 1 1 0 1
                                          → 0 1 0 1 1 1 0 1
                                              0 1 0 1 1 1 0 1
                                                0 1 0 1 1 1 0 1
```

В этом примере даже с довольно коротким вероятным словом несовпадающий бинарный перебор не дал ни одного промаха, тогда как обычный несовпадающий перебор не исключил 21 из этих позиций, и таким образом сделал 20 промахов. Даже более короткое вероятное слово /comp/ с бинарным шаблоном 0101 имеет лишь три промаха; — сравните с ожидаемыми $26 \cdot (1/2)^4 = 1.62$ возможными позициями. Для шифрования ПОРТА вероятные слова даже с небольшим числом символов, вообще говоря, уже достаточны для различения истинного попадания и полного промаха.

Слово /oberkommandoderwehrmacht/ содержит 11-символьное слово /kommandoder/ с бинарным шаблоном 01000101001. Его нельзя обнаружить в открытом тексте данного примера, так как каждая позиция приводит к краху:

```
P R G B F I O Z G P L Y C N E E D G W S A I K Q D O B K J Q C M P
1 1 0 0 0 0 1 1 0 1 0 1 0 1 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0 1 0 0 1

0 1 0 0 0 1 0 1 0 0 1
  0 1 0 0 0 1 0 1 0 0 1
    0 1 0 0 0 1 0 1 0 0 1
      0 1 0 0 0 1 0 1 0 0 1
        0 1 0 0 0 1 0 1 0 0 1
          0 1 0 0 0 1 0 1 0 0 1
            0 1 0 0 0 1 0 1 0 0 1
              0 1 0 0 0 1 0 1 0 0 1
                0 1 0 0 0 1 0 1 0 0 1
                  0 1 0 0 0 1 0 1 0 0 1
                    0 1 0 0 0 1 0 1 0 0 1
                      0 1 0 0 0 1 0 1 0 0 1
                        0 1 0 0 0 1 0 1 0 0 1
                          0 1 0 0 0 1 0 1 0 0 1
                            0 1 0 0 0 1 0 1 0 0 1
                              0 1 0 0 0 1 0 1 0 0 1
                                0 1 0 0 0 1 0 1 0 0 1
                                  0 1 0 0 0 1 0 1 0 0 1
                                    0 1 0 0 0 1 0 1 0 0 1
                                      0 1 0 0 0 1 0 1 0 0 1
                                        0 1 0 0 0 1 0 1 0 0 1
```

14.3. Атака Де Виари

Даже против ленточных и цилиндрических мультиплексных шифров, работая с одноцикловыми неродственными алфавитами, допускающими несовпадающий перебор в соответствии с разд. 14.1, незаконный дешифровальщик может достичь большего. Это открытие было, к несчастью, сделано Этьеном Базерье, великим практиком криптоанализа, когда его, якобы невзламывае-

мое устройство было осмеяно его оппонентом Де Виари. Так случилось, что атака Де Виари в 1893 г. и атака Фридмана в 1918 г. не предполагали, что алфавиты были одноцикловыми.

14.3.1. Но даже общий метод предполагает, что шифрующее устройство доступно незаконному дешифровальщику. Поскольку ничто не препятствует использованию не менее двух дюжин дисков или лент, то можно предположить, что число перестановок всегда достаточно велико, чтобы исключить тривиальный перебор. Максимальный период определяется системой, и известные алфавиты можно тестировать по столбцам односимвольно зашифрованных символов (разд. 18.2.5); но глубина материала, как правило, недостаточна для успешного применения частотного анализа.

Для частного случая ленточных и цилиндрических устройств существует кажущееся усложнение, вызываемое омофонией. Оказывается, что омофония мешает незаконному дешифровальщику не больше, чем законному. Допустим на минуту, что какой-то фрагмент (омофонного) криптотекста читается с k-й строки под строкой открытого текста (в k-й образующей (generatrix), разд. 7.5.3). Начиная с малых значений k, потребовалось бы, самое худшее, две дюжины вариантов проб и ошибок (по числу образующих). Допустим также для простоты, что вероятное слово будет достаточно коротким, чтобы не быть обрезанной периодической сменой шифрования.

Тогда для фиксированного k и данного вероятного слова открытого текста мы определяем множество всех символов, встречающихся на дисках или полосах k-й образующей. С этой основной информацией мы исследуем все позиции вероятного слова, чтобы обнаружить, для каких из них криптотекст вообще мог быть получен. Для короткого вероятного слова мы ожидаем несколько возможностей доведения до конца. Если же вероятное слово достаточно длинное, то может случиться, что не будет найдено ни одной возможности. В таком случае переходим к другой отмеченной образующей. Если это оказывается безуспешным для каждой образующей, значит, вероятное слово, вопреки нашему предположению, не представлено в открытом тексте (это может также означать, что оно было оборвано).

Для цилиндра Базерье с 20 дисками и криптотекста Живержа:

```
F S A M C  R D N F E  Y H L O E  R T X V Z
L R M Q U  U X R G Z  N B O M L  N D N P V
R T M U K  H R D O X  L A X O D  C R E E H
V R E X Z  G U G L A  B S E S T  V F N G H
```

относящегося к военной тематике, Де Виари начинает атаку следующим образом.

Допустим, что вероятным словом является /division/. Для 20 циклов Базерье (рис. 56) рис. 97 а показывает зашифрования слова /division/ для первой образующей (первой строки рис. 56). Таким образом, множества встречающихся криптосимволов должны читаться вертикально под символами открытого текста /d/, /i/, /v/, ... слова /division/.

(a)	d	i	v	i	s	i	o	n
1	E	J	X	J	T	J	P	O
2	**F**	O	X	O	T	O	U	P
3	**F**	O	X	O	T	O	J	P
4	C	H	U	H	R	H	N	M
5	C	Q	T	Q	R	Q	I	M
6	C	E	T	E	R	E	I	M
7	P	J	B	J	E	J	N	S
8	T	E	D	E	P	E	Y	H
9	E	J	X	J	L	J	P	O
10	I	E	X	E	V	E	T	C
11	B	T	I	T	**C**	T	U	D
12	G	C	Y	C	A	C	R	P
13	N	R	Y	R	A	**R**	I	S
14	**F**	B	X	B	V	B	N	E
15	**F**	N	X	N	T	N	P	S
16	K	M	X	**M**	V	M	G	F
17	**F**	E	X	E	O	E	N	A
18	I	U	X	U	T	U	M	Q
19	**F**	J	L	J	U	J	N	T
20	G	J	X	J	T	J	U	O

(b)	d	i	v	i	s	i	o	n
1	**H**	M	A	M	X	A	S	R
2	J	B	E	B	Z	B	E	S
3	I	**L**	E	L	Z	L	M	Q
4	Z	E	R	**E**	O	E	K	J
5	U	M	**O**	M	Q	M	N	E
6	U	X	Q	X	N	X	Z	J
7	J	V	K	V	D	V	F	T
8	**M**	U	J	U	D	U	F	X
9	L	N	H	N	I	N	V	U
10	P	R	D	R	Z	R	A	J
11	**H**	S	L	S	**R**	S	N	G
12	K	B	R	B	U	B	Z	**V**
13	U	T	L	T	B	**T**	M	X
14	K	F	H	F	Z	F	R	T
15	K	R	N	R	E	R	**X**	U
16	S	O	J	O	Z	O	R	H
17	J	O	Y	O	B	O	C	D
18	B	C	L	C	U	C	R	Y
19	J	Q	F	Q	M	Q	Z	A
20	J	P	N	P	A	P	E	T

Рис. 98. Зашифрования слова /division/, (a) первой образующей (с **FSAMCR**DN); (b) 4-й образующей (с **HLOERTXV**)

Скользя по бумажной ленте вдоль этих множеств, мы можем для каждой позиции вероятного слова решить, будут ли все соответствующие буквы криптотекста находиться среди имеющихся. Например, это не выполняется для следующей позиции с фрагментом FSAMCRDN:

d i v i s i o n
F S A **M C** **R** D N F E Y H L O E R T X V Z,

где находятся лишь четыре буквы, выделенные жирным шрифтом (а не восемь). То же самое имеет место и для следующей позиции с фрагментом SAMCRDNF:

d i v i s i o n
F S A **M C** **R** D **N** **F** E Y H L O E R T X V Z,

где снова находятся лишь 4 буквы (вместо 8), выделенные жирным шрифтом.

Для следующей позиции с фрагментом AMCRDNFE,

d i v i s i o n
F S A **M C** **R** D **N** **F** **E** Y H L O E R T X V Z

снова получается промах. Продолжая таким образом, мы вообще исключаем первую образующую. На рис. 97 b показаны зашифрования слова /division/ для четвертой образующей. Снова множество встречающихся криптосимволов должно быть прочитано вертикально под символами /d/, /i/, /v/, ... слова /division/. Снова начиная слева, мы, наконец, впервые получаем попадание на 12-й позиции с фрагментом **HLOERTXV**:

```
                    d i v i s i o n
F S A M C R D N F E Y H L O E R T X V Z
```

Все 8 букв (выделенные жирным шрифтом) находятся среди имеющихся, причем, семь из них только по одному разу; они и определяют соответствующий алфавит. Однако для **H** существует выбор между первым и одиннадцатым алфавитами, как показывает рис. 97 b.

14.3.2. В данный момент было бы полезно дать о системе некоторые дополнительные сведения. В принципе, любой алфавит, родственный или неродственный, можно было бы использовать несколько раз. Для схем шифрования ВИЖЕНЕР это было бы вполне нормальным. Последовательное же шифрование (разд. 8.4.2) ограничивает этот произвол, чтобы предотвратить накопление материала зашифрованного одним и тем же алфавитом. Для цилиндрических и ленточных устройств последовательное шифрование является систематическим; кажется, чтобы увеличить надежность, надо воспользоваться каждым цилиндром или лентой только раз, чтобы таким образом использовать его на протяжении периода в точности один раз. Но как только эти алфавиты попадают в руки врага, это действительно становится *иллюзорной сложностью*.

В предположении последовательности шифрования, которое справедливо для цилиндров Базерье, равенство $H \hat{=} \mathrm{d}$ исключает одиннадцатый алфавит, так как для него необходимая единственность уже исчерпана равенством $R \hat{=} \mathrm{s}$. Таким образом, пока порядок цилиндров частично определен следующим образом:

```
* * * * *  * * * * *  * 1 3 5 4  11 13 15 12 *
```

К тому же, чтобы исключить возможность промаха, мы проверим, появятся ли на расстоянии двадцати символов значительные результаты дешифрования. Для фрагмента BOMLNDNV результат

```
L R M Q U  U X R G Z  N B O M L  N D N P V
                     1 3 5 4  11 13 15 12
                     z h p n  r  m  y  k      24.
                     a i n m  a  t  i  n   → 0.
                     B O M L  N  D  N  P      1.
                     c j l k  d  n  s  q      2.
                     d k k j  b  s  t  t      3.
```

показывает уверенное дешифрование /ainmatin/; для первого раунда была использована первая образующая. Далее, на расстоянии в 40 символов для фрагмента AXODGREE результат

```
R T M U K  H R D O X  L A X O D  C R E E H
                     1 3 5 4  11 13 15 12
                     A X O D  C R E E      22.
                     b z i c  o e z z      23.
                     c a q b  u m l l      24.
                     d e p a  r t a s  →   0.
                     e b n z  a d j a       1.
```

дает дешифрование /departas/: для второго раунда использовалась 22-я образующая. Фрагмент открытого текста /departas/ можно дополнить двумя способами: до слов /departasixheueres/ или /departaseptheueres/. Учитывая тот факт, что 6 часов — это довольно рано, мы попробуем /departaseptheuer/ в качестве следующего вероятного слова, которое разрежем на /departase/ и /ptheures/. Последнее из них испытывается на рис. 98; для третьей образующей выставлены зашифрования слова /ptheures/. В следующей позиции криптотекста, с фрагментом VREXZGUG (шансы на успех 1 : 206) имеет место настоящее попадание:

```
p t h e u r e s
V R E X Z  G U G L A  B S E S T  V F N G H,
```

которое определяет положение еще четырех, пока не рассматривавшихся цилиндров: R $\hat{=}$ t требует 7-го, X $\hat{=}$ e — 6-го, G $\hat{=}$ r — 10-го, G $\hat{=}$ s — 9-го цилиндра. На рис. 98 цилиндры, определенные к настоящему моменту, отмечены. Из оставшихся соответствий Z $\hat{=}$ u требует однозначно 17-цилиндра, тогда как три случая остаются открытыми: V $\hat{=}$ p требует 16-го или 20-го цилиндра, E $\hat{=}$ h — 14-го или 18-го, U $\hat{=}$ e — 2-го или 8-го цилиндров.

Результатом является следующее распределение девятнадцати из общего числа 20 цилиндров:

$$\begin{matrix}16\\20\end{matrix}\ 7\ \begin{matrix}14\\18\end{matrix}\ 6\ 17\quad 10\ \begin{matrix}2\\8\end{matrix}\ 9\ **\quad *\ 1\ 3\ 5\ 4\quad 11\ 13\ 15\ 12\ *$$

Остающееся дешифрование — пустячное дело: с 13 цилиндрами, положение которых выявлено к настоящему времени, получаем следующее фрагментарное дешифрование:

```
F S A M C  R D N F E  Y H L O E  R T X V Z
* a * r o  i * i * *  * d i v i  s i o n *
L R M Q U  U X R G Z  N B O M L  N D N P V
* p * r t  e * a * *  * a i n m  a t i n *
R T M U K  H R D O X  L A X O D  C R E E H
* r * e i  m * s * *  * d e p a  r t a s e
V R E X Z  G U G L A  B S E S T  V F N G H
p t h e u  r e s * *  * p x x x  x x x x *
```

	p	t	h	e	u	r	e	s
• 1	S	X	K	H	Y	U	H	V
2	S	Z	L	U	C	V	**U**	X
• 3	Q	Z	J	D	R	V	D	X
• 4	M	Q	**E**	B	R	O	B	P
• 5	L	O	D	H	V	Q	H	I
• 6	L	Q	D	**X**	E	N	X	P
• 7	J	**R**	Q	D	B	U	D	T
8	D	M	X	U	L	S	**U**	V
• 9	**V**	F	Y	L	**Z**	D	L	**G**
•10	T	A	M	R	O	**G**	R	Y
•11	Y	S	M	T	N	D	T	U
•12	**V**	F	N	S	D	Z	S	C
•13	Y	S	P	D	C	T	D	X
14	B	I	**E**	T	P	A	T	Y
•15	X	E	O	A	L	Z	A	U
16	**V**	B	C	G	D	U	G	Y
17	U	X	P	O	**Z**	L	O	A
18	T	U	**E**	S	C	D	S	I
19	S	A	B	C	M	X	C	Y
20	**V**	A	K	C	E	Y	C	L

Полная вероятность попадания

$$\frac{12}{25}\times\frac{13}{25}\times\frac{14}{25}\times\frac{13}{25}\times\frac{13}{25}\times\frac{14}{25}\times\frac{13}{25}\times\frac{11}{25}\approx$$

$$\approx 1:206$$

$c \mapsto \{A, B, C, D, G, H, L, O, R, S, T, U, X\}$

$h \mapsto \{B, C, D, E, J, K, L, M, N, O, P, Q, X, Y\}$

$p \mapsto \{B, D, J, L, M, Q, S, T, U, V, X, Y\}$

$r \mapsto \{A, D, G, L, N, O, Q, S, T, U, V, X, Y, Z\}$

$s \mapsto \{A, C, G, I, L, P, T, U, V, X, Y\}$

$t \mapsto \{A, B, E, F, I, M, O, Q, R, S, U, X, Z\}$

$u \mapsto \{A, B, E, F, I, M, O, Q, R, S, U, X, Z\}$

Рис. 99. Зашифрования /ptheures/, 3-я образующая (с **VREXZGUG**)

которое сразу наводит на два дальнейших фрагмента: /troisieme/ и /demain/, позволяющие нам дополнить все позиции, кроме 20-й, а после этого и 20-ю позицию тоже. Теперь можно восстановить полный порядок цилиндров.

16 7 18 6 17 10 8 9 20 19 2 1 3 5 4 11 13 15 12 14

Полное дешифрование имеет вид (заметно заполнение конца пустышками):

F S A M C R D N F E Y H L O E R T X V Z
l a t r o i s i e m e d i v i s i o n s
L R M Q U U X R G Z N B O M L N D N P V
e p o r t e r a d e m a i n m a t i n s
R T M U K H R D O X L A X O D C R E E H
u r r e i m s s t o p d e p a r t a s e
V R E X Z G U G L A B S E S T V F N G H
p t h e u r e s t t o p x x x x x x x x

14.3.3. Даже если вероятное слово отсутствует, атака Де Виари может работать. Следуя Живаржу (1925 г.), можно использовать частоты биграмм, триграмм и тетраграмм. Покажем это на стандартном английском и французском окончании /ation/. Для каждой образующей, для каждой буквы открытого текста заранее изготавливаются возможные буквы криптотекста. На

рис. 99 это показано для первой образующей. Так как эти соответствующие множества заключают в себе, грубо говоря, лишь половину букв, то опасность промаха снова не слишком велика

	a	t	i	o	n
1	B	U	J	P	O
2	E	V	O	U	P
3	E	V	O	J	P
4	Z	S	H	N	M
5	J	S	Q	I	M
6	Z	S	E	I	M
7	L	D	J	N	S
8	V	O	E	Y	H
9	R	S	J	P	O
10	F	G	E	T	C
11	N	Z	T	U	D
12	I	V	C	R	P
13	U	D	R	I	S
14	I	P	B	N	E
15	J	R	N	P	S
16	I	H	M	G	F
17	B	U	E	N	A
18	B	D	U	M	Q
19	C	E	J	N	T
20	C	E	J	U	O

Полная вероятность попадания

$$\frac{12}{25} \times \frac{11}{25} \times \frac{12}{25} \times \frac{10}{25} \times \frac{12}{25} \approx$$

$$\approx 1:51$$

$a \mapsto \{B, C, E, F, I, J, L, N, R, U, V, Z\}$

$t \mapsto \{D, E, G, H, O, P, R, S, U, V, Z\}$

$i \mapsto \{B, C, E, H, J, M, N, O, Q, R, T, U\}$

$o \mapsto \{G, I, J, M, N, P, R, T, U, Y\}$

$n \mapsto \{A, C, D, E, F, H, M, O, P, Q, S, T\}$

Рис. 100. Зашифрования слова /ation/, первая образующая

Можно сказать, что атаки Де Виари и Фридмана и, в частности, вариант Живержа, пытаются обнаружить много маленьких островов, которые затем могут быть увеличены в архипелаги, которые, в свою очередь, могут слиться в континенты, и так далее.

14.3.4. Как отмечалось выше, общая атака Де Виари заранее не предполагает, что алфавиты должны быть одноцикловыми. Мы можем теперь ясно видеть, что как бинарные несовпадающие переборы (разд. 14.2), так и несовпадающие переборы (разд. 14.1) являются частными случаями, в которых множества символов криптотекста, соответствующих символам открытого текста, формируются систематически. Для того, чтобы исключить промахи, дешифровальщику лучше брать более мелкие множества соответствующих криптосимволов. С другой стороны, по этой причине общая атака Де Виари терпела неудачу, если каждое из этих соответствующих множеств являлось полным словарем криптотекста. Криптосистему с таким защитным свойством мы будем называть транзитивной. В этом случае необходимо, чтобы число алфавитов было больше или равно N. Цилиндр Базерье с 20 дисками нарушает это условие. Американское устройство М-138-А применяло 30 лент, и можно

было ожидать, что алфавиты всегда выбирались из заданных (от 50 до 100) транзитивных криптосистем.

Если число алфавитов равно N, то для транзитивной криптосистемы МУЛЬТИПЛЕКС каждый алфавит с N символами образует латинский квадрат (разд. 7.5.4). У аппарата М-94 с 25 дисками почти все алфавиты построены таким образом (табл. 2). Это эквивалентно тому, что для каждой пары символов ОТ и КТ (открытого и криптотекстов) определен (и притом однозначно) соответствующий алфавит. Но это в точности условие, характеризующее криптосистему Шеннона (см. 2.6.4).

14.3.5. Французский маркиз Де Виари родился 13 февраля 1847 г. в Шербуре в семье капитана артиллерии. Девятнадцати лет Виари поступил в Политехническую школу, в 21 год стал моряком, позднее он был префектом полиции и, наконец, пехотным офицером. Его интерес к криптологии возник примерно в 1885 г., он первым заслужил репутацию изобретателя шифровальной машины. Он же был первым после Бэббиджа, кто использовал математику в криптологии, а именно, когда рассматривал линейные подстановки в ряде статей в 1888 г. В 1893 г. он написал криптоаналитическое эссе «Искусство шифрования и дешифрования секретных сообщений», которое сделало его известным. В 1898 г. он опубликовал также один коммерческий код. Умер Де Виари 18 февраля 1901 г.

Живерж был майором и помощником полковника Картье в начале 1914 г. Когда разразилась Первая мировая война, он создал во французском генеральном штабе дешифровальное бюро. В 1925 г., когда была опубликована его книга «Курс криптографии», он был уже полковником; позднее был повышен в звании до генерала. Не без гордости он отмечал, что одно «mot probable» [вероятное слово (*фр.*)], изобретенное им, дороже квинтиллиона проб.

Под руководством Живержа работал майор Пэнви, гениальный дешифровальщик, который изучал палеонтологию и впоследствии стал крупным промышленным и финансовым деятелем.

14.3.6. Один из немногих в XX в. случаев открытого сообщения об успешном профессиональном криптоанализе произошел благодаря окончанию Второй мировой войны, когда был написан «FIAT Обзор немецкой науки». В серии приложений математики Рорбах поместил сообщение по криптологии, куда вошли детали взлома «Американского ленточного шифра О-2» (так Рорбах назвал вариант шифра М-138-А используемого госдепартаментом США), выполненного Германской Sonderdienst Dahlem [Специальная служба Далема (*нем.*)] Министерства иностранных дел. В 1979 г. это обстоятельное сообщение, написанное во второй половине 1945 г., было, наконец, опубликовано. Оно описывало работу математиков Кунце, Рорбаха, Хюнке, Паннвиц, Круга, Грунски и Шульца — не следует забывать и лингвистов Мюллера, Фридрихс, Цигенрюкера и Дойбнера.

Эта работа началась в ноябре 1943 г. со сбора и сортировки большого количества криптотекстов, главным образом направленных посольству США в Берне (Швейцария) или отправленных оттуда. Там располагался офис стра-

тегической службы OSS Аллена В. Даллеса, шефа американской шпионской сети OSS (SI) в Европе. Эта работа выявила:

1) частые параллели, в том числе довольно длинные, но никогда не длиннее тридцати символов, чаще длиной 15,

2) частые параллели между сообщениями одного и того же дня — но никогда между сообщениями разных дней одного и того же месяца,

3) никаких параллелей между двумя сообщениями, если одно было раньше, а другое позже 1 августа 1942 г.

Были сделаны выводы, что после 15 символов (иногда после 30) делалось изменение в шифровании, что пароль менялся каждый день и что 1 августа 1942 г. было сделано фундаментальное изменение в криптосистеме. Это позволило выдвинуть рабочую гипотезу о многоалфавитном олносимвольном шифровании периода 15. Используемая шифросистема была неизвестна Рорбаху, но было известно, что американские криптологи предпочитают цилиндрические и ленточные шифры. Однако даже это не слишком помогало: у Рорбаха не было алфавитов. Поэтому нельзя было начать с прямой атаки Де Виари. Дальнейшее изучение с использованием перфокарточных машин Холлерита показало, что

4) если сообщения разбиты на блоки (Zielen [строки (*нем.*)]) по 15 символов, то все повторения по крайней мере 8 символов появляются вертикально в тех же столбцах.

Это подтверждало предположение о многоалфавитном односимвольном шифровании периода 15; более того, из стереотипных повторений («От Мэрфи», «Строго секретно») в начале сообщения можно было вывести, что ни одна буква не представляла себя и что один и тот же открытый текст в одной и той же позиции должен давать криптотексты без совпадений. Все это приводило к подозрению о полифоническом шифровании с одноцикловыми алфавитами, выполненном при помощи цилиндрического или ленточного шифра, без использования какой-либо машины. Таким образом, для каждого дня должны были определяться 15 или 30 алфавитов.

Но «наследство» было богатым, в среднем около 15 сообщений в день, каждое с 40 блоками из 15 символов. Эти блоки надо было сгруппировать в «семейства», зашифрованные одним и тем же множеством алфавитов, предположительно отобранных из некоторого запаса, содержащего более чем 15 символов, в одном и том же порядке. Кроме того, блоки должны были группироваться в классы, соответствующие образующей, которой они принадлежат. Как только блоки были сгруппированы требуемым образом, криптоаналитики получили уже одноалфавитные шифры: работа проводилась с применением перфокарточных машин и оборудования, построенного Кругом; в работе использовался Хи тест (гл. 16; разд. 18.3). Наиболее объемный класс («Класс III») окончательно состоял из 3000 блоков, сгруппированных в 25 семейств — из 60 и 150 блоков. Для окончательного восстановления алфавитов они использовали фрагменты вероятных слов /tion/ и /ation/, подкрепленные тройками биграмм /in/, /an/, /on/ и /in/, /an/, /un/. Рорбах живо описал этапы этой работы. Спустя примерно один год дело продолжилось под его

собственным руководством. Сначала был полностью определен Класс III и его 2×15 лент. Некоторые из этих лент встречались также в других классах, например, лента 18 в Классе I. В конце работы группа Рорбаха обнаружила, что всего было использовано 50 лент, — и сегодня мы знаем, что это соответствует действительности. Далее классический метод Де Виари позволил определить выбор и порядок лент, соответствующих ежедневным паролям, 40 из которых были идентифицированы.

Таким образом, все сообщения, зашифрованные О-2, могли быть прочитаны. К несчастью для германской стороны и к счастью для англо-американской, вскоре после того, как стало возможным полное использование результатов взлома шифра О-2, госдепартамент США в середине 1944 г. изменил этот шифр на более современные и надежные машины SIGTOT с индивидуальными ключами, которыми была снабжена армия США. Итак, к сентябрю 1944 г. источник высох. Кроме того, производительность группы Sonderdienst Dahlem страдала от частых бомбардировок союзников. К тому же русская армия все ближе и ближе подходила к месту их эвакуации в Силезии. К концу войны группа переместилась в Тюрингию, а затем в Марбург. Между тем, геринговское агентство прослушивания телефонных разговоров (Служба исследований министерства авиации) находилось не в лучших условиях.

Японцы тоже пытались взломать шифр CSP-642, но без особого успеха. Фридман усилил его против атак Де Виари и Живержа, после чего этот шифр можно было сломать только несовпадающим перебором позиции вероятного слова. Насколько успешно это делали русские — неизвестно.

14.4. Перебор зигзагом позиции вероятного слова

Некоторые методы используют вероятное слово для восстановления ключа, что возможно, разумеется, лишь тогда, когда открытый текст и криптотекст определяют ключ однозначно. Это тривиально при одноалфавитном шифровании даже в случае многосимвольного шифра большой ширины. Но это также справедливо и для многоалфавитного шифрования, удовлетворяющего условию Шеннона (разд. 9.1.1). Наиболее известными представителями являются все линейные подстановки, где ключ — периодический он или нет — можно просто вычислить вычитанием, если известен алфавитный порядок. Эта возможность была изучена Бэббиджем в 1846 г. для шифров ВИЖЕНЕР и БОФОРТ. Среди нелинейных подстановок, удовлетворяющих условию Шеннона, шифры АЛЬБЕРТИ и ПОРТА не приводят к осложнениям, если известны упомянутые алфавиты.

Вдвойне опасно использование осмысленного ключевого текста. Если для многоалфавитной системы Шеннона — периодической или нет — схемы шифрования известны, то возможными позициями вероятного слова в открытом тексте являются те, которые дают осмысленные фрагменты ключевого текста; они могут быть определены перебором.

В этом случае, однако, можно также поменять местами роли открытого и ключевого текстов. В осмысленном ключевом тексте скорее всего также

существует вероятное слово, которое приводит к фрагменту осмысленного открытого текста.

Например, допустим, что криптотекст

B A W I S M E W O O P G V R S F I B B T J T W L H W W A H T M J V B

зашифрован над $\mathbb{Z}_{26}$ схемой шифрования ВИЖЕНЕР. В качестве вероятного слова для ключевого текста выберем частое слово THAT, которое находится на 7-м месте в частотном списке английского языка. Перебор позиций этого слова в ключевом тексте дает следующие фрагменты открытого текста:

itwp, hpiz, dbst, plml, zfcd, txwv, lpovdhow, vhpu, ... , dtha, hatt,

среди которых многообещающими выглядят восьмой, предпоследний и последний. Угадав теперь, что dhow может быть продолжено как dhowever, находим в продолженном ключе фрагмент *THATCANB*. Последние два фрагмента dtha и hatt являются взаимно исключающими, тем не менее получаем, что dtha — правильный фрагмент. Со стороны открытого текста тоже можно угадывать; например, should приводит к

$$JTIOHJ,\ IPUYBB,\ EBESTT,\ QLYKLL,\ \ldots$$

где попаданием является третья позиция. Таким образом, should и dhowever пересекаются, образуя shouldhowever, что соответствует *EBESTTHATCANB*. Расширяя ключевой фрагмент до *THEBESTTHATCANBE*, мы получаем расширение открытого текста, которое является попаданием и читается как itshouldhoweverb. Последнее в свою очередь приводит к новому расширению itshouldhoweverbe и так далее.

Такое взаимодействие «зигзагом» часто приводит в результате к полному дешифрованию (Фридман, 1918 г.); в данном примере открытый и ключевой тексты читаются следующим образом (см. цитаты во введении к части II):

i t s h o u l d h o w e v e r b e e m p h a s i z e d t h a t c r y
T H E B E S T T H A T C A N B E E X P E C T E D I S T H A T T H E D

Непериодический ключевой текст не мешает зигзагообразному перебору; важно, чтобы как ключевой, так и открытый тексты безусловно имели более чем 50-процентную избыточность.

14.5. Метод изоморфов

Недостатком шифрований, основанных на вращаемых алфавитах, является также то, что эти алфавиты не обязательно формируют латинские квадраты. Это обстоятельство открывает широкий путь для атаки.

14.5.1. Кнокс и Кандела. Это можно продемонстрировать с помощью метода изоморфов для взлома шифра РОТОР. Метод применялся уже в 1935 г., если не раньше, Кноксом для взлома итальянской коммерческой машины

ЭНИГМА, а позднее — против Франко в Испании, но был описан в открытой литературе лишь в 1946 г. Канделой. В случае коммерческой машины ЭНИГМА без штепсельного коммутатора метод изоморфов назывался méthod des bâtons [метод палок (*фр.*)], или методом стержня (или «группы на стержнях»); его появление было главной причиной введения (еще в 1930 г.!) штепсельных коммутаторов для машины ЭНИГМА Вермахта. В гражданской войне 1938–1939 гг. в Испании стороны использовали коммерческие машины ЭНИГМА, и метод изоморфов служил английским, немецким, итальянским и испанским республиканским криптоаналитическим службам.

Допустим, что многоалфавитная подстановка имеет вид

$$c_i = p_i S_i U S_i^{-1},$$

где p_i — символ открытого текста, c_i — символ криптотекста, S_i — известный алфавит, порядок которого тоже известен. Неизвестным ключом является начальный индекс последовательности (S_i) и, возможно, подстановка U. Последовательности $p_i S_i$ и $c_i S_i$ изоморфны (разд. 2.6.3), так как

$$c_i S_i = p_i S_i \cdot U,$$

т. е. последовательность $(c_i S_i)$ является одноалфавитным образом последовательности $(p_i S_i)$ при отображении U.

Вообще говоря, две произвольно выбранных последовательности неизоморфны. Например, последовательность (a l l e...) неизоморфна последовательности (g a n g...), так как пары (l, a) и (l, n) (так же, как и пары (a, g) и (e, g)) противоречивы (они «кричат»).

Криптоанализ требует для заданного вероятного слова $(p_i,\ p_{i+1},\ \ldots,\ p_{i+k})$ некоторого подходящего начального индекса i, такого, чтобы последовательности $(p_i S_i,\ p_{i+1} S_{i+1},\ \ldots,\ p_{i+k} S_{i+k})$ и $(c_i S_i,\ c_{i+1} S_{i+1},\ \ldots,\ c_{i+k} S_{i+k})$ были изоморфны. Противоречия приводят к исключению данного индекса. Среди подходящих индексов, конечно, имеется и правильный (если встретилось настоящее вероятное слово). При этом, чем длиннее вероятное слово, тем меньше должно ожидаться промахов.

Указанное выше предварительное условие выполняется в случае (b^0) из разд. 7.2.2 при $S_i = \rho^i$. Такая ситуация, в частности, имеет место для коммерческих машин ЭНИГМА C и ЭНИГМА D без штепсельных коммутаторов, с тремя роторами и с фиксированными или движущимися отражателями (разд. 7.3.2, где штепсельный коммутатор T тождественен); в этом случае имеем:

$$S_{(i_1,i_2,i_3)} = \rho^{-i_1} R_N \rho^{i_1-i_2} R_M \rho^{i_2-i_3} R_L \rho^{i_3},$$

или

$$S_{(i_1,i_2,i_3,i_4)} = \rho^{-i_1} R_N \rho^{i_1-i_2} R_M \rho^{i_2-i_3} R_L \rho^{i_3-i_4},$$

если все роторы известны (а это, во всяком случае, справедливо для коммерческих машин) и порядок их тоже известен (в худшем случае, для трехроторной машины ЭНИГМА существует всего 6 порядков, которые нужно опробовать).

Кроме того, в случае ЭНИГМА подстановка U должна быть взаимообратной, что приводит к дальнейшим возможностям «крика» и исключению индекса (т. е. позиции роторов) или к подтверждению подходящего индекса при появлении взаимообратной подстановки. Возможность промахов уменьшается, так что инволютивный характер машины ЭНИГМА помогает дешифровальщику.

Кроме того, благодаря регулярности (разд. 8.4.4) движения ротора машины ЭНИГМА нет нужды обязательно проверять все $26^3 = 17\,576$ или $26^4 = 456\,976$ роторных алфавитов. Обычно достаточно рассмотреть лишь 26 позиций закрепленного ротора R_N, лежащих между двумя движениями среднего ротора R_M. Два других ротора остаются в это время фиксированными и образуют вместе с U псевдоотражатель

$$U'_{(i_2,i_3)} = \rho^{-i_2} R_M \rho^{i_2-i_3} R_L \rho^{i_3} \cdot U \cdot \rho^{-i_3} R_M \rho^{i_3-i_2} R_L \rho^{i_2}$$

или

$$U'_{(i_2,i_3,i_4)} = \rho^{-i_2} R_M \rho^{i_2-i_3} R_L \rho^{i_3-i_4} \cdot U \cdot \rho^{i_4-i_3} R_M \rho^{i_3-i_2} R_L \rho^{i_2}$$

с

$$S_i = \rho^{-i} R_N \rho^i\colon \quad c_i S_i = p_i S_i \cdot U'_{(i_2,i_3)} \quad \text{или} \quad c_i S_i = p_i S_i \cdot U'_{(i_2,i_3,i_4)}.$$

14.5.2. Ленточный метод. Для практического исполнения метода изоморфов снова применяется ленточный метод, где в качестве лент берутся столбцы вращаемых алфавитов. В следующем примере, принадлежащем Девуру и Кру, данное вероятное слово надо сравнить с фрагментом криптотекста:

r e c o n n a i s s a n c e
U P Y T E J O J Z E G B O T

Испытание нужно провести с ротором I из машины ЭНИГМА Вермахта, столбцы которой даны в разд. 7.3.5. Слово открытого текста и фрагмент криптотекста формируются лентами. Сопоставление символа за символом происходит так, как показано на рис. 100.

В каждой строке, кроме строки, обозначенной роторной позицией $i = 2$, имеются противоречия (одно из них всегда отмечено жирным шрифтом). Например, в строке $i = 0$ пары A Q и H Q так же, как и пары B N и D N нарушают инъективность, а пары R Y и R D нарушают единственность схемы шифрования; пары U A и A Q, F W и W I, Q R и R D, X U и U A, A Q и Q R, B N и N G, D N и N G, H Q и Q R нарушают инволютивное свойство. С другой стороны, в строке $i = 2$ имеются 2-циклы (J U), (M C), (S E), и инволютивное свойство не нарушается; это единственное попадание дает следующие пары изоморфов:

j g m g f u h r w c n s e w
U Z C Z B J O T A M Q E S A

Таким образом, ротор I подтверждается в качестве «быстрого» ротора R_N. Кроме того, 14 пар символов входа и выхода определяют уже девять 2-циклов псевдо-рефлектора $U'_{(i_2,i_3)}$ или $U'_{(i_2,i_3,i_4)}$, а именно (A W), (B F), (C M), (E S), (G Z), (H O), (J U), (N Q), (R T). Этот метод, очевидно, не нуждается в очень длинных вероятных словах.

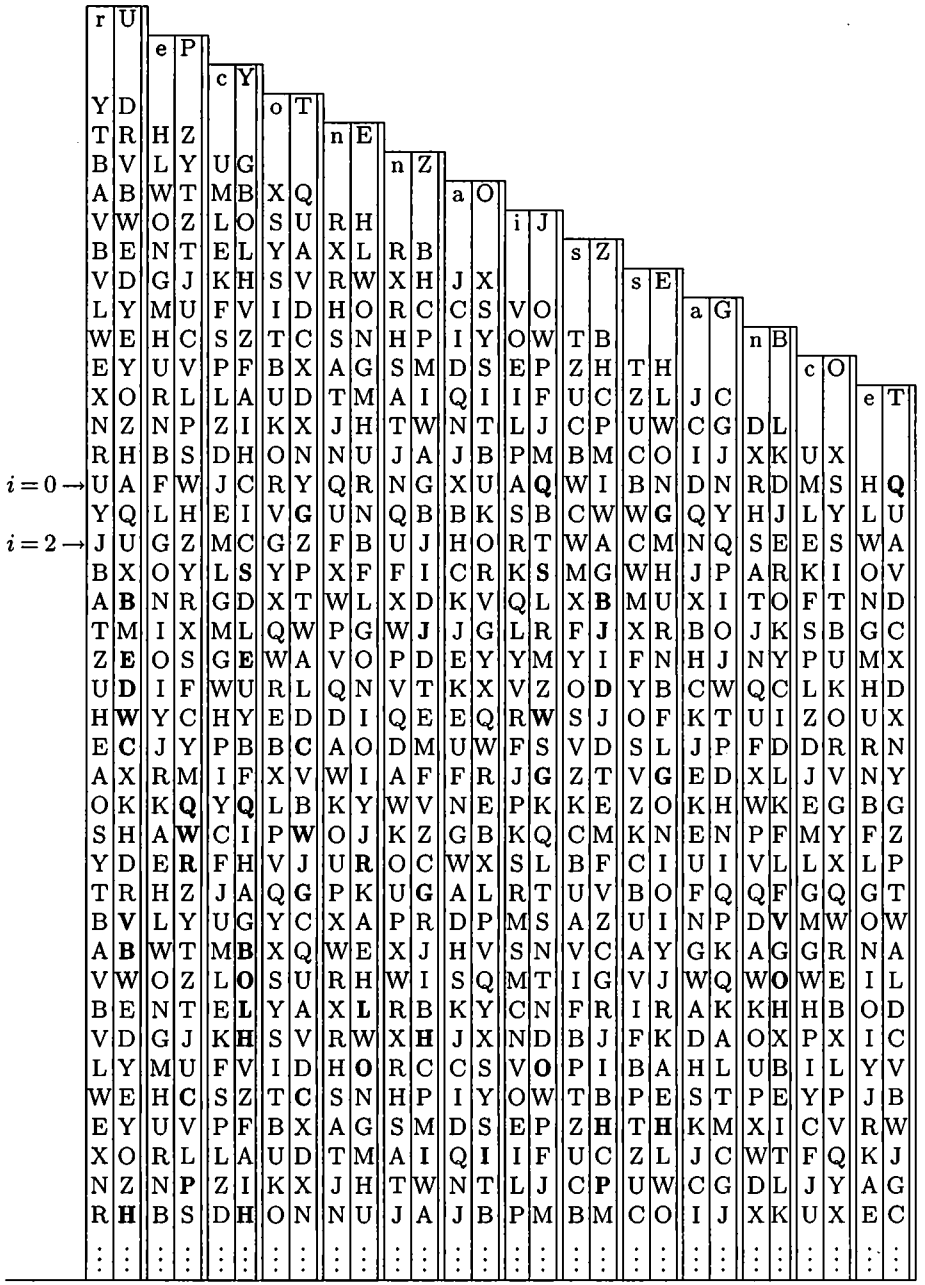


Рис. 101. Метод изоморфов с лентами (метод стержней)

Из изготовленного фабричным способом каталога с $2 \times 26^2 = 1352$ элементами всех $U'_{(i_2,i_3)}$ и $2 \times 26^3 = 35152$ элементами всех $U'_{(i_2,i_3,i_4)}$ можно определить позиции и порядок двух роторов II и III, служащих в качестве R_M и R_L. С таким указателем установки роторов, используя копию машины ЭНИГМА, можно провести дешифрование. Приведенный метод называется «встречей посредине». Швейцария, как и другие малые страны, использовала машину ЭНИГМА без штепсельных коммутаторов (с изменением роторных электрических схем) в течение (и частично после) Второй мировой войны (американское кодовое название INDIGO). С помощью каталогов немцы таким образом читали все их новости.

14.5.3. Исследование по частям. Проведенный выше анализ был основан на предположении, что средний ротор R_M при обусловленной краткости вероятного слова остается неподвижным. Если же он движется, то исследование разбивается на две части, не становясь при этом существенно труднее. Более того, в машине ЭНИГМА Вермахта положение выступа на роторе определяет его установку (в коммерческой машине ЭНИГМА положение выступа фиксировано и для каждого ротора оно известно). Если имеются два изоморфных текста (c', p') и (c'', p'') до и после углубления, то известны 2-циклы псевдоотражателя $U^{(1)}$ до и некоторые 2-циклы псевдоотражателя $U^{(2)}$ после углубления; это помогает найти позицию и порядок среднего ротора R_M. Для машины ЭНИГМА D это снижает объем каталога до $2 \times 26^2 = 1352$ элементов. Все это в пределах досягаемости.

Приведенные рассуждения можно проиллюстрировать примером (Девура). Мы постараемся исследовать следующую пару из (довольно длинного) вероятного слова и фрагмента криптотекста:

g e n e r a l f e l d m a r s c h a l l k e s s e l r i n g
L S H X B T F W U I O V B C A R X S N C V Z V X N J B F W B

Предположим, что исследование этого фрагмента в целом от позиции $i = 17$ до $i = 23$ дало попадание и два изоморфа. Продолжение до $i = 26$ (но не дальше) подтверждает это и дает

e x o v l y l x r u
M R H F D T D R X G

т. е. в псевдоотражатель $U^{(1)}$ входят 2-циклы (EM), (RX), (HO), (FV), (DL), (TY), (GU). Оставшийся текст должен быть связан со значениями от $i = 1$ до $i = 10$, что дает в действительности два изоморфа

b d a q r w r l j b s p f q c h m o o z
N R W X D A D J L N M Y H X I F S E E T

и 2-циклы (BN), (DR), (AW), (QX), (JL), (SM), (PY), (FH), (CI), (EO), (TZ), входящие в псевдорефлектор $U^{(2)}$. Эти два множества имеют 11 общих символов D, E, F, H, L, M, O, R, T, X и Y. Для ротора, который, по предположению, является средним ротором R_M, принимается следующая таблица вращений P-алфавитов:

i	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	L	W	F	T	B	A	X	J	D	S	C	K	P	R	Z	Q	Y	O	E	H	U	G	M	I	V	N
1	O	M	X	G	U	C	B	Y	K	E	T	D	L	Q	S	A	R	Z	P	F	I	V	H	N	J	W
2	X	P	N	Y	H	V	D	C	Z	L	F	U	E	M	R	T	B	S	A	Q	G	J	W	I	O	K
3	L	Y	Q	O	Z	I	W	E	D	A	M	G	V	F	N	S	U	C	T	B	R	H	K	X	J	P
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Для одиннадцати общих символов D, E, F, H, L, M, O, R, T, X, Y имеющийся там результат изображает таблицы псевдоотражателей $U^{(1)}$ и $U^{(2)}$:

$U^{(1)}$:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
i												l	m	v	o			d	e	h	x	y			r	t
0												K	P	G	Z			T	B	J	I	V			O	H
1												D	L	V	S			G	U	Y	N	J			Z	F
2												U	E	J	R			Y	H	C	I	O			S	Q
3												G	V	H	N			O	Z	E	X	J			C	B
⋮												⋮	⋮	⋮	⋮			⋮	⋮	⋮	⋮	⋮			⋮	⋮

$U^{(2)}$:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
i												r	o	h	f			j	s	e	d	z			q	p
0												O	Z	J	A			S	E	B	T	N			Y	Q
1												Z	S	Y	C			E	P	U	G	W			R	A
2												S	R	C	V			L	A	H	Y	K			B	T
3												C	N	E	I			A	T	Z	O	P			U	S
⋮												⋮	⋮	⋮	⋮			⋮	⋮	⋮	⋮	⋮			⋮	⋮

Сравнивая теперь $i=0$ для $U^{(1)}$ и $i=1$ для $U^{(2)}$, мы находим общую букву Z в строке $i=0$ под h и в строке $i=1$ под d:

i	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0							K	P	G	**Z**			T	B		J			I		V				O	H
1							**Z**	S	Y		C		E	P		U			G		W				R	A

Однако в соответствующей вырезке из таблицы вращений P-алфавитов

i	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	L	W	F	T	B	A	X	**J**	D	S	C	K	P	R	Z	Q	Y	O	E	H	U	G	M	I	V	N
1	O	M	X	**G**	U	C	B	Y	K	E	T	D	L	Q	S	A	R	Z	P	F	I	V	H	N	J	W

совпадений нет. Таким образом, эти роторные позиции «кричат».

Сравнивая, с другой стороны, $i=1$ для $U^{(1)}$ и $i=2$ для $U^{(2)}$, находим

букву L в строке $i=1$ под e, а в строке $i=2$ под l,
букву V в строке $i=1$ под f, а в строке $i=2$ под h,
букву S в строке $i=1$ под h, а в строке $i=2$ под d,
букву Y в строке $i=1$ под o, а в строке $i=2$ под r:

i	a b c d e f g h i j k l m n o p q r s t u v w x y z
1	D **L V** S G U **Y** N J Z F
2	S R C **V** **L** A H **Y** K Z F

В соответствующих строках, взятых из таблицы вращений P-алфавитов

i	a b c d e f g h i j k l m n o p q r s t u v w x y z
1	O M X **G U C** B **Y** K E T D L Q **S** A R Z P F I V H N J W
2	X P N **Y** H V D **C** Z L F **U** E M R T B **S** A Q G J W I O K

имеются совпадения во всех случаях (с **Y** для **S**, **U** для **L** и т. д.). Таким образом, эта роторная позиция является попаданием, и правильная позиция среднего ротора R_M найдена. Практически это определение роторной позиции можно было бы провести также с помощью лент, содержащих столбцы вращаемых алфавитов.

14.5.4. Съемный отражатель. Вариант указанного метода позволяет определить все 2-циклы истинного отражателя U. Это стало необходимым для союзников, когда в начале 1944 г. немцы время от времени стали использовать в ЭНИГМЕ Вермахта (см. разд. 7.3.3) «съемный» отражатель. Теперь метод изоморфов проводится через все $26^3 = 17\,576$ начальных индексов с соответствующими вероятными словами. Эта задача требовала специальной механизации. Релейная машина AUTOSCRITCHER (1944 г.) и электронная машина SUPERSCRITCHER (1946 г.) были построены в США отделом F военного Агентства безопасности связи, возглавляемым полковником Розеном[1)].

14.5.5. Использование штекеринга. Штепсельный коммутатор погубил метод изоморфов, потому что неизвестная коммутаторная связь (штекеринг) скрывала вероятное слово открытого текста. Теперь требовалось найти повторяющиеся пары символов открытого текста и соответствующего криптотекста. Каждая группа таких пар отображалась при коммутаторной перестановке в группу соответствующих символов из двух изоморфов. Изоморфизм требует, чтобы группы не расщеплялись, когда опробуются все роторные позиции. Машины AUTOSCRITCHER и SUPERSCRITCHER были спроектированы с учетом этого требования. Удовлетворяли ему также машина DUENNA американского военно-морского флота и английская GIANT.

14.6. Скрытый компромисс OT-КТ

Целью метода вероятного слова является восстановление открытого текста. OT-КТ компромисс (открытый текст—криптотекст компромисс[2)]) для

[1)] В открытой литературе слова «scritch, scritchmus» использовались без подробного объяснения, например, Таунт (1993 г.), когда описывал атмосферу работы в Блетчли Парк со ссылкой на обязанности Дениса Бэббиджа, упомянул также съемный отражатель. Поэтому происхождение этого термина следует искать в Великобритании. Кроуфорд и Фокс сообщали в 1992 г., что они построили AUTOSCRITCHER и SUPERSCRITCHER, но не дали информации о криптоаналитической подоплеке. В работе Девура (1995 г.) устанавливалась связь с методом изоморфов. «Scritch» — это диалектный вариант «screech».

[2)] Слово «компромисс», согласно словарю Мерриама-Вебстера, среди других значений имеет и такое: «приводить к опасности, подвергать опасности некоторыми действиями, которые нельзя вернуть обратно, оставлять беззащитным перед каким-нибудь злом». В криптологии слово «компромисс» имеет именно этот смысл.

этого не обязателен, хотя в случае удачи он дает шанс восстановить ключ и таким образом получить намного больше, чем лишь открытый текст. Технически все методы, которые работают для вероятных слов, применимы и здесь, причем можно выбирать длинные слова.

Но подозрение (или надежда, — смотря о какой стороне идет речь) о том, что имеет место прямой ОТ-КТ компромисс возникает не слишком часто. Однако бывают непрямые случаи, когда, например, открытый текст получается дешифрованием и вдруг находится некий другой криптотекст, полученный зашифрованием того же открытого текста с помощью другой криптосистемы. Существует много видов халатности и глупости, которые могут привести к такой ситуации, которая начинается как безобидный КТ-КТ (криптотекст—криптотекст) компромисс.

Имеется огромное число возможных путей, приводящих к такому компромиссу. Например, он может встретиться при возникновении организационных проблем управления ключами. Радикальное изменение в криптосистеме не всегда можно осуществить гладко, и может случиться так, что сообщение, уже посланное в старом ключе, повторяется в новом. Как серьезна эта опасность, можно судить по поговорке: «Риск сломать криптосистему никогда не выше, чем риск сломать жизнь». Хюттенхайн сообщил, что между 1942 г. и сентябрем 1944 г. множество так называемых CQ-сообщений («call to quarters» — сигналы общего назначения), посланных из госдепартамента в Вашингтоне своим дипломатическим аванпостам, прочитывались немцами. Наборы CQ-лент для машин М-138 были идентичны для всех посольств. Таким образом, как только шифр был взломан, компромисс происходил почти неизбежно, после перехода к новому набору лент.

Кроме того, для многих практически используемых методов, когда система известна, ОТ-КТ компромисс становится особенно опасным, так как ставится под удар даже ключ, и таким образом возможна глубокая брешь в криптосистеме. Хюттенхайн в своей ретроспективе (1978 г.) заключил, что ни один метод шифрования не должен использоваться, если он восприимчив к ОТ-КТ компромиссу («Es dürfen also keine Chiffrierverfahren verwendet werden, die gegen Klar-Gehein-Kompromisse anfällig sind». В комбинации с принципом Керкхоффса принцип Хюттенхайна вычеркивает многие излюбленные классические криптосистемы; он исключает все те, которые обладают свойством Шеннона (см. разд. 2.6.4).

Ошибки обязательно случаются всюду. Кан заметил в этой связи: «Немцы не имели монополии на криптографические неудачи. В этом отношении англичане были точно так же нелогичны, как и немцы».

Глава 15

Анатомия языка: частоты

Мы можем только сказать, что дешифрование любого шифра, даже простейшего, будет иногда приводить к сюрпризам.

Элен Фуше Гейнс, 1939 *г.*

Дешифрование, рассматриваемое до сих пор, основывалось на шаблонах, использовало «костяк» языка, лежащего в основе открытого текста. Стратегия же дешифрования, к которой мы приступаем теперь, использует его «внутренние органы». Это нацеливает на стохастические законы языка, в частности, на частоты символов и мультиграмм. Этот аспект криптографии был известен уже Леону Баттисте Альберти (Trattati in cifra [Трактат о шифрах (*итал.*)], 1470 г.).

Прежде всего, имеет место

Теорема инвариантности 2. Для всех простых перестановок *частоты индивидуальных символов в тексте являются инвариантными.*

15.1. Исключение из шифровальных методов

Теорема 2 может быть использована негативно для исключения перестановок, а именно, если частоты индивидуальных символов криптотекста являются определенно не теми, которые имеет предполагаемый язык открытого текста. При этом следует проявлять осмотрительность. Например, данные технических измерений вполне могут иметь частоты символов, отличные от частот обычных естественных языков.

15.1.1. Криптотекст

FDRJN	UHVXX	URDMD	SKVSO	PJRKZ	DYFZJ
XGSRR	VTQYR	WDARW	DFVRK	VDRKV	TDFSZ
ZDYFR	DNNVO	VTSXS	AWVZR		

показывает, что R, D, V, S — наиболее часто встречающиеся символы, тогда как B, C, E, I, L очень редки (в действительности они просто отсутствуют).

Такой криптотекст нельзя получить из английского, немецкого, французского или итальянского открытого текста перестановкой. На самом деле это простая подстановка, см. разд. 13.3.1.

15.1.2. С другой стороны, нельзя исключить, что криптотекст (разд. 12.5, табл. 6)

SAEWS	HRCNU	ODKLN	ELIAS	HNCIO	NBNNA
AKIHM	CWNZA	MCGIM	IHEEN	NAUFK	NNCTI
TIHMD	RTEWO	ATAIM	TALKB	UEAFZ	LNUSE
ASDEN	...				

с E и T среди наиболее частых индивидуальных символов, а V, P, J, Q, X и Y отсутствуют, получен перестановкой из немецкого текста.

15.1.3. Теорема 2 также используется (логически недопустимо) в смысле правдоподобной аргументации: если распределение частот индивидуальных символов такое же, как в соответствующем естественном языке, то *предположительно* имела место перестановка. Наивный аргумент: что же еще, какая другая процедура могла бы оставить инвариантным распределение частот индивидуальных символов? Это могло бы быть правдоподобным, если бы была уверенность, что никто не мог взять на себя хлопоты, чтобы придумать шифрование, совершенно отличное от перестановки, но оставляющее распределение частот инвариантным, — но это не доказательство, и приговор суда не может основываться на этом. На самом деле, омофонную многосимвольную подстановку, скажем, некоторый код, можно легко сделать так, чтобы она обладала предписанными частотами символов. Хоман описал в 1948 г. метод кодирования, который дает все символы с равными частотами («равночастотный шифр»). То же самое достигается уничтожающими избыточность кодированиями Хаффмана и Шеннона.

Такие трюки, однако, не могут надолго задержать профессионального дешифровальщика. Хотя в 1892 г. великий Базерье, пытаясь разгадать сообщение, захваченное у группы французских анархистов, потратил две лишних недели, так как он был введен в заблуждение шестью пустышками, добавленными к началу и концу и отдельными буквами, подмешанными в сообщение. На самом деле это был шифр ВИЖЕНЕРа с периодом 6, и в другом случае его взлом был бы тривиальным для Этьена Базерье. Возможно, это было просто ошибкой — добавить ровно столько букв, каков был период, поэтому может быть Базерье просто не мог вообразить такой глупости. Часть ужасного сообщения, между тем, читалась так: «La femme et lui sont des mouchards, s'il m'arrive quelque chose, songe à les supprimer [он и женщина шпионы; если со мной что-то случится, позаботьтесь об их ликвидации (*фр.*)]».

15.2. Инвариантность разбиений

Разбиения этой главы играют роль шаблонов гл. 13: это абстрактное средство для характеризации инвариантности частот. Разбиение — это разложение

натурального числа M в сумму натуральных чисел m_i:

$$M = m_1 + m_2 + \ldots + m_N.$$

Каждому тексту длины M соответствует некоторое разбиение числа M, образованное числами вхождений в этот текст N индивидуальных символов из словаря Z_N, пустышки при этом обычно подавляются, таким образом, текст

w i n t e r s e m e s t e r

имеет разбиение

$$14 = 4 + 2 + 2 + 2 + 1 + 1 + 1 + 1.$$

Поэтому мы говорим о разбиении по числам символов в тексте.

Имеет место следующая фундаментальная теорема, соответствующая теореме 1 (разд. 13.1).

Теорема инвариантности 3. Для всех одноалфавитных функциональных простых подстановок, в частности, для одноалфавитных линейных простых подстановок (включая сложение ЦЕЗАРЯ и реверсивность), *разбиение длины текста по числам идивидуальных символов является инвариантом.*

Одноалфавитное шифрование открытого текста w i n t e r s e m e s t e r функциональными простыми подстановками, какими бы они ни были, всегда состоит из 4 экземпляров какого-то одного символа, 2 экземпляров какого-то другого символа, 2 экземпляров третьего символа и т. д. Разбиение же $4 + 2 + 2 + 2 + 1 + 1 + 1 + 1$ остается инвариантным.

Пусть дан криптотекст

Z L Q W H U V H P H V W H U

и предположим, что незаконному дешифровальщику известны частоты определенных букв открытого текста, а именно четыре раза /e/, дважды /t/, дважды /r/, дважды /s/ и по разу /n/, /i/, /w/ и /m/, тогда он знает, что

$$\mathrm{H} \mathrel{\hat{=}} \mathrm{e}, \quad \{\mathrm{U\ V\ W}\} \mathrel{\hat{=}} \{\mathrm{r\ s\ t}\}, \quad \{\mathrm{L\ P\ Q\ Z}\} \mathrel{\hat{=}} \{\mathrm{i\ m\ n\ w}\},$$

и имеет полифоническое дешифрование

```
i i i i
m m m r  r r  m  r r  r
n n n s e s s e n e s s e s
w w w t  t t  w  t t  t
```

На самом деле незаконный дешифровальщик знает намного меньше, так как он знает частоты всех букв открытого текста лишь приблизительно. Согласно присущим языку законам, каждый символ χ_i появляется с определенной вероятностью p_i («стохастический источник» Q), так что частота $m_i = Q[\chi_i]$ его встречаемости в открытом тексте размера M близка к $M \cdot p_i$.

15.3. Интуитивный метод. Частотный профиль

Чтобы преуспеть в интуитивном методе дешифровки одноалфавитных подстановок, рекомендуется наглядно представить себе «частотный профиль» рассматриваемого языка.

В английском языке (рис. 101) частотный профиль показывает заметный e-пик и немного меньший a-пик. Имеется также заметное возвышение хребта r-s-t и еще два хребта поменьше l-m-n-o и h-i.

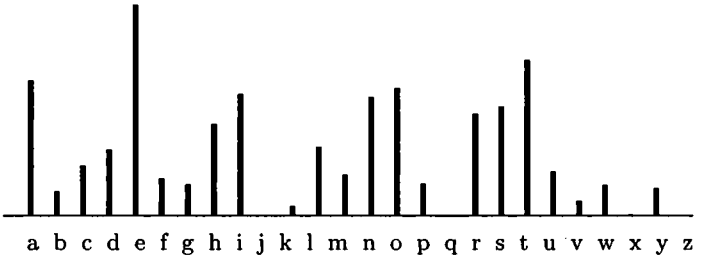


Рис. 102. Частотный профиль английского языка

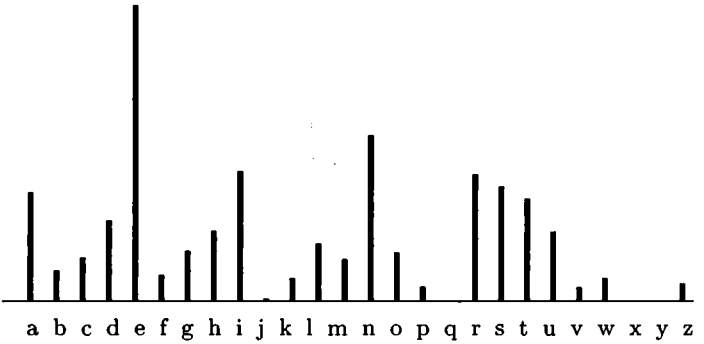


Рис. 103. Частотный профиль немецкого языка

В немецком языке (рис. 102) частотный профиль довольно похож, но e-пик выделяется сильнее, имеется хребет пошире r-s-t-u и еще один хребет пошире f-g-h-i. Оба языка показывают j-k впадину, p-q впадину и очень заметную низменность v-w-x-y-z. Различия между любыми большими европейскими языками (вроде французского, итальянского или испанского) не больше, чем между английским и немецким. В романских языках a-пик более заметен и есть изолированный i-пик. На взгляд они довольно похожи.

15.3.1. Для перестановки частотный подсчет дает профиль, близкий к профилю рассматриваемого языка. Но также и одноалфавитная простая подстановка с $q = 1$ и сложение ЦЕЗАРЯ обнаруживаются с первого взгляда.

Теорема инвариантности 4. Для всех сложений ЦЕЗАРЯ *частотный профиль текста циклически сдвигается.*

Криптотекст из 349 символов

HVZDU	VFKRQ	GXQNH	ODOVL	FKLQE	RQQDQ
NDPLF	KCZDQ	JPLFK	PHLQH	DQNXQ	IWQLF
KWPLW	GHUDX	WRPDW	LNDEO	DXIHQ	CXODV
VHQGL	HVLFK	LQIXH	QIMDH	KULJH	PXQWH
UZHJV	VHLQK	HUDXV	JHELO	GHWKD	WEDKQ
VWHLJ	WUHSS	HUXQW	HUEDK	QVWHL	JWUHS
SHUDX	IUHLV	HWDVF	KHDEV	WHOOH	QIDKU
NDUWH	DXVGH	UPDQW	HOWDV	FKHQH	KPHQU
HLVHW	DVFKH	DXIQH	KPHQI	DKUND	UWHDE
JHEHQ	CXPCH	LWXQJ	VVWDQ	GDEHQ	GCHLW
XQJHQ	NDXIH	QQDFK	GUDXV	VHQJH	KHQXQ
GHLQW	DALKH	UDQZL	QNHQ		

имеет частотный профиль, изображенный на рис. 103. Очевидно, что шифрованием является сложение ЦЕЗАРЯ со сдвигом на 3. Здесь зашифрован фрагмент из романа Генриха Бёлля.

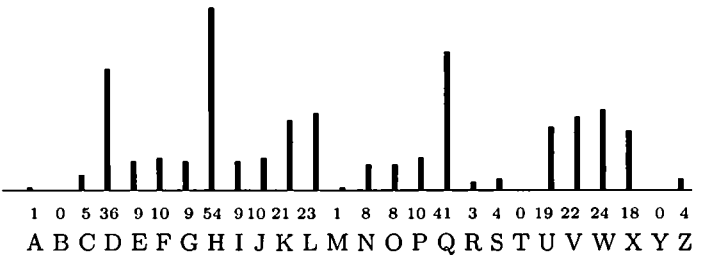


Рис. 104. Частотный профиль криптотекста Бёлля (разд. 15.3.1)

15.3.2. Но следует заметить, что теорему 4 нельзя обратить. Циклически сдвинутый частотный профиль совместим с композицией перестановки и сложения ЦЕЗАРЯ.

15.3.3. Интуитивный метод может при исключительных обстоятельствах ввести в заблуждение, как, например, в следующем криптотексте из 175 символов.

VQPOU	TKTKB	IKTCB	NHPKO	HUPTI	PXZPV
IPXBC	VODIP	GCSKH	IUZPV	OHGPM	LTEKE
GKOEB	DIBNQ	KPOBN	BOXKU	ICPZT	BOEHK
SMTPG	IKTPX	OBNBO	PGTPE	PNKOU	KOHBO
EIBQQ	ZKOEK	WKEVB	MKUZU	IBUZP	VUIKT
ESBXO	UPIKN	BTKTB	GMZUP	BTVHB	SCPXM

Его частотный профиль (рис. 104) показывает значительное отклонение от рис. 101 или рис. 102. Можно подумать о каком-либо экзотическом языке. Подозрение, что это тоже сложение ЦЕЗАРЯ (над Z_{25}) со сдвигом на 1: «upon this basis i am going to show you... [на этом основании я собираюсь показать

вам (*англ.*)]» — возникает под влиянием ленточного метода из разд. 12.7, который уже дает «upon» для первых четырех символов без какого-либо сомнения. Провал интуитивного метода, ориентированного на частоты, вызван искажением частот символов: открытый текст является липограммой (текстом, в котором не встречаются какие-либо буквы), взятой из Gadsby (разд. 13.3.2, рис. 88).

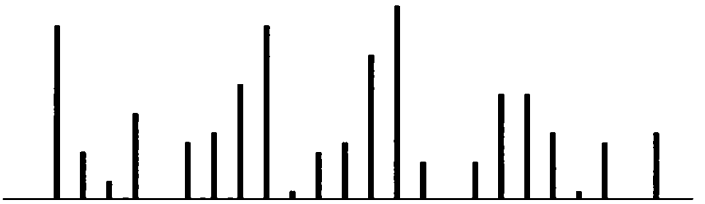


Рис. 105. Частотный профиль криптотекста из разд. 15.3.3

15.4. Частотное упорядочение

Для одноалфавитной линейной простой подстановки с $q = -1$ (особенно для реверсивной) частотный профиль просто отражается справа налево. Для значений q, отличных от 1 и -1, а также для нелинейных простых подстановок частотный профиль бесполезен: соседство букв разрывается. Наивная интуиция в этом случае использует частотное упорядочение: наиболее частый символ в криптотексте будет соответствовать наиболее частой букве рассматриваемого языка. После замены этого символа криптотекста соответствующим символом открытого текста процедура повторяется до тех пор, пока все символы не будут исчерпаны и не будет установлена вся схема шифрования.

15.4.1. Изъяны частотного упорядочения. Теоретически этот метод должен был бы работать — по крайней мере для достаточно длинных текстов — достаточно длинных для того, чтобы немногочисленные липограммы, которые там могут существовать, утонули в массе «нормальных» текстов. Но пример, приведенный в разд. 15.2, приводит к полифонической ситуации, даже если известны верные частоты букв открытого текста, что показывает фундаментальную ограниченность этой процедуры: существует много букв открытого текста с одной и той же частотой, и тогда выбор является недетерминированным.

Более того, даже длинные тексты обычно показывают значительные колебания частот символов. Поэтому «распределение частот» английского языка — фикция, и в лучшем случае в военном, дипломатическом, коммерческом или литературном подъязыках имеет место определенная однородность; действительно, даже один и тот же человек может разговаривать на разных «английских» языках в зависимости от окружения. Соответственно, статисти-

ка на частоты букв в различных языках совершенно различна. Кроме того, большинство старых подсчетов основывались на текстах, содержащих не более 10 000 букв. И по вопросу о частотной упорядоченности уже имеется большое расхождение в литературе.

Для английского языка

eaoidhnrstuycfglmwbkpqxz	(E. A. Poe, 1843)
etaoinshrdlucmfwypvbgkqjxz	(O. Mergenthaler, 1884)
etoanirshdlcfumpywgbvkxjqz	(P. Valério, 1893)
etaonisrhldcupfmwybgvkqxjz	(H. F. Gaines, O. P. Meaker, 1939)
etoanirshdlcwumfygpbvkxqjz	(L. D. Smith, 1943)
etoanirshdlufcmpywgbvkxzjq	(L. Sacco, 1951)
etaonirshdlucmpfywgbvjkqxz	(D. Kahn, 1967)
etaonrishdlfcmugpywbvkxjqz	(A. G. Konheim, 1981)
etaoinsrhldcumfpgwybvkxjqz	(C. H. Meyer, S. M. Matyas, 1982)

Для французского языка

eusranilotdpmcbvghxqfjyzkw	(Ch. Vesin de Romanini, 1840)
ensautorilcdvpmqfgbhxyjzkx	(F. W. Kasiski, 1863)
esriantouldmcpvfqgxjbhzykw	(A. Kerckhoffs, 1883)
easintrulodcpmvqfgbhjxyzkw	(G. de Viaris, 1893)
enairstuoldcmpvfbgqhxjyzkw	(P. Valério, 1893, M. Givierge, 1925)
eaistnrulodmpcvqgbfjhzxykw	(H. F. Gaines, 1939)
etainroshdlcfumgpwbyvkqxjz	(Ch. Eyraud, 1953)

Для немецкого языка

enrisdutaghlobmfzkcwvjpqxy	(Ch. Vesin de Romanini, 1840)
enirsahtudlcgmwfbozkpjvqxy	(F. W. Kasiski, 1863)
enirstudahgolbmfzcwkvpjqxy	(E. B. Fleissner von Wostrowitz, 1881)
enritsduahlcgozmbwfkvpjqxy	(P. Valério, 1893)
enrisatdhulcgmobzwfkvpjyqx	(F. W. Kaeding, 1898)
enritsduahlcgozmbwfkvpjyqx	(M. Givierge, 1925)
enirstudahgolbmfczwkvpjqxy	(A. Figl, 1926)
enirsadtugholbmcfwzkvpjyqx	(H. F. Gaines, J. Arthold, 1939)
enristudahglocmbzfwkvpjqxy	(L. D. Smith, 1943)
enritsudahlcgozmbwfkvpjqxy	(L. Sacco, 1951)
enisrtahduglcofmbwkzvpjyqx	(Ch. Eyraud, 1953)
enristdhaulcgmobzwfkvüpäöjyqx	(K. Küpfmüller, H. Zemanek, 1954)
enisratduhglcmwobfzkvpjqxy	(W. Jensen, 1955)
enisratdhulcgmobwfkzpvjyxq	(A. Beutelspacher, 1987)
enirsatdhulgocmbfwkzpvjyxq	(F. L. Bauer, 1993)

Подсчет, произведенный Кюпфмюллером и Земанеком, включающий модифицированные гласные, является криптологически несколько неуместным

и приводится здесь просто для сравнения. Аналогичные таблицы для итальянского, испанского, датского и латинского языков можно найти в книге Ланга и Сударa, 1935 г.

Для первой дюжины (приблизительно) букв существуют приятные мнемонические строфы вроде

Английский:	etaoinshrdlu	(LYNOTYPE)
Французский:	esarintulo	(Базерье, Живерж)
Немецкий:	enirstaduhl	(Хюттенхайн)
Итальянский:	eiaorlnts	(Сакко)

Частотное распределение в английском языке было отражено уже в длине символов кода Морзе, используемого на телеграфе — Морзе подсчитал литеры в ящике типографского шрифта и нашел там 12000 /e/, 9 000 /t/, 8 000 /a/, /i/, /n/, /o/, /s/, 6 400 /h/. По техническим причинам распределение частот букв английского языка также влияет на распределение клавиш пишущей машинки LINOTYPE Оттмара Мергенталера (1854–1899 гг.) (рис. 105).

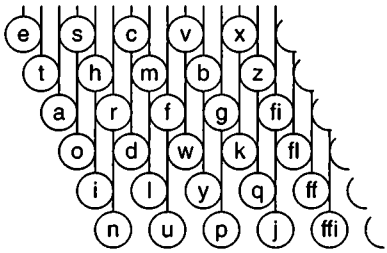


Рис. 106. Оригинальные клавиши пишущей машинки LINOTYPE Оттмара Мергенталера, 1884 г.

15.4.2. Подсчет частот. Для немецкого языка в 1898 г. Кэдинг провел обширный подсчет частот. Для стенографических целей он изучил тексты, включающие в совокупности 20 миллионов слогов и 62 069 452 буквы (с ä, ö, ü, замененными на ae, oe, ue соответственно). Мы можем предположить, что этот подсчет был достаточно большим, чтобы быть пристрастным.

Упомянутое выше частотное упорядочение основывается на этом подсчете. Если мы сопоставим его с упорядочением частот для криптотекста Бёлля в разд. 15.3.1 (рис. 103), располагая буквы с одной и той же частотой в порядке, соответствующем алфавиту, то получим следующую таблицу дешифрования:

54 41 36 24 23 22 21 19 18 10 10 10 9 9 9 8 8 5 4 4 3 1 1 0 0 0
H Q D W L V K U X F J P E G I N O C S Z R A M B T Y
e n r i s a t d h u l c g m o b z w f k v p j y q x.

Расшифровывая начало криптотекста Бёлля

HVZDU VFKRO GXQNH ODOLV FKLQE RQQDQ

с помощью построенной выше таблицы, мы получаем абсолютно неприемлемый «открытый» текст:

eakrd autvn mhnbe zrzai uting vnnrn,

перестановки букв с равными частотами в разном порядке не улучшает ситуации. Фактически же таблица дешифрования (точнее, расшифрования) имеет вид

H Q D W L V K U X F J P E G I N O C S Z R A M B T Y
e n a t i s h r u c g m b d f k l z p w o x j y q v,

так что соответствующий открытый текст был таким:

eswar schon dunke lalsi chinb onnan.

15.5. Клики и подгонка разбиений

Рис. 106 показывает, что в предыдущем примере правильное упорядочение частот отличается от теоретического, основанного на вероятностях: происходят отдельные перестановки, когда /r/ и /v/ перемещаются на 5 позиций, а /d/, /e/ и /o/ — на 6. Другие буквы перемещаются только на одну или две позиции, — но будь это перемещение большим или малым — оно нарушает правильное соответствие символов OT и KT (открытого текста и криптотекста). Только несколько букв (среди них /e/ и /n/) были правильно расставлены в пары. Конечно, виновата в этом колебании краткость криптотекста Бёлля, но не только она, как мы сейчас увидим.

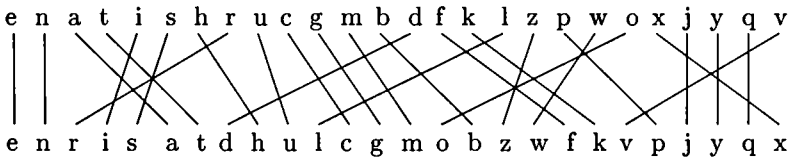


Рис. 107. Соответствие между наблюдаемыми и основанными на вероятности частотами для примера из разд. 15.4.2

Не существует полностью автоматического дешифрования, основанного только на частотном упорядочении. Причина состоит в том, что даже в длинных текстах встречаются колебания частот, поэтому эмпирически найденные вероятности также колеблются. Это и приводит к перемещениям в частотном упорядочении.

15.5.1. Колебания. На самом деле, не только порядок частот, но также и сами частоты индивидуальных символов, приводимые в литературе, показывают отклонения. Поэтому мы исследуем колебания, которые следует ожидать, на немецких текстах из 100, 1000, 10 000 и 100 000 символов.

Типичный результат показан на рис. 107 (пунктирная линия указывает среднюю частоту). Текст, содержащий 681 972 символов, был собран из всех политических комментариев ежедневных газет в марте 1992 г. (с тех пор он называется SZ3-92). Ясно видно налегание областей колебания, и как колебание убывает с ростом объема текста. Само колебание убывает, грубо говоря, про-

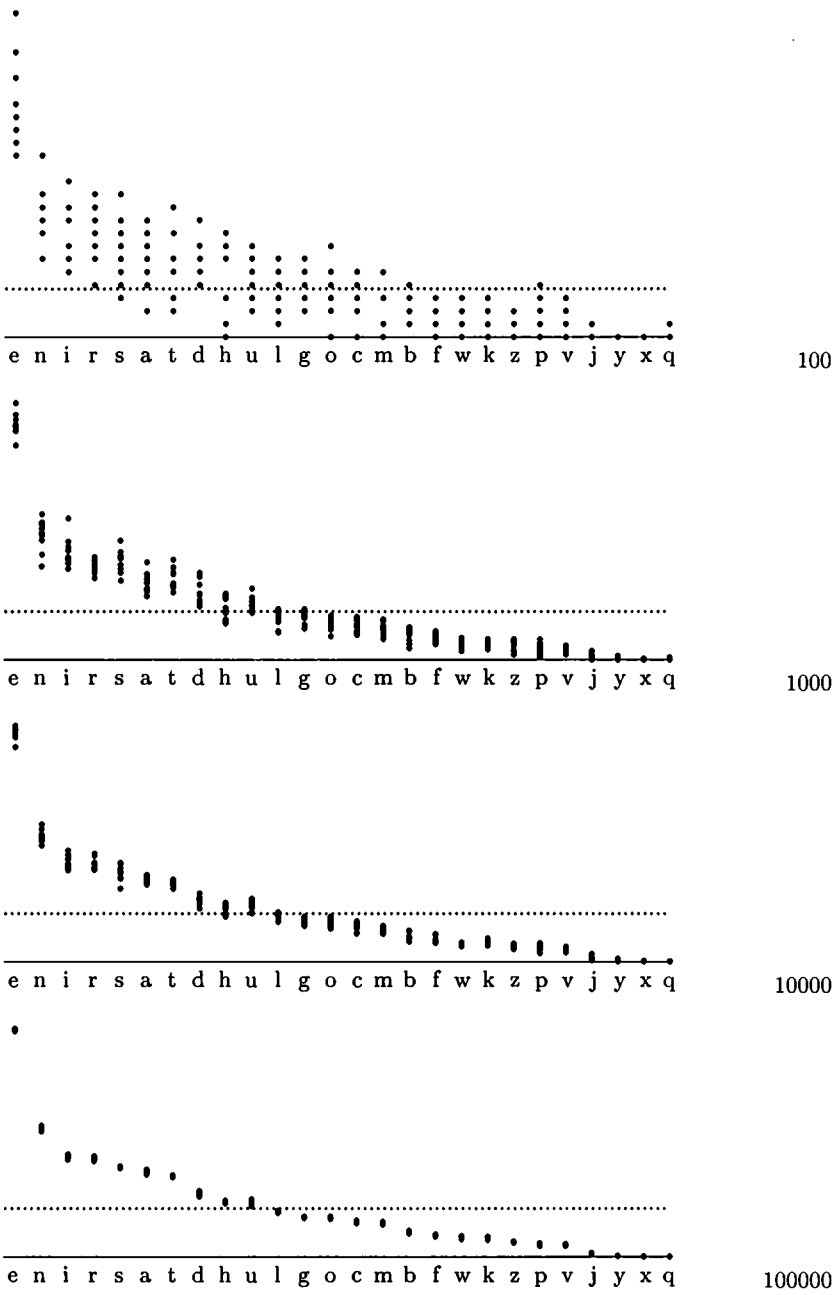


Рис. 108. Колебания частот индивидуальных букв в немецких газетах

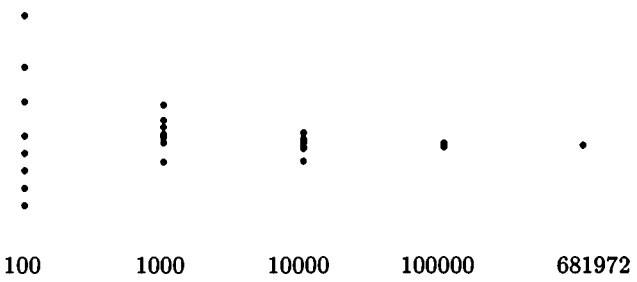


Рис. 109. Колебание частоты буквы /e/ немецкого языка в зависимости от длины текста

порционально квадратному корню из длины текста, как показано на рис. 108 для буквы /e/.

Мы специально проделали эксперимент сопоставления порядка частот английского текста по Мейеру–Матиашу (разд. 15.4.1) с довольно длинным английским текстом из 29 272 символов (взятым из этой книги), частоты букв которого приведены на рис. 109.

3879	2697	2240	2151	2133	2082	1910	1907	1415	1095	1035	995	780
e	t	a	n	o	i	r	s	h	d	l	c	m
765	719	687	620	551	469	404	277	230	101	55	45	30
u	f	p	y	g	w	b	v	k	x	z	q	j

Рис. 110. Частотное распределение в английском тексте из 29 272 символов (из этой книги)

Результаты сопоставления представлены на рис. 110. Перемещений немного, но все же они есть.

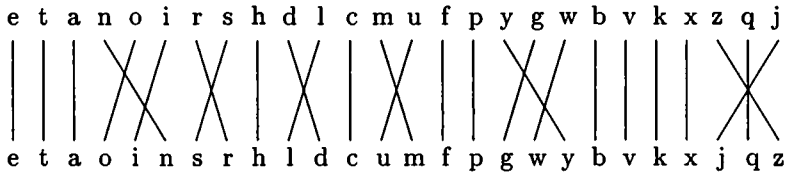


Рис. 111. Сопоставление наблюдаемых частот в английском тексте из 29 272 символов с частотами, основанными на вероятностях (Мейер—Матиаш)

Если длинный английский текст зашифрован простой подстановкой, а затем дешифруется посредством сопоставления наблюдаемых и основанных на вероятности частотных порядков, то может получиться текст, фрагмент которого приводится ниже:

```
i v e s t h e c e o t m s n e r c s g p t i d i w g h a r c i d d e c t
e l a t s e a r m s g i f e x p e s n e o c e r e v e o t h e i p e o d
n t e s a t m s e r h i y r t h n r t h e r e e x p e s n e o c e r o i
s u a d d g r c a t t e s e l c a o b e c i o c e o t s a t e l n o t i
a f e y u a x n u r f i s c s g p t i w s a p h n c y i s k n o p a s t
```

Его нельзя прочесть гладко, и только после выполнения раздражающего 3-цикла между /i/, /o/ и /п/ и 2-цикла между /r/ и /s/ результат может служить в качестве грубой расшифровки. Верный текст (см. начало разд. 11.2) таков:

```
o v e r t h e c e n t u r i e s c r y p t o l o g y h a s c o l l e c t
e d a t r e a s u r y o f e x p e r i e n c e s e v e n t h e o p e n l
i t e r a t u r e s h o w s t h i s t h e s e e x p e r i e n c e s n o
r m a l l y s c a t t e r e d c a n b e c o n c e n t r a t e d i n t o
a f e w m a x i m s f o r c r y p t o g r a p h i c w o r k i n p a r t
```

15.5.2. Клики. Пожалуй, работе с частичными упорядочениями символов следует предпочесть работу с упорядочениями «равночастотных» клик символов, которые трудно разделить на основе подсчета их частот.

Для английского языка существует разложение множества букв на клики, определенные Смитом в 1943 г.:

$$\{\text{etaoin}\},\quad \{\text{srh}\}\,\{\text{ld}\}\,\{\text{cumfpgwyb}\}\quad \{\text{vk}\}\quad \{\text{xjqz}\},$$

или, в немного более подробном разложении,

$$\{\text{e}\}\,\{\text{t}\}\,\{\text{aoin}\}\quad \{\text{srh}\}\,\{\text{ld}\}\,\{\text{cumf}\}\,\{\text{pgwyb}\}\quad \{\text{vk}\}\quad \{\text{xjqz}\},$$

которое для длинных и «нормальных» текстов может быть разложено еще дальше:

$$\{\text{e}\}\{\text{t}\}\{\text{ao}\}\{\text{in}\}\{\text{srh}\}\{\text{ld}\}\{\text{cu}\}\{\text{mf}\}\{\text{p}\}\{\text{gwy}\}\{\text{b}\}\{\text{v}\}\{\text{k}\}\{\text{xjqz}\}.$$

Для немецкого языка также существует разложение множества букв в клики, данные Лангом и Сударом в 1935 г.:

$$\{\text{e}\}\,\{\text{nirsatdhu}\}\quad \{\text{lgosmbfwkz}\}\quad \{\text{pvjyxq}\},$$

или, в чуть более подробно разложенные клики

$$\{\text{e}\}\,\{\text{n}\}\,\{\text{irsat}\}\,\{\text{dhu}\}\quad \{\text{lgocm}\}\quad \{\text{bfwkz}\}\quad \{\text{pv}\}\,\{\text{jyxq}\},$$

которые для длинных и нормальных текстов можно разложить дальше в

$$\{\text{e}\}\{\text{u}\}\{\text{ir}\}\{\text{sat}\}\{\text{dhu}\}\{\text{lgo}\}\{\text{cm}\}\{\text{bfwkz}\}\{\text{pv}\}\{\text{jyxq}\}.$$

Можно указать процедуру перебора, выполненную при помощи компьютера, которая последовательно испытывает клики одну за другой. В частности, если разложение содержит клики из 2–3 элементов, то перебор легко выполним.

Например, для двух длинных английских текстов (см. рис. 110) клики сопоставляются на рис. 111.

В этом случае грубое дешифрование было бы все же достаточно хорошим, потому что клики не слишком налагаются и существуют зазоры между /a/ и /n/, /s/ и /h/, /c/ и /m/. В таких случаях достаточно лишь нескольких попыток перебора.

{e}	{t}	{a}	{noi}	{rs}	{h}	{dl}	{c}	{mufp}	{ygw}	{b}	{v}	{k}	{xzqj}
{e}	{t}		{aoin}		{srh}	{ld}		{cumf}	{pgwyb}		{vk}		{xjqz}

Рис. 112. Сопоставление клик

15.5.3. Пример. Для короткого криптотекста Бёлля из разд. 15.3.1 с частотами, приведенными на рис. 103, никакое хорошее разложение на клики работать не будет. Частоты 54 H и 41 Q указывают на $\mathrm{H} \mathrel{\hat{=}} \mathrm{e}$, и $\mathrm{Q} \mathrel{\hat{=}} \mathrm{n}$, а частоты 36 D, 24 W, 23 L, 22 V и 21 K показывают, что нельзя ожидать разделения клик {i r} и {s a t}. С другой стороны, D хорошо отделяется, и представляется разумным положить $\mathrm{D} \mathrel{\hat{=}} \mathrm{i}$. Это позволило бы сопоставить {r s a t} с {WLVK} и ограничиться $4! = 24$ опробованиями. К несчастью, ни одно из них не приводит к осмысленным текстам. В действительности, следующие две частоты 19 U и 18 X так близки, что следует рассмотреть клику {d h u}. Но это означает, что должно быть сделано $8! = 40\,320$ опробований, что является верхней границей числа переборов.

{H}	{Q}	{D}	{LUVWKX}	{GOJFPEIN}	{RZCS}	{YMBAT}	
{e}	{n}	{ir}	{ast}	{hud}	{lgo}	{cm}	{bfwkz}	{pv}	{jyxq}

Рис. 113. Сопоставление клик для криптотекста Бёлля из разд. 15.3.1

Сопоставление клик показано на рис. 112. Очевидно, что для коротких текстов механическое дешифрование на основе частот индивидуальных букв не работает. По крайней мере, требуется принять во внимание другие стохастические особенности языка, например, частоты биграмм. Это будет изучено в разд. 15.7.

15.5.4. Эмпирические частоты. Для английского языка табл. 8 дает эмпирические относительные частоты $\mu_i = m_i/M$, это результат подсчета Мейера—Матиаша, основанного на обработке 4 миллионов символов повседневного языка Англии. Кульбак указал в 1976 г., что жанр сообщений приводит к возникновению сильных колебаний в частоте символов и отделил «литературную Англию» с частотой (относительной) 12, 77% для /e/ от «телеграфной Англии» с частотой 13, 19% для /e/.

Для немецкого языка текстовая база SZ3-92 с полным массивом из $M = 681\,972$ символов дает результаты, подробно представленные в табл. 8.

Числовые значения в табл. 8 могут служить в качестве гипотетического вероятностного распределения стохастической выборки. Для немецкого языка

символ	Английский	Немецкий	символ	Английский	Немецкий
a	8.04%	6.47%	n	7.09%	9.84%
b	1.54%	1.93%	o	7.60%	2.98%
c	3.06%	2.68%	p	2.00%	0.96%
d	3.99%	4.83%	q	0.11%	0.02%
e	12.51%	17.48%	r	6.12%	7.54%
f	2.30%	1.65%	s	6.54%	6.83%
g	1.96%	3.06%	t	9.25%	6.13%
h	5.49%	4.23%	u	2.71%	4.17%
i	7.26%	7.73%	v	0.99%	0.94%
j	0.16%	0.27%	w	1.92%	1.48%
k	0.67%	1.46%	x	0.19%	0.04%
l	4.14%	3.49%	y	1.73%	0.08%
m	2.53%	2.58%	z	0.09%	1.14%

Таблица 8. Гипотетические вероятности символов стохастической выборки из английского и немецкого текстов

частота /e/ искажена в связи с криптографическим обычаем разлагать /ä/, /ö/, /u/ соответственно на /ae/, /oe/, /ue/.

Ципф и Мандельброт опубликовали эмпирические формулы для относительной частоты k-й буквы, которые удивительно подходят ко многим языкам, а именно:

$$p(k) \propto 1/k \quad \text{и} \quad p(k) \propto 1/(k+c)^m$$

для подходящих положительных c и m.

Фактические значения относительных частот символов для английского языка показаны графически на рис. 113. Убедительных теоретических объяснений этого распределения нет.

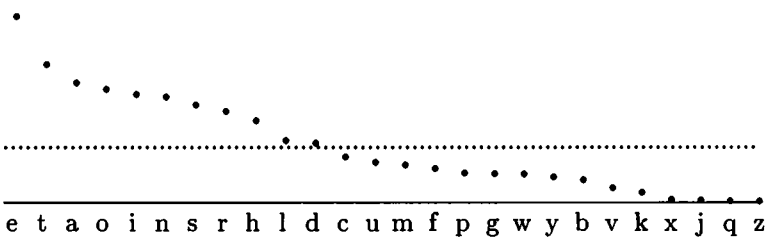


Рис. 114. Относительные частоты символов английского языка (подсчет Мейера—Матиаша)

15.6. Оптимальное соответствие

15.6.1. Квадрат расстояния. Отклонение частот символов двух текстов одинаковой длины — данным текстом T с M символами и текстом T^Q, порожденным стохастическим источником Q — может быть измерено квадратом расстояния $d(T, T^Q)$:

$$d(T, T^Q) = \sum_{i=1}^{N} (m_i - M \cdot p_i)^2.$$

Здесь p_i означает вероятность появления i-го символа χ_i $(i = 1, \ldots, N)$, m_i — частоту χ_i в тексте T, где $m_1 + \ldots + m_N = M$.

Пусть T — заданный криптотекст, а σ — дешифрующая (точнее, расшифровывающая) подстановка, преобразующая символы криптотекста T в символы открытого текста. Значение величины

$$d_\sigma = d(T, T^{\sigma(Q)}) = \sum_{i=1}^{N} (m_i - M \cdot p_{\sigma(i)})^2$$

измеряет соответствие между криптотекстом T с наблюдаемыми частотами m_i и ожидаемым криптотекстом $T^{\sigma(Q)}$ источника $\sigma(Q)$; величина

$$\min_\sigma d_\sigma = \min_\sigma \sum_{i=1}^{N} (m_i - M \cdot p_{\sigma(i)})^2$$

характеризует подстановки, достигающие оптимального соответствия, которые поэтому являются кандидатами для дешифрования. Из-за колебаний подстановки, приводящие величину d_σ близко к минимуму, тоже являются кандидатами, но с ростом d_σ данные подстановки становятся все менее и менее интересными.

15.6.2. Минимизация. Очевидно, $\sum\limits_{i=1}^{N} p_{\sigma(i)}^2 = \sum\limits_{i=1}^{N} p_i^2$, поэтому

$$\min_\sigma \sum_{i=1}^{N} (m_i - M \cdot p_{\sigma(i)})^2 = \sum_{i=1}^{N} m_i^2 + M^2 \sum_{i=1}^{N} p_i^2 - 2M \max_\sigma \sum_{i=1}^{N} m_i \cdot p_{\sigma(i)}.$$

Поэтому чтобы найти кандидатов для дешифрования, достаточно рассмотреть следующий максимум:

$$\max_\sigma \sum_{i=1}^{N} m_i \cdot p_{\sigma(i)}.$$

Теорема. Допустим, что $m_i \geqslant m_{i+1}$ для всех i. Тогда сумма $\sum\limits_{i=1}^{N} m_i \cdot p_{\sigma(i)}$ максимальна только тогда, когда $p_{\sigma(i)} \geqslant p_{\sigma(i+1)}$ для всех i.

Доказательство. Поскольку каждую подстановку можно представить в виде последовательности перестановок пар элементов, то достаточно исследовать вклад перестановки двух элементов χ_j, χ_k, где $j < k$, в эту сумму. Видим, что

$$m_j \cdot p_{\sigma(j)} + m_k \cdot p_{\sigma(k)} \geqslant m_j \cdot p_{\sigma(k)} + m_k \cdot p_{\sigma(j)}$$

только в том случае, когда

$$(m_j - m_k) \cdot (p_{\sigma(j)} - p_{\sigma(k)}) \geqslant 0,$$

т. е. тогда и только тогда, когда $p_{\sigma(j)} \geqslant p_{\sigma(k)}$, так как $m_j \geqslant m_k$. Отсюда легко вытекает требуемое утверждение.

Ожидаемый результат, что оптимальное соответствие σ достигается соответствием в упорядочении частот (т. е. что частотному упорядочению

$$m_{i_1} \geqslant m_{i_2} \geqslant \ldots \geqslant m_{i_N}$$

соответствует упорядочение вероятностей, переставленных по подстановке σ:

$$p_{\sigma(i_1)} \geqslant p_{\sigma(i_2)} \geqslant \cdots \geqslant p_{\sigma(i_N)}),$$

дополняется стратегией нахождения других кандидатов путем перестановок пар символов χ_j, χ_k так, чтобы всякий раз произведение

$$(m_j - m_k) \cdot (p_{\sigma(j)} - p_{\sigma(k)})$$

было минимально.

15.6.3. Пример. В предположении, что вероятности p_i заданы немецкой частью табл. 8, исследуем криптотекст Бёлля из разд. 15.3.1 (частоты m_i указаны на рис. 103). Применяя дешифрование σ, соответствующее порядку частот табл. 8, получаем для $\sum_{i=1}^{N} m_i \cdot p_{\sigma(i)}$ значение

$$2634.56\% = M \cdot 7.5489\%.$$

Проведя перестановку пары символов D $\hat{=}$ i, W $\hat{=}$ r, т. е. положив D $\hat{=}$ r, W $\hat{=}$ i (это соответствует порядку частот Кэдинга), получим, что сумма уменьшилась на

$$(36 - 24) \cdot (7.73\% - 7.54\%) = 2.28\% = M \cdot 0.0065\%$$

и стала равной

$$2632.28\% = M \cdot 7.5424\%.$$

Для правильного дешифрования получаем ее значение

$$2585.80\% = M \cdot 7.4092\%.$$

Значение суммы $\sum\limits_{i=1}^{N} m_i^2$ в примере равно

$$9347\% = M^2 \cdot 7.67\%,$$

а значение суммы $\sum\limits_{i=1}^{N} p_i^2$, соответствующее табл. 8, равно 7, 62%.

15.7. Частоты мультиграмм

Даже больше, чем частоты индивидуальных символов, в языке закреплены частоты биграмм. Их значения характеризуются следующей теоремой.

Теорема инвариантности 3^n. Для всех одноалфавитных функциональных простых подстановок, в частности, для всех одноалфавитных линейных простых подстановок (включая сложение ЦЕЗАРЯ и реверсивности) — *частоты n-грамм внутри текста инвариантны.*

15.7.1. Частоты мультиграмм. Согласно приведенной выше теореме, частоты n-грамм криптотекста можно использовать для его дешифровки. Однако для $N = 26$ существует 676 биграмм и 17576 триграмм; лишь в довольно длинных криптотекстах можно найти достаточное число биграмм и триграмм, а в коротких текстах даже биграммы являются большой редкостью, поэтому влияние колебаний значительно. Криптоанализ односимвольных шифров на основе лишь биграмм вместо отдельных символов не дает большого преимущества.

	.a	.b	.c	.d	.e	.f	.g	.h	.i	.j	.k	.l	.m	.n	.o	.p	.q	.r	.s	.t	.u	.v	.w	.x	.y	.z
a.	1	32	39	15		10	18		16		10	77	18	172	2	31	1	101	67	124	12	24	7		27	1
b.	8				58				6	2		21	1		11			6	5		25				19	
c.	44		12		55	1		46	15		8	16			59	1		7	1	38	16		1			
d.	45	18	4	10	39	12	2	3	57	1		7	9	5	37	7	1	10	32	39	8	4	9		6	
e.	65	11	64	107	39	23	20	15	40	1	2	46	43	120	46	32	14	154	145	80	7	16	41	17	17	
f.	21	2	9	1	25	14	1	6	21	1		10	3	2	38	3		4	8	42	11	1	4		1	
g.	11	2	1	1	32	3	1	16	10			4	1	3	23	1		21	7	13	8		2		1	
h.	84	1	2	1	251	2		5	72			3	1	2	46	1		8	3	22	2		7		1	
i.	18	7	55	16	37	27	10				8	39	32	169	63	3		21	106	88		14	1	1		4
j.					2										4						4					
k.					28				8					3	3				2	1			3		3	
l.	34	7	8	28	72	5	1		57	1	3	55	4	1	28	2	2	2	12	19	8	2	5		47	
m.	56	9	1	2	48			1	26				5	3	28	16			6	6	13		2		3	
n.	54	7	31	118	64	8	75	9	37	3	3	10	7	9	65	7		5	51	110	12	4	15	1	14	
o.	9	18	18	16	3	94	3	3	13		5	17	44	145	23	29		113	37	53	96	13	36		4	2
p.	21	1			40			7	8			29			28	26	42	3	14	7		1		2		
q.																					20					
r.	57	4	14	16	148	6	6	3	77	1	11	12	15	12	54	8		18	39	63	6	5	10		17	
s.	75	13	21	6	84	13	6	30	42		2	6	14	19	71	24	2	6	41	121	30	2	27		4	
t.	56	14	6	9	94	5	1	315	128			12	14	8	111	8		30	32	53	22	4	16		21	
u.	18	5	17	11	11	1	12	2	5			28	9	33	2	17		49	42	45				1	1	1
v.	15				53				19						6											
w.	32		3	4	30	1		48	37			4	1	10	17	2		1	3	6	1	1	2			
x.	3		5		1				4						1	4				1	1					
y.	11	11	10	4	12	3	5	5	18			6	4	3	28	7		5	17	21	1	3	14			
z.					5				2			1														1

Таблица 9. Частоты биграмм английского языка (в %%) по Микеру

Частоты биграмм (как видно из табл. 9 и 10) и триграмм еще более несбалансированы, чем частоты единичных символов. 19 наиболее частых биграмм английского языка и 18 наиболее частых биграмм немецкого языка (они составляют 92,93% всех биграмм) приведены в таблицах 11 и 12; 98 наиболее частых триграмм английского языка и 112 наиболее частых триграмм немецкого языка (они составляют около 52,11% всех триграмм) представлены в табл. 13 и 14.

	.a	.b	.c	.d	.e	.f	.g	.h	.i	.j	.k	.l	.m	.n	.o	.p	.q	.r	.s	.t	.u	.v	.w	.x	.y	.z
a.	8	31	27	11	64	15	30	20	5	1	7	59	28	102		4		51	53	46	75	2	3		1	2
b.	16	1		1	101		3	1	12		1	9		1	8			9	6	4	14		1		1	1
c.	2			2	1			242	1		14	1			2				1							
d.	54	3	1	13	227	3	4	2	93	1	3	5	4	6	9	3		10	11	6	16	3	4			3
e.	26	45	25	51	23	26	50	57	193	3	19	63	55	400	6	13	1	409	140	55	36	14	23	2	1	11
f.	19	2		9	25	12	3	1	7		1	5	1	2	9	1		18	4	20	24	1	1			1
g.	20	3		12	147	2	3	3	19	1	3	9	3	5	6	1		14	18	18	11	4	3			3
h.	70	4	1	14	102	2	4	3	23	1	3	25	11	19	18	1		37	11	47	11	4	9			3
i.	7	7	76	20	163	5	38	12	1	1	12	25	27	168	20	2		17	79	78	3	5	1			5
j.	9				9										2						5					
k.	26	1		2	26	1	1	1	7		1	10	1	1	24	1		13	5	14	9	1	1			1
l.	45	7	2	14	65	5	6	2	61	1	7	42	3	4	14	2		2	22	27	13	3	2			3
m.	40	6	1	8	50	4	4	3	44	2	3	4	23	3	15	7		2	10	8	14	4	3			2
n.	68	23	5	187	122	19	94	17	65	5	25	10	23	43	18	10		10	74	59	33	18	29			25
o.	3	8	15	7	25	6	5	9	1	1	3	31	17	64	1	6		50	19	9	3	3	7		1	6
p.	16			3	10	6		2	4			4			11	5		23	1	3	4					
q.																					2					
r.	80	25	9	67	112	18	27	19	52	4	23	18	20	31	30	9		15	54	49	48	12	17			14
s.	36	10	89	20	99	7	13	9	65	2	11	9	12	7	28	22		8	76	116	15	9	10		2	7
t.	57	8	1	35	185	5	10	14	59	2	4	11	9	9	15	3		31	50	23	26	8	21		1	26
u.	3	8	16	5	78	27	8	4	2		3	7	21	119		5		33	48	23	1	3	2			1
v.	3				37				9						43											
w.	34				48				36	1				1	17				1		9					
x.									1						1					1						
y.					1							1	1						1							
z.	4	1		1	28		1		11		1	2	1		2				1	7	43	1	9			1

Таблица 10. Частоты биграмм немецкого языка (в %%) на базе текста SZ3-92

Сравнивая значения, приводимые в литературе, важно знать, принимается ли во внимание пробел между словами; иногда (например, у Пратта, 1939 г.) подсчитываются лишь биграммы и триграммы внутри слов. Эти подсчеты показывают даже большие колебания, чем подсчеты частот отдельных символов, как можно видеть из табл. 11 и 12. Частотные таблицы для некоторых других индоевропейских языков были опубликованы Гэйнс и Эйраудом.

Одного взгляда на табл. 9 и 10 достаточно, чтобы убедиться в том, что матрицы частот биграмм не симметричны. Широко распространены биграммы с редкими обращениями (по-немецки Dreher) вроде /th/, /he/, /ea/, /nd/, /nt/, /ha/, /ou/, /ng/, /hi/, /eo/, /ft/, /sc/, /rs/; они полезны для разруше-

	Таблица 9	Кульбак	Синков	Эйрауд
th	315	156	270	330
he	251	40	257	270
an	172	128	152	167
in	169	150	194	202
er	154	174	179	191
re	148	196	160	169
on	145	154	154	134
es	145	108	115	149
ti	128	90	108	126
at	124	94	127	127
st	121	126	103	116
en	120	222	129	146
or	113	128	108	91
nd	118	104	95	122
to	111	100	95	79
nt	110	164	93	124
ed	107	120	111	125
is	106	70	93	79
ar	101	88	96	83

Таблица 11. Девятнадцать наиболее частых английских биграмм (частоты в %%)

	Таблица 10	Bauer-Goos	Valerio	Эйрауд
er	409	340	337	375
en	400	447	480	443
ch	242	280	266	280
de	227	214	231	233
ei	193	226	187	242
nd	187	258	258	208
te	185	178	222	178
in	168	204		197
ie	163	176	222	188
ge	147	168	160	196
es	140	181		168
ne	122	117		143
un	119	173	169	139
st	116	124		118
re	112	107	213	124
he	102	117		124
an	102	92		82
be	101	96		104

Таблица 12. Восемнадцать наиболее частых немецких биграмм (частоты в %%)

the 353	hat 55	man 40	ant 32	rom 28	str 25	nte 23
ing 111	ers 54	red 40	hou 31	ven 28	tic 25	rat 23
and 102	his 52	thi 40	men 30	ard 28	ame 24	tur 23
ion 75	res 50	ive 38	was 30	ear 28	com 24	ica 23
tio 75	ill 47	rea 38	oun 30	din 27	our 24	ich 23
ent 73	are 47	wit 37	pro 30	sti 27	wer 24	nde 23
ere 69	con 46	ons 37	sta 30	not 27	ome 24	pre 23
her 68	nce 45	ess 36	ine 29	ort 27	een 24	enc 22
ate 66	all 44	ave 34	whi 28	tho 26	lar 24	has 22
ver 64	eve 44	per 34	ove 28	day 26	les 24	whe 22
ter 63	ith 44	ect 33	tin 28	ore 26	san 24	wil 22
tha 62	ted 44	one 33	ast 28	but 26	ste 24	era 22
ati 59	ain 43	und 33	der 28	out 25	any 23	lin 22
for 59	est 42	int 32	ous 28	ure 25	art 23	tra 22

Таблица 13. 98 наиболее частых английских триграмм (частоты в %%)

ein 122	das 47	erd 33	ese 27	eni 23	ner 20	hei 18
ich 111	hen 47	enu 33	auf 26	ige 23	nds 20	lei 18
nde 89	ind 46	nen 32	ben 26	aen 22	nst 20	nei 18
die 87	enw 45	rau 32	ber 26	era 22	run 20	nau 18
und 87	ens 44	ist 31	eit 26	ern 22	sic 20	sge 18
der 86	ies 44	nic 31	ent 26	rde 22	enn 19	tte 18
che 75	ste 44	sen 31	est 26	ren 22	ins 19	wei 18
end 75	ten 44	ene 30	sei 26	tun 22	mer 19	abe 17
gen 71	ere 43	nda 30	and 25	ing 21	rei 19	chd 17
sch 66	lic 42	ter 30	ess 25	sta 21	eig 18	des 17
cht 61	ach 41	ass 29	ann 24	sie 21	eng 18	nte 17
den 57	ndi 41	ena 29	esi 24	uer 21	erg 18	rge 17
ine 53	sse 39	ver 29	ges 24	ege 20	ert 18	tes 17
nge 52	aus 36	wir 29	nsc 24	eck 20	erz 18	uns 17
nun 48	ers 36	wie 28	nwi 24	eru 20	fra 18	vor 17
ung 48	ebe 35	ede 27	tei 24	mme 20	hre 18	dem 17

Таблица 14. 112 наиболее частых немецких триграмм (частоты в %%)

ния клик. С другой стороны, следующие пары биграмм показывают примерно одинаковую частоту: /er/-/re/, /es/-/se/, /an/-/na/, /ti/-/it/, /on/-/no/, /in/-/ni/, /en/-/ne/, /at/-/ta/, /te/-/et/, /or/-/ro/, /to/-/ot/, /ar/-/ra/, /st/-/ts/, /is/-/si/, /ed/-/de/, /of/-/fo/.

15.7.2. Частоты слов. Очень интересны частоты слов, т. е. мультиграмм с пробелами в начале и конце. Порядок наиболее частых слов такой:

в английском языке:

the, of, and, to, a, in, that, it, is, I, for, as, with, was, his, he, be, not, by, but, have, you, which, are, on, or, her;

в немецком языке:

die, der, und, den, am, in, zu, ist, daß, es;

во французском языке:

de, il, le, et, que, je, la, ne, on, les, en, ce, se, son, mon, pas, lui, me, au, une, des, sa, qui, est, du;

в итальянском языке:

la, di, che, il, non, si, le, una, lo, in, per, un, mi, io, piu, del, ma, se;

в испанском языке:

de, la, el, que, en, no, con, un, se, su, las, los, es, me, al, lo, si, mi, una, di, por, sus, muy, hay, mas.

Однобуквенными словами в английском языке являются только a и I; двухбуквенными:

an, at, as, be, in, is, it, on, or, to, of, do, go, no, so, my.

Самыми частыми словами во многих индоевропейских языках являются бессодержательные слова[1)] (фр. *mots vides*, нем. *Fortwörter, inhaltsleere Wörter*), а именно, артикли, приставки, союзы и другие вспомогательные частицы — в противоположность словам, имеющим осмысленное значение (нем. *Begriffswörter*) — существительным, прилагательным и глаголам. 70 из наиболее частых слов английского языка являются бессодержательными, а среди 100 наиболее частых слов лишь 10 являются осмысленными словами.

15.7.3. Позиции. Частоты букв нередко зависят от их позиции внутри слова. Например, для буквы /e/ в немецком языке эти частоты таковы:

на первой позиции	7,7%
на второй позиции	21,7%
на третьей позиции	16,5%
⋮	
на третьей позиции от конца	8,8%
на второй предпоследней позиции	7,7%
на последней позиции	15,0%

15.7.4. Средняя длина слова. Хотя в криптографии пробелы между словами профессионально запрещены, средняя длина слова является важной характеристикой языка (табл. 15). В немецком языке длины слов распределены следующим образом:

1 0.05%	6 11.66%	11 3.24%	16 0.32%
2 8.20%	7 6.04%	12 2.06%	17 0.38%
3 28.71%	8 4.43%	13 1.40%	18 0.16%
4 13.49%	9 3.67%	14 0.59%	19 0.10%
5 11.55%	10 2.64%	15 0.65%	

[1)] В английском языке бессодержательные слова это единственные слова, которые в заголовках пишутся со строчной буквы.

15.7.5. Строение слов. Гласные и согласные буквы обычно чередуются. В любом языке гласные обеспечивают напевный звуковой строй. Во французском языке они могут встречаться целыми группами: ouïe, aïeul. Согласные в арабских языках образуют основу письма и встречаются группами, как и в славянских языках: czyszczenie (польский), cvŕčak (сербско-хорватский), nebezpečenství (чешский). У валлийцев встречаются странные образы:

Llanfairpwllgwyngyllgogerychwyrndrobwllllantysiliogogogoch

— название железнодорожной станции в Уэльсе, в rhy ddrwg («чересчур плохом» месте), y и w обозначают гласные.

В английском языке слова с четырьмя согласными подряд типа sixths довольно редки, в немецком языке Schlacht, schlecht, schlicht, Schlucht — 8-буквенные слова с единственной гласной, и слова типа Erstschlag с группой из 7 подряд идущих согласных можно получить соединением двух слов. Расстояния между гласными тоже выявляют типичные частоты: в немецком языке (без учета пробелов) они таковы:

1	20.77%	5	2.63%
2	25,06%	6	1.03%
3	35.95%	7	0.15%
4	14.75%	8	0.03%

Таблица 15 дает сравнительные значения средней длины слова, частоты гласных, частоты преобладающих согласных {l n r s t} и редких букв для шести важнейших языков.

15.7.6. Пробелы. Для шифров, которые сохраняют пробелы между словами и, возможно, также пунктуацию, биграммные таблицы содержат также частоты для букв в начале и конце слова, а триграммные таблицы содержат частоты для биграмм в начале и конце слова и однобуквенных слов. Если пробелы между словами не подавляются, они также включаются в шифрование. Таким образом, пробел может стать самым частым символом. В немецком языке пробелы почти так же часты, как и /e/, тогда как в английском языке они гораздо более часты, чем /e/.

Язык	Средняя длина слова	Частота гласных	Частота {l n r s t}	Редкие буквы
Английский	4.5	40%	33%	j q x z
Французский	4.4	45%	34%	k w
Немецкий	5.9	39%	34%	j q x y
Итальянский	4.5	48%	30%	j k w x y
Испанский	4.4	47%	31%	k w
Русский	6.3	45%		

Таблица 15. Характеристики строения слов

Если, как в «аристократах», пробелы сохраняются, они шифруются тоже пробелами, и таким образом, один символ расшифровывается с самого начала. Это значительно упрощает вход. Опытный криптолог-любитель иногда может читать такие неформальные шифры с первого взгляда. «Нередко криптограмма, которая сохраняет деление на слова, может быть прочитана сразу же, невзирая на ее возможную краткость» (Элен Фуше Гэйнс, 1939 г.).

В профессиональной криптологии имеются достаточные основания для применения формальных шифров. Если по техническим причинам, как при телетайпной связи или с кодом ASCII, полезны специальные контрольные символы, то они должны применяться с осмотрительностью и не смешиваться с криптографическим процессом — хорошо подготовленный и ответственный шифровальщик отлично это знает.

15.8. Комбинированный метод частотного соответствия

При попытке механизировать дешифрование одноалфавитных подстановок, особенно в случае коротких текстов, разумно сочетать информацию о частоте индивидуальных символов, биграмм и, возможно, триграмм — в том смысле, что частоты биграмм принимаются в расчет, только если некоторая клика символов не допускает разделения при помощи одних монограммных частот, а частоты триграмм — если и частоты биграмм не в состоянии разделить клику. Маловероятно, что возможно применение сочетаний более чем из трех символов. И в том случае, когда при этом не используется вероятное слово, атака является атакой только криптотекста, которая не использует ничего, кроме предположения о лежащем в его основе естественном языке.

15.8.1. Пример. Для криптотекста из 280 символов (Кан, 1967 г.)

GJXXN	GGOTZ	NUCOT	WMOHY	JTKTA	MTXOB
YNFGO	GINUG	JFNZV	QHYNG	NEAJF	HYOTW
GOTHY	NAFZN	FTUIN	ZANFG	NLNFU	TXNXU
FNEJC	INHYA	ZGAEU	TUCQG	OGOTH	JOHOA
TCJXK	HYNUV	OCOHQ	UHCNU	GHHAF	NUZHY
NCUTW	JUWNA	EHYNA	FOWOT	UCHNP	HOGLN
FQZNG	OFUVC	NZJHT	AHNGG	NTHOU	CGJXY
OGHTN	ABNTO	TWGNT	HNTXN	AEBUF	KNFYO
HHGIU	TJUCE	AFHYN	GACJH	OATAE	IOCOH
UFQXO	BYNFG				

частотный подсчет букв дает следующие величины:

17	4	13	0	7	17	23	26	5	12	3	2	2	36	25	1	5	0	0	23	20	3	6	9	13	8
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Указаний о сдвиге частотного профиля нет, и гипотеза о сложении ЦЕЗАРЯ исключается коротким переборным тестом. Упорядочим частоты

36 26 25 23 23 20 17 17 13 13 12 9 8 7 6 5 5 4 3 3 2 2 1 0 0 0
N H O G T U A F C Y J X Z E W I Q B K V L M P D R S

Допустим, что в задаче речь идет об английском тексте. Сопоставление с порядком частот английского языка, данных Каном в 1967 г.:

e t a o n i r s h d l u c m p f y w g b v j k q x z

наводит на мысль (ввиду заметного убывания от 17 до 13) о кликах {e}, {t}, {a o n i r s}. Таким образом,

$$\mathrm{N} \hat{=} \mathrm{e}, \quad \mathrm{H} \hat{=} \mathrm{t} \quad \text{и} \quad \{\mathrm{OGTUAF}\} \hat{=} \{\mathrm{aonirs}\}.$$

Для разделения последней клики можно применить частоты биграмм. Таблица 16 дает нужный фрагмент биграммной таблицы для английского языка (из 1000 символов).

В табл. 17 посчитаны биграммы представленного текста.

В табл. 16 можно заметить, что символы /a/, /i/ и /o/ не встречаются парами (избегают контактов), за исключением биграммы /io/. Биграмма /io/ редка. В табл. 17 не встречаются пары символов O, U и A, только OA встречается дважды, тогда как AO не встречается вообще. Это указывает на то, что $\mathrm{O} \hat{=} \mathrm{i}$, $\mathrm{A} \hat{=} \mathrm{o}$, и, значит, $\mathrm{U} \hat{=} \mathrm{a}$.

Это предположение хорошо согласуется с тем, что OU становится биграммой /ia/, которая встречается несколько раз. Кроме того, биграмма NU, которая представляет /ea/, является частой биграммой, тогда как UN, которая отсутствует, представляет редкую биграмму /ae/. Таким образом, большая клика разбита, и остается только маленькая клика {G T F}={n r s}, которую можно легко перебрать.

	.e	.t	.a	.o	.n	.i	.r	.s	.h	.l	.d	.u	.c
e.	39	80	131	46	120	40	154	145	15	46	107	7	64
t.	94	53	56	111	8	128	30	32	315	12	9	22	6
a.	-	124	1	2	172	16	101	67	-	77	15	12	39
o.	3	53	9	23	145	13	113	37	3	17	16	96	18
n.	64	110	54	65	9	37	5	51	9	10	118	12	31
i.	37	88	18	63	169	-	21	106	-	39	16	-	55
r.	148	63	57	54	12	77	18	39	3	12	16	6	14
s.	84	121	75	71	19	42	18	41	30	6	6	30	21
h.	251	22	84	46	2	72	8	3	5	3	1	2	2
l.	72	19	34	28	1	57	2	12	-	55	28	8	8
d.	39	39	45	37	5	57	10	32	3	7	10	8	4
u.	11	45	18	2	33	5	49	42	2	28	11	-	17
c.	55	38	44	59	-	15	7	1	46	16	-	16	12

Таблица 16. Таблица биграмм для тринадцати наиболее частых букв английского языка

	.N	.H	.O	.G	.T	.U	.A	.F	.C	.Y	.J	.X	.Z
N.	-	1	-	5	4	5	5	7	1	-	-	1	3
H.	3	2	4	1	2	1	1	-	1	9	1	-	-
O.	-	4	-	4	7	1	2	1	2	-	-	-	-
G.	4	2	6	2	-	-	2	-	-	-	3	-	-
T.	1	4	1	-	-	3	3	-	1	-	1	3	1
U.	-	1	-	2	4	-	-	3	5	-	-	-	1
A.	1	1	-	-	2	-	-	4	1	-	1	-	1
F.	3	2	1	2	1	2	-	-	-	1	-	-	1
C.	2	1	4	1	-	1	-	-	-	-	2	-	-
Y.	7	-	3	-	-	-	1	-	-	-	1	-	-
J.	-	2	2	-	1	2	-	2	1	-	-	3	-
X.	3	-	2	-	-	1	-	-	-	1	-	1	-
Z.	3	1	-	1	-	-	1	-	-	-	1	-	-

Таблица 17. Таблица подсчета биграмм наиболее частых символов криптотекста из разд. 15.8.1

Рассмотрение пар гласных является особенностью английского языка. Термин «vowel solution method» [метод разрешения с помощью гласных (*англ.*)], который встречается в английской литературе (Элен Фуше Гэйнс, 1939 г.), не является общепринятым и не описывает какой-либо общий метод, так как в других языках гласные вовсе не имеют тенденции избегать контакта.

В английском языке (и в некоторых других) согласная /n/ тоже имеет контактные предпочтения: /n/ регулярно предшествует гласной. Это соображение делает кандидатом на /n/ скорее T, чем G, F или что-нибудь из следующей клики {CY}. Итак, можно предположить, что

$$\mathrm{T} \mathrel{\hat{=}} \mathrm{n} \quad \text{и} \quad \{\mathrm{GF}\} \mathrel{\hat{=}} \{\mathrm{rs}\}.$$

Другой удобный случай представляется с буквой /h/: /th/ является очень частой биграммой, часты также /he/ и /ha/. Остающиеся символы G, F и C не показывают таких контактов, и это приводит к предположению $\mathrm{Y} \mathrel{\hat{=}} \mathrm{h}$.

Действительно, HY (для /th/) очень частая биграмма, часты также YN (для /he/) и YO (для /ha/).

К данному моменту определены в порядке рабочей гипотезы семь из десяти наиболее частых символов (и один тест на клику $\{\mathrm{GF}\} \mathrel{\hat{=}} \{\mathrm{rs}\}$ должен будет решить вопрос еще о двух):

```
N H U A T O * * Y * * * * * * * * * * * * * * * *
e t a o n i r s h d l u c m p f y w g b v j k q x z
```

15.8.2. Продолжение примера. После этого входа, который был прогулкой пешком, дальнейшее решение будет продвигаться рысью. Действительно, частичное дешифрование

```
GJXXe  GGinZ  eaCin  WMith  JnKno  MnXiB
heFGi  GIeaG  JFeZV  QtheG  eEoJF  thinW
Ginth  eoFZe  FnaIe  ZoeFG  eLeFa  nXeXa
FeEJC  Ietho  ZGoEa  naCQG  iGint  Jitio
nCJXK  theaV  iCitQ  atCea  GttoF  eaZth
eCanW  JaWeo  Etheo  FiWin  aCteP  tiGLe
FQZeG  iFaVC  eZJtn  otegG  entia  CGJXh
iGtne  oBeni  nWGen  tenXe  oEBaF  KeFhi
ttGIa  nJaCE  oFthe  GoCJt  ionoE  IiCit
aFQXi  BheFG
```

наводит на целую серию предположений: в первой строке Mith означает with, откуда M ≙ w; JnKnown означает unknown, откуда J ≙ u, K ≙ k. thinW во второй строке означает thing, откуда W ≙ g; в четвертой строке JethoZ означает method, откуда I ≙ m, Z ≙ d; это слово интуитивно подходит. Имеются несколько фрагментов, которые могут помочь найти оставшиеся буквы из клики {h d l c w u m}, а именно /1/ и /c/. Однако сначала мы должны сделать выбор между (G, F) = (r, s) и (G, F) = (s, r). Встречаемость FG во второй строке и тот факт, что биграмма /sr/ очень редка, дает

$$\text{G} \hat{=} \text{s} \quad \text{и} \quad \text{F} \hat{=} \text{r}.$$

Теперь мы имеем такую частичную расшифровку:

```
suXXe  ssind  eaCin  gwith  unkno  wnXiB
hersi  smeas  uredV  Qthes  eEour  thing
sinth  eorde  rname  doers  eLera  nXeXa
reEuC  metho  dsoEa  naCQs  isint  uitio
nCuXk  theaV  iCitQ  atCea  sttor  eadth
eCang  uageo  Etheo  rigin  aCteP  tisLe
rQdes  iraVC  edutn  otess  entia  CsuXh
istne  oBeni  ngsen  tenXe  oEBar  kerhi
ttsma  nuaCE  orthe  soCut  ionoE  miCit
arQXi  Bhers
```

В первой строке suXXess означает success, откуда X ≙ c, а deaCing означает dealing, откуда C ≙ l. В общем, мы теперь знаем все, кроме нескольких редких букв:

```
NHUATOFGYZCJXI***MW***K***
etaonirshdlucmpfywgbvjkqxz.
```

Полученный текст уже можно бегло прочесть:

succe	ssind	ealin	gwith	unkno	wnciB
hersi	smeas	uredV	Qthes	eEour	thing
sinth	eorde	rname	doers	eLera	nceca
reEul	metho	dsoEa	nalQs	isint	uitio
nluck	theaV	ilitQ	atlea	sttor	eadth
elang	uageo	Etheo	rigin	alteP	tisLe
rQdes	iraVl	edutn	otess	entia	lsuch
istne	oBeni	ngsen	tence	oEBar	kerhi
ttsma	nualE	orthe	solut	ionoE	milit
arQci	Bhers				

и продолжение нашей процедуры галопом дает почти автоматически следующие соответствия: B $\hat{=}$ p, V $\hat{=}$ b, Q $\hat{=}$ y, E $\hat{=}$ f. Правда, в третьей строке обескураживает фраза /rname doers eLera nceca/, но зато в шестой–седьмой строках мы читаем: /rigin altex tisve rydes/, т. е. P $\hat{=}$ x, L $\hat{=}$ v. Все действительно встретившиеся буквы определены и остается открытым лишь вопрос о буквах /j/, /q/ и /z/.

Во время этого галопа мы нашли три ошибки шифрования:

в третьей строке четвертая группа должна читаться ZBNFG;
в седьмой строке третья группа должна читаться NVJHT;
в восьмой строке первая группа должна читаться OGHYN.

15.8.3. Окончательный результат. Если это недостаточно убедило сомневающегося читателя, то мы можем дополнительно восстановить еще и пароль для полученной подстановки. Заметим, что пока за исключением трех пропущенных букв алфавитное упорядочение букв открытого текста дает результат:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
U	V	X	Z	N	E	W	Y	O	*	K	C	I	T	A	B	*	F	G	H	J	L	M	P	Q	*

Нельзя не заметить пароля NEWYORKCITY; он дополнительно определяет значения R $\hat{=}$ j, D $\hat{=}$ q и S $\hat{=}$ z символов, не встречающихся в тексте.

Расшифрованное сообщение в удобной для читателя форме, избавленной от трех ошибок шифрования (положения которых отмечены подчеркиванием) оказалось ценным соображением:

«Success in dealing with unknown ciphers is measured by these four things in the order named: perseverance, careful methods of analysis, intuition, luck. The ability at least to read the language of the original text is very desirable, but not essential». Such is the opening sentence of Parker Hitt's *Manual for the Solution of Military Ciphers.*

[«Успех работы с неизвестными шифрами определяется четырьмя факторами в названном порядке: упорство, точные методы анализа, интуиция, везение. Способность владеть языком оригинального текста очень желательна, но не обязательна». Таково вступительное предложение из книги Хитта *Руководство по дешифрованию военных шифров*.]

Полковник Паркер Хитт (1877–1971 гг.) опубликовал в 1916 г. одну из первых в США книг по криптологии и в ней впервые систематически разобрал дешифрование шифра ПЛЕЙФЕЙР (разд. 4.2.1). Позднее Хитт стал вице-президентом AT&T и президентом ее филиала International Communication Laboratories. Цитированное выше предложение Хитта констатирует, что семантическая поддержка не является решающей для успеха дешифрования, и оно было понято как поощрение механизации его трудоемкой части (чистого криптоанализа).

Заметим, что дешифрование было осуществлено, исходя только из соображений частоты, т. е. на основании теоремы 3 и теоремы $3^{(2)}$. Другие вспомогательные средства, например, нахождение шаблонов или использование вероятных слов не применялись. Еще о смешанных методах см. в разд. 15.9.

15.8.4. Окончательное соответствие. Правильное дешифрование показывает соответствие наблюдаемых частот биграмм и ожидаемых частот, основанных на вероятностях биграмм. Это наблюдение детализировано в табл. 18 для тринадцати наиболее частых букв после подходящих перестановок. Описанный метод можно интерпретировать как способ разбивания монограммных клик и сведение к комбинаторной задаче. Простая механическая процедура, эффективно находящая оптимальное соответствие, неизвестна.

15.8.5. Другой подход. Вместо поиска открытого текста, соответствующего криптотексту, иногда легче восстановить зашифрованный алфавит непосредственно, если известно, что он был порожден некоторым паролем. В предыдущем примере это восстановление можно было начать после того, как были найдены первые девять букв:

U * * * N * * Y O * * * * T A * * F G H * * * * * *
a b c d e f g h i j k l m n o p q r s t u v w x y z.

Имеется брешь из двух букв между A и F, в которую могут быть поставлены две буквы из {B,C,D,E}, тогда две оставшиеся буквы примут участие в формировании пароля. Это приводит к шести вариантам, которые надо перебрать; верный вариант такой:

U * * * N E * Y O * * C * T A B D F G H * * * * * *
a b c d e f g h i j k l m n o p q r s t u v w x y z.

Это эвристический метод, не имеющий достаточного теоретического основания.

	e	t	a	o	n	i	r	s	h	l	d	u	c
e	1.1	2.2	3.7	1.3	3.4	1.1	4.3	4.1	0.4	1.3	3.0	0.2	1.8
t	2.6	1.5	1.6	3.1	0.2	3.6	0.8	0.9	8.8	0.3	0.3	0.6	0.2
a	-	3.5	-	0.1	4.8	0.4	2.8	1.9	-	2.2	0.4	0.3	1.1
o	0.1	1.5	0.3	0.6	4.1	0.4	3.2	1.0	0.1	0.5	0.5	2.7	0.5
n	1.8	3.1	1.5	1.8	0.3	1.0	0.1	1.4	0.3	0.3	3.3	0.3	0.9
i	1.0	2.4	0.5	1.8	4.7	-	0.6	3.0	-	1.1	0.4	-	1.5
r	4.1	1.8	1.6	1.5	0.3	2.2	0.5	1.1	0.1	0.3	0.4	0.2	0.4
s	2.4	3.4	4.4	2.0	0.5	1.2	0.5	1.1	0.8	0.2	0.2	0.8	0.6
h	7.0	0.6	2.4	1.3	0.1	2.0	0.2	0.1	0.1	0.1	-	0.1	0.1
l	2.0	0.5	1.0	0.8	-	1.6	-	0.3	-	1.5	0.8	0.2	0.2
d	1.1	1.1	1.3	1.0	0.1	1.6	0.3	0.9	0.1	0.2	0.3	0.2	0.1
u	0.3	1.3	0.5	0.1	0.9	0.1	1.4	1.2	0.1	0.8	0.3	-	0.5
c	1.5	1.1	1.2	1.7	-	0.4	0.2	-	1.3	0.4	-	0.4	0.3

	N	H	U	A	T	O	F	G	Y	C	Z	J	X
N	-	1	5	5	4	-	7	5	-	1	3	-	1
H	3	2	1	1	2	4	-	1	9	1	-	1	-
U	-	1	-	-	4	-	3	2	-	5	1	-	-
A	1	1	-	-	2	-	4	-	-	1	1	1	-
T	1	4	1	-	-	3	3	-	1	-	1	3	1
O	-	4	1	2	7	-	1	4	-	2	-	-	-
F	3	2	2	-	1	1	-	2	1	-	1	-	-
G	4	2	-	2	-	6	-	2	-	-	-	3	-
Y	7	-	-	1	-	3	-	-	-	-	-	1	-
C	2	1	1	-	-	4	-	1	-	-	-	2	-
Z	3	1	-	1	-	-	-	1	-	-	-	1	-
J	-	2	2	-	1	2	2	-	-	1	-	-	3
X	3	-	1	-	-	2	-	-	1	-	-	-	1

Таблица 18. Ожидаемые частоты биграмм и наблюдаемые частоты после подходящего согласования символов

15.9. Частотное соответствие для многосимвольных подстановок

Многосимвольные подстановки можно рассматривать как простые подстановки, если m-граммы считать индивидуальными символами. Как бы то ни было, результатом является большой алфавит из N^m символов. Однако, из 676 биграмм стандартного английского языка реально используется (см. табл. 9) всего несколько сотен, а из 17 576 триграмм — не намного больше; m-граммы имеют сильно смещенное распределение частот, облегчающее незаконный вход.

15.9.1. Приводимый случай. Специальные биграммные подстановки имеют специфические методы решения. Тривиальным случаем является шифрование при помощи стандартной матрицы и перестановки номеров ее строк и столбцов:

	a	m	e	r	i	c	...
e	AA	AB	AC	AD	AE	AF	...
q	BA	BB	BC	BD	BE	BF	...
u	CA	CB	CC	CD	CE	CF	...
a	DA	DB	DC	DD	DE	DF	...
l	EA	EB	EC	ED	EE	EF	...
i	FA	FB	FC	FD	FE	FF	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Это можно свести (разд. 4.1.2) к односимвольному двухалфавитному шифру с периодом 2, который рассматривается в гл. 17.

15.9.2. Использование скрытой симметрии. Шифр ПЛЕЙФЕЙР, некогда широко используемый британской и германской армиями, не говоря уже о любителях, — является не только шифром ограниченной сложности, но обладает также скрытой тороидальной симметрией. Это имеет такое следствие: если некоторая биграмма открытого текста содержит букву X, то для зашифрования этой биграммы можно использовать всего лишь 8 букв, а именно отличные от X буквы строки и столбца, содержащих X:

P A L M E	**L M E P** A	U H **I** K Q
R **S** T O N	T O N R **S**	Z V **W** X Y
B **C** D F G	D F G B **C**	**E P A L M**
H **I** K Q U	K Q U H **I**	N R **S** T O
V **W** X Y Z	X Y Z V **W**	G B **C** D F

К тому же, шифрование обратной биграммы часто бывает обращением шифрования исходной — во всяком случае так бывает всегда, когда применяется «перекрестный шаг».

В то время как частоты биграмм при шифровании ПЛЕЙФЕЙР сохраняются, частоты индивидуальных символов — нет.

Интуитивные атаки против шифра ПЛЕЙФЕЙР основаны на частотах биграмм в связи с только что упомянутыми его особенностями. На практике используется также вероятное слово. Систематическое исследование впервые было начато в 1916 г. полковником Паркером Хиттом, в 1918 г. Ланги и в 1922 г. Смитом. Во время Второй мировой войны ПЛЕЙФЕЙР взламывали везде, где он использовался, модифицированный ПЛЕЙФЕЙР (разд. 4.2.2), применяемый Немецким африканским корпусом в качестве полевого шифра, был не лучше.

Незаконный дешифровальщик многосимвольного шифра как правило имеет значительные основания предполагать, что ему известна позиция пробела

мультиграмм. Но это может быть ошибкой в случае, когда сообщение, зашифрованное ПЛЕЙФЕЙРом, дополнено вначале нечетным числом пустышек. Тем не менее, это не слишком усложняет расшифрование, так как испытать два случая ненамного труднее, но сначала следует найти правильную идею.

15.10. Смешанные методы

Классификация и разделение методов, как это делается в настоящей книге, дает новое знание, которое необходимо для автоматизации вскрытия шифров. С другой стороны, совместное использование различных методов может значительно увеличить интенсивность атаки. Поэтому опытные криптоаналитики, работая «вручную», должны обязательно сочетать подходящие методы. Эта мысль подтверждается высказываниями как специалистов, подобных Базерье, Хитту, Фридману, так и одаренных воображением любителей вроде Бэббиджа.

15.10.1. Знаменитая криптограмма. Изящный пример вошел в мировую литературу. В 1843 г. Эдгар Аллан По (1809–1849 гг.) написал короткий детективный рассказ, «The Gold Bug» [Золотой жук], содержащий зашифрованное сообщение и его дешифрование. Алфавит представляет собой забавную мешанину из фигур и других символов, пригодных для печати — По был homme de lettres [литератор]. Криптотекст из 203 букв напоминал следующий[2]:

```
5 3 ‡ ‡ † 3 0 5 ) ) 6 * ; 4 8 2 6 ) 4 ‡ . ) 4 ‡ ) ; 8 0 6 * ; 4 8 † 8 ¶
6 0 ) ) 8 5 ; 1 ‡ ( ; : ‡ * 8 † 8 3 ( 8 8 ) 5 * † ; 4 6 ( ; 8 8 * 9 6 *
? ; 8 ) * ‡ ( ; 4 8 5 ) ; 5 * † 2 : * ‡ ( ; 4 9 5 6 * 2 ( 5 * - 4 ) 8 ¶
8 * ; 4 0 6 9 2 8 5 ) ; ) 6 † 8 ) 4 ‡ ‡ ; 1 ( ‡ 9 ; 4 8 0 8 1 ; 8 : 8 ‡
1 ; 4 8 † 8 5 ; 4 ) 4 8 5 † 5 2 8 8 0 6 * 8 1 ( ‡ 9 ; 4 8 ; ( 8 8 ; 4 (
‡ ? 3 4 ; 4 8 ) 4 ‡ ; 1 6 1 ; : 1 8 8 ; ‡ ? ;
```

Эдгар По предоставляет Леграну, герою рассказа, начать с замечания, что криптосистема (он называет ее «криптограмма») была адекватна умственным способностям капитана Кидда, «страшилы» этого рассказа, и, таким образом, непостижима для простого моряка, хоть она и не слишком сложна. Легран, который хвастает, что он может прочесть секретные сообщения и в тысячу раз более сложные, заключает, что, учитывая географическое положение, следовало бы рассмотреть французский или испанский языки, но, к счастью, подпись «Kidd» ясно указывала на английский. Он также заметил отсутствие пробелов, разделяющих слова, и выразил недовольство, что это сделает работу более трудной. Поэтому он начал с таблицы частот индивидуальных символов:

33	26	19	16	16	13	12	11	10	8	8	6	5	5	4	4	3	2	1	1
8	;	4	‡	)	*	5	6	(	†	1	0	9	2	:	3	?	¶	-	.

[2]В большом числе перепечаток и переводов эти шесть строк содержат много ошибок. Можно понять, как трудна работа наборщика, если он в своей работе не учитывает семантику.

Его первым предположением было 8 ≙ e, что подтверждалось частой встречаемостью удвоенного /e/ в английском языке — биграммный довод. Потом он нашел наиболее частую триграмму /the/ — шаблон 123 с 8 в конце. Он так же нашел семь появлений ; 48, и потому предположил, что ; ≙ t, 4 ≙ h.

Таким образом, был сделан большой шаг. Получен вход. Частично расшифрованная предпоследняя строка читается:

1 t h e † e 5 t h) h e 5 † 5 2 e e 0 6 * e 1 (‡ 9 t h e t (e e t h (

Сочетание thet(ee в конце строки навело Леграна на (≙ r. Это дало ему thetreethr‡?3hthe и догадку /thetreethroughthe/. Поэтому ‡ ≙ o, ? ≙ u, 3 ≙ g. Затем во второй строке †83(88, т. е. †egree, что намекает на /degree/ и † ≙ d, а четырьмя символами позже ; 46(; 88*, т. е. th6rtee*, что читается /thirteen/, откуда 6 ≙ i, и * ≙ n. Теперь почти все частые символы (кроме a и s) определены. Частичная расшифровка текста такова:

5 g o o d g 0 5)) i n t h e 2 i) h o .) h o) t e 0 i n t h e d e ¶
i 0)) e 5 t 1 o r t : o n e d e g r e e) 5 n d t h i r t e e n 9 i n
u t e) n o r t h e 5) t 5 n d 2 : n o r t h 9 5 i n 2 r 5 n – h) e ¶
e n t h 0 i 9 2 e 5) t) i d e) h o o t 1 r o 9 t h e 0 e 1 t e : e o
1 t h e d e 5 t h) h e 5 d 5 2 e e 0 i n e 1 r 0 9 t h e t r e e t h r
o u g h t h e) h o t 1 i 1 t : 1 e e t o u t

Сразу же Легран находит 5 ≙ a,) ≙ s, а также 0 ≙ 1, 2 ≙ b, . ≙ p, ¶ ≙ v, 1 ≙ f, : ≙ y, 9 ≙ m, – ≙ c.

Отображение, соответствующее этой одноалфавитной схеме шифрования, содержит 20 букв:

8 ; 4 ‡) * 5 6 († 1 0 9 2 : 3 ? ¶ - .
e t h o s n a i r d f l m b y g u v c p

и открытый текст в более приемлемой для чтения форме получается таким:

«A good glass in the Bishop's hostel in the Devil's seat — forty-one degrees and thirteen minutes — northeast and by north — main branch seventh limb east side — shoot from the left eye of the death's-head — a bee-line from the tree through the shot fifty feet out».

[Хорошее стекло на чертовом стуле двадцать один градус и тринадцать минут северо-северо-восток седьмая ветка восточная сторона стреляй из левого глаза мертвой головы прямая от дерева через выстрел на пятьдесят футов.]

15.10.2. Замечание. Заметим, что не было сказано ни одного слова, что этот шифр мог быть многоалфавитным. Эдгар По был одноалфавитно мыслящим человеком.

15.11. Снова расстояние единственности

Знание вероятностей n-грамм помогает понять, как при помощи перебора в разд. 12.7 происходит выборка «правильного» открытого текста и почему существует расстояние единственности. Маловероятная последовательность символов едва ли является «правильным» сообщением, но мы можем надеяться, что последовательность символов с вероятностью, близкой к единице, может быть «правильной». Расстояние единственности — это наименьшая длина текста, при которой только одно из возможных дешифрований имеет вероятность, близкую к единице, а вероятность других дешифрований близка к нулю.

	длина 1	длина 2	длина 3	длина 4	длина 5
V F K R Q	0.76	0.02			
W G L S R	2.03	0.04			
X H M T S	0.01	0.01			
Y I N U T	0.01	0.06	0.06		
Z J O V U	1.21	0.03	0.05		
A K P W V	5.96	1.88	0.01		
B L Q X W	1.77	2.35			
C M R Y X	3.17	0.03			
D N S Z Y	5.22	1.44	0.01		
E O T A Z	17.98	1.58	0.11	1.27	
F P U B A	1.23	0.19	0.03		
G Q V C B	3.25				
H R W D C	4.61	9.54	0.45		
I S X E D	7.97	20.30	0.01		
J T Y F E	0.06				
K U Z G F	1.12	2.34			
L V A H G	3.19	0.71	0.11		
M W B I H	2.47	0.86			
N X C J I	11.06	0.03			
O Y D K J	2.00	0.14	0.01		
P Z E L K	0.59	0.05			
Q A F M L	0.01				
R B G N M	6.42	6.38	0.05		
S C H O N	7.48	22.84	90.51	98.73	100.00
T D I P O	5.55	9.09	8.56		
U E J Q P	4.87	20.09	0.03		
	100.00	100.00	100.00	100.00	100.00

Таблица 19. Последовательное определение «правильного» открытого текста в соответствии с вероятностями n-грамм (в %)

Человек выделяет «правильный» открытый текст (пробегая список глазами), используя оптический и мысленный процессы распознавания. Аналогичная работа может быть выполнена при помощи статистического анализа.

Для примера из табл. 5, начиная с 6-го столбца, это показано в табл. 19. Вероятности мультиграмм определяются текстовой основой SZ3-92 и приводятся к 100%, пустые места означают вероятности ниже 0.005%. Ясно, что расстояние единственности в этом примере равно 5.

Этот перебор, однако, имеет свои пределы, если он приводит к десяткам тысяч проб, и не подходит для полного одноалфавитного шифрования, если нельзя использовать дополнительную информацию.

> Никакая одноалфавитная подстановка не является безопасной при передаче больших объемов информации.
>
> *Дэвид Кан*, 1967 *г.*

Глава 16

Каппа и Хи

Сообщение «Riverbank Publication № 22», написанное в 1920 г., когда Фридману было 28 лет, следует рассматривать как самую важную единичную публикацию по криптографии.

Дэвид Кан, 1967 *г.*

Поразительно, но для данного одноалфавитно зашифрованного криптотекста легче определить его язык — английский, французский или немецкий, чем дешифровать его. Это верно также для открытого текста: для достаточно длинного текста существует достоверный метод тестирования, позволяющий отнести этот текст к известному языку, не вникая в него, т. е. не рассматривая его грамматики и семантики, и существует родственный ему тест, чтобы решить, принадлежат ли два данных текста к одному и тому же языку без глубокого их изучения.

На самом деле существует один особый инвариант текста при одноалфавитном шифровании, который мы рассмотрим в этой главе, и один родственный инвариант пары текстов, который остается инвариантным даже при многоалфавитном шифровании обоих текстов с одним и тем же ключом. И эти инварианты имеют специфические значения, различающиеся между собой для большинства общих индоевропейских языков.

16.1. Определение и инвариантность Каппа

Зададим пару текстов $T = (t_1, t_2, t_3, \ldots, t_M)$ и $T' = (t'_1, t'_2, t_3, \ldots, t'_M)$ равной длины M над одним и тем же словарем Z_N.

Каппой пары текстов называется относительная частота нахождения в этих текстах одного и того же символа в одной и той же позиции (совпадение символов отмечаемое $*$) (Фридман, 1925 г., «index of coincidence [индекс совпадения, *англ.*]», часто сокращаемый IC). Таким образом,

$$Kappa(T, T') = \sum_{\mu=1}^{M} \delta(t_\mu, t'_\mu)/M,$$

где индикаторная функция («дельта-функция») определяется равенством

$$\delta(x, y) = \begin{cases} 1, & \text{если} \quad x = y, \\ 0 & \text{в противном случае.} \end{cases}$$

Пример 1. ($M = 180$).

```
T:  t h e p r e c e d i n g c h a p t e r h a s i n d i c a t e d h o w a m
T': w o u l d s e e m t h a t o n e w a y t o o b t a i n g r e a t e r s e
                *                               *     *
    o n o a l p h a b e t i c c i p h e r c a n b e s o l v e d e v e n i f
    c u r i t y w o u l d b e t o u s e m o r e t h a n o n e a l p h a b e
                                    *                 *
    t h e o r i g i n a l w o r d l e n g t h s a r e c o n c e a l e d a n
    t i n e n c i p h e r i n g a m e s s a g e t h e g e n e r a l s y s t
    *                       *               *   *   * *
    d t h e s u b s t i t u t i o n a l p h a b e t i s r a n d o m i t i s
    e m c o u l d b e o n e t h a t u s e s a n u m b e r o f d i f f e r e
                        *               *       *   *
    p o s s i b l e t o f i n d a s o l u t i o n b y u s i n g f r e q u e
    n t a l p h a b e t s f o r e n c i p h e r m e n t w i t h a n u n d e
                                                    *           *
```

Пример 2. ($M = 180$).

```
T:  e s t a u c h t v o n z e i t z u z e i t i m m e r w i e d e r e i n m
T': u n t e r s c h w e i z e r p o l i t i k e r n w a e c h s t d i e a n
      *             * *           *
    a l a u f u m k u r z d a r a u f e i l f e r t i g d e m e n t i e r t
    g s t d e n n a e c h s t e n z u g r i c h t u n g e g z u v e r p a s
                                                *
    z u w e r d e n d a s g e r u e c h t d a s s s i c h d i e o e l e x p
    s e n a u s s e n m i n i s t e r r e n e f e l b e r s a h s i c h j e
                                *
    o r t i e r e n d e n l a e n d e r v o m d o l l a r l o e s e n w o l
    t z t u e b e r r a s c h e n d e i n e r f o r d e r u n g a u s d e m
      *   *   *           * * * *         *       *
    l e n z u v e r d e n k e n w a e r e e s i h n e n f r e i l i c h n i
    s t a e n d e r a t a u s g e s e t z t e i n b e i t r i t t s g e s u
              * *               *       *     *     *
```

В примере 1 (английский язык) получаем результат $Kanna(T, T') = 17/180 = 9.44\%$, в примере 2 (немецкий язык) $Kanna(T, T') = 21/180 = 11.67\%$.

16.1.1. Очевидно,

$$Kanna(T, T') \leqslant 1,$$

при этом

$$Kappa(T, T') = 1 \quad \text{тогда и только тогда, когда} \quad T' = T.$$

Существует эмпирический результат, говорящий о том, что если достаточно длинные тексты (одной и той же длины) T и T' принадлежат одному и тому же языку $\mathcal{S}$ (или, лучше, одному и тому же жанру этого языка), то они имеют величину $Kappa(T, T')$, близкую к некоторому числу $\varkappa_{\mathcal{S}}$, в то время как $\varkappa_{\mathcal{S}}$ изменяется от языка к языку. В литературе даются следующие значения величины $\varkappa_{\mathcal{S}}$:

$\mathcal{S}$	N	Кульбак 1976	Эйрауд 1953
Английский	26	6.61%	6.75%
Немецкий	26	7.62%	8.20%
Французский	26	7.78%	8.00%
Итальянский	26	7.38%	7.54%
Испанский	26	7.75%	7.69%
Японский (Romaji)	26	8.19%	
Русский	32	5.29%	4.70%

В разных источниках встречаются разные значения $\varkappa_{\mathcal{S}}$: 6.5–6.9% для английского языка и 7.5–8.3% для немецкого. Для упомянутого в разд. 15.5.4 набора английских текстов значение $\varkappa_{\mathrm{e}} = 6.58\%$, а для немецких текстов из SZ3-92 значение $\varkappa_{\mathrm{d}} = 7.62\%$. Эти величины хорошо согласуются со значениями, данными Кульбаком. Для французского и испанского языков у него получаются очень близкие значения.

Возможно, что значения $\varkappa_{\mathcal{S}}$ для языка $\mathcal{S}$ отчасти отражают избыточность языков: перевод Евангелия от Святого Марка, составляет 30 000 слогов в английском, 36 000 во французском, в среднем 40 000 в тевтонских и в среднем 35 500 в славянских языках. Но строгой связи здесь нет.

16.1.2. Выделим два результата.

Теорема инвариантности 5. Для всех *многоалфавитных*, функциональных простых подстановок, в частности, для всех многоалфавитных линейных простых подстановок (включая сложения ВИЖЕНЕРА и вычитания БОФОРТА) *Каппа двух текстов равной длины, зашифрованных с одним и тем же ключом, является инвариантом.*

Теорема инвариантности 6. Для всех перестановок *Каппа двух текстов равной длины, зашифрованных с одним и тем же ключом, является инвариантом.*

16.1.3. Ожидаемое значение *Каппы* двух текстов равной длины M над одним и тем же словарем Z_N подсчитывается из вероятностей p_i и p'_i появления i-го символа в «стохастических источниках» Q и Q' этих текстов. Вероятность появления символа χ_i в μ-й позиции обоих текстов равна $p_i p'_i$, поэтому

вероятность одинаковых символов в μ-й позиции определяется формулой

$$\langle Kanna(T, T')\rangle_{QQ'} = \sum_{i=1}^{N} p_i \cdot p'_i.$$

Если оба источника идентичны, $Q' = Q$, то $p'_i = p_i$ и, следовательно,

$$\langle Kanna(T, T')\rangle_{Q} = \sum_{i=1}^{N} p_i^2. \qquad (*)$$

Это равенство соответствует определению *Kanna* при помощи классических урновых схем теории вероятностей.

Теорема. Для идентичных источников $Q' = Q$ имеем

$$\frac{1}{N} \leqslant \langle Kanna(T, T')\rangle_{Q} \leqslant 1.$$

Нижняя оценка достигается в случае равновероятного распределения Q_{R}: $p_i = 1/N$ для всех i, и только в этом случае; верхняя граница достигается для каждого детерминированного распределения Q_j: $p_j = 1$, $p_i = 0$ для $i \neq j$, и ни для какого другого распределения.

Как было сказано выше, для гипотетического вероятностного распределения в разд. 15.5.4, табл. 8,

$$\langle Kanna(T, T')\rangle_{\text{англ}} = 0.06577 = \varkappa_{\mathrm{e}},$$
$$\langle Kanna(T, T')\rangle_{\text{нем}} = 0.07619 = \varkappa_{\mathrm{d}}.$$

Для источника с равновероятным распределением Q_{R} ($N = 26$)

$$\langle Kanna(T, T')\rangle_{\mathrm{R}} = \varkappa_{\mathrm{R}} = 0.03846 = \frac{1}{26}.$$

Таким образом, *Kanna* тест ясно отличает английский и немецкий источники от источника с равновероятным распределением:

$$\varkappa_{\mathrm{d}}/\varkappa_{\mathrm{R}} = N \cdot \varkappa_{\mathrm{d}} = 1.98, \qquad \varkappa_{\mathrm{e}}/\varkappa_{\mathrm{R}} = N \cdot \varkappa_{\mathrm{e}} = 1.71.$$

Эмпирическое правило для этих языков состоит в следующем: *отношение* $\langle Kanna(T, T')\rangle_{\mathrm{S}}/\langle Kanna(T, T')\rangle_{\mathrm{R}}$ *близко к* 2.

16.2. Определение и инвариантность Хи

Снова зададим два текста равной длины M над одним и тем же словарем Z_N из N символов, $T = (t_1, t_2, t_3, \ldots, t_M)$, $T' = (t'_1, t'_2, t_3, \ldots, t'_M)$. Пусть m_i и m'_i обозначают частоты символа χ_i в текстах T и T' соответственно; тогда

$$\sum_{i=1}^{N} m_i = M, \qquad \sum_{i=1}^{N} m'_i = M.$$

Пусть Xu обозначает следующую нормализованную сумму произведений (Кульбак, 1935 г.):

$$Xu(T, T') = \left(\sum_{i=1}^{N} m_i \cdot m'_i\right)/M^2,$$

или в однородной форме

$$Xu\,(T, T') = \left(\sum_{i=1}^{N} m_i \cdot m'_i\right) \Big/ \left(\left(\sum_{i=1}^{N} m_i\right) \cdot \left(\sum_{i=1}^{N} m'_i\right)\right).$$

Два текста из примера 1 разд. 16.1 дают следующие частоты:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
T	15	6	8	9	21	3	4	10	17	0	0	8	2	14	13	6	1	8	11	14	5	2	2	0	1	0
T'	15	6	4	5	30	4	4	9	8	0	0	6	6	15	12	4	0	10	10	17	8	0	4	0	3	0

Получаем результат $Xu\,(T, T') = 2151/(180 \cdot 180) = 6.64\%$.
Для двух текстов из примера 2 разд. 16.1 частоты таковы:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
T	10	0	4	11	35	4	2	5	15	0	2	10	6	14	7	1	0	15	7	9	9	3	4	1	0	6
T'	11	3	6	6	33	2	7	7	12	1	1	2	2	16	2	2	0	15	18	16	10	1	2	0	0	5

Это приводит к $Xu\,(T, T') = 2492/(180 \cdot 180) = 7.69\%$.

16.2.1. По аналогии с разд. 16.1.1 получаем, что

$$Xu(T, T') \leqslant 1,$$

причем $Xu\,(T, T') = 1$ тогда и только тогда, когда T и T' образованы одним и тем же символом.

Если же все m_i равны, $m_i = M/N$ для всех i, то (для любых m'_i)

$$Xu(T, T') = \frac{1}{N} = \varkappa_{\mathrm{R}}.$$

Эмпирически установлено, что для достаточно длинных текстов из одного и того же языка S (и даже из одного и того же жанра этого языка) не только значение Xu довольно близко к некоторой величине, типичной для данного языка, но также эта величина близка к значению *Kanna* для этого языка. Этот факт прояснится в разд. 16.3.

16.2.2. Существует важный частный случай $T' = T$, $m'_i = m_i$. Пусть

$$\text{Пси}\,(T) = Xu\,(T, T) = \sum_{i=1}^{N} \frac{m_i^2}{M^2}.$$

Из теоремы Штейнера

$$\sum_{i=1}^{N} \left(m_i - \frac{M}{N}\right)^2 / M^2 = \sum_{i=1}^{N} \frac{m_i^2}{M^2} - \frac{1}{N}.$$

получаем, что

$$\frac{1}{N} \leqslant \text{Пси}\,(T) \leqslant 1,$$

причем $\text{Пси}\,(T) = 1$ тогда и только тогда, когда T образовано одним и тем же символом, а $\text{Пси}\,(T) = \text{Пси}\,(T) = 1/N = \varkappa_{\text{R}}$ тогда и только тогда, когда все m_i равны.

Более того, для экстремального случая $M \leqslant N$ имеет место неравенство

$$\frac{1}{M} \leqslant \text{Пси}\,(T),$$

причем $\text{Пси}\,(T) = 1/M$ тогда и только тогда, когда $0 \leqslant m_i \leqslant 1$ для всех i.

16.2.3. Хи и Пси обладают свойствами инвариантности.

Теорема инвариантности 7. Для всех одноалфавитных функциональных простых подстановок, в частности, для всех одноалфавитных линейных простых подстановок (включая сложение ВИЖЕНЕРА и вычитание БОФОРТА), *Хи двух текстов равной длины, зашифрованных с одним и тем же ключом, так же как и Пси одного текста, являются инвариантами.*

Теорема инвариантности 8. Для всех перестановок *Хи двух текстов равной длины, зашифрованных с одним и тем же ключом, так же как и Пси одного текста, являются инвариантами.*

Поскольку *Каппа*, *Хи* и Пси характеризуют язык, то этот язык можно определить из криптотекста.

16.2.4. Ожидаемое значение для *Хи* текста T из языка S и T' из языка S' длины M над одним и тем же словарем Z_N подсчитывается из вероятностей p_i и p'_i появлений i-го символа χ_i в стохастических источниках Q, Q' этих текстов: ожидаемое число появлений символа χ_i в тексте T равна $p_i \cdot M$, а в тексте T' равно $p'_i \cdot M$, что дает для $Xu\,(T, T')$ величину, равную

$$\langle Xu\,(T, T')\rangle_{QQ'} = \sum_{i=1}^{N} p_i \cdot p'_i.$$

Если два источника идентичны, $Q' = Q$, то $p'_i = p_i$ и

$$\langle Xu\,(T, T')\rangle_{Q} = \sum_{i=1}^{N} p_i^2. \qquad (*)$$

В частности,

$$\langle \text{Пси}\,(T)\rangle_{Q} = \sum_{i=1}^{N} p_i^2.$$

Теорема. Для идентичных источников $Q' = Q$ имеем

$$\frac{1}{N} \leqslant \langle Xu\,(T, T')\rangle_{Q} \leqslant 1, \qquad \frac{1}{N} \leqslant \langle \text{Пси}\,(T)\rangle_{Q} \leqslant 1.$$

Нижняя оценка здесь достигается для случая равновероятного распределения Q_R: $p_i = 1/N$ для всех i и только в этом случае; верхняя оценка достигается для каждого детерминированного распределения Q_j: $p_j = 1$, $p_i = 0$ для $i \neq j$, и ни для каких других распределений.

Удивительно, что величины $\langle Kanna(T, T')\rangle_Q$ и $\langle Xu(T, T')\rangle_Q$ (отмеченные символом (*)) совпадают. Ниже мы увидим, что между $Kanna(T, T')$ и $Xu(T, T')$ существует определенная связь.

16.3. Теорема Каппа-Хи

В дальнейшем нам понадобятся две вспомогательные функции $g_{i,\mu}$ и $g'_{i,\mu}$. Пусть T — некоторый текст и t_μ — μ-й символ этого текста. Пусть далее

$$g_{i,\mu} = \begin{cases} 1, & \text{если } t_\mu = \chi_i, \\ 0 & \text{в противном случае} \end{cases}$$

и пусть $g'_{i,\mu}$ определяется аналогичным образом для текста T'. Тогда

$$\delta(t_\mu, t'_\nu) = \sum_{i=1}^{N} g_{i,\mu} \cdot g'_{i,\nu} \quad \text{и} \quad \delta m_i = \sum_{i=1}^{N} g_{i,\mu}, \quad m'_i = \sum_{i=1}^{N} g'_{i,\nu}.$$

16.3.1. Пусть $T^{(r)}$ — текст T, циклически сдвинутый на r позиций вправо. Тогда число совпадений между $T^{(r)}$ и T' равно

$$Kanna(T^{(r)}, T') = \sum_{\mu=1}^{M} \delta(t_{(\mu-r-1) \bmod M+1}, t'_\mu)/M.$$

В частности, $Kanna(T^{(0)}, T') = Kanna(T, T')$.

16.3.2. Теперь сформулируем теорему.

Теорема Каппа-Хи.

$$\frac{1}{M} \sum_{\rho=0}^{M-1} Kanna(T^{(\rho)}, T') = Xu(T, T').$$

Таким образом, $Xu(T, T')$ есть среднее арифметическое всех $Kanna(T^{(r)}, T')$.

Следствие.

$$\frac{1}{M} \sum_{\rho=0}^{M-1} Kanna(T^{(\rho)}, T) = \text{Пси}(T).$$

Доказательство. Имеем:

$$\frac{1}{M}\sum_{\rho=0}^{M-1} Kanna\,(T^{(\rho)}, T') = \frac{1}{M}\cdot\frac{1}{M}\cdot\sum_{\rho=0}^{M-1}\sum_{\mu=1}^{M}\delta(t_{(\mu-\rho-1) \bmod M+1}, t'_\mu) =$$

$$= \frac{1}{M}\cdot\frac{1}{M}\cdot\sum_{\nu=1}^{M}\sum_{\mu=1}^{M}\delta(t_\mu, t'_\nu) = \frac{1}{M}\cdot\frac{1}{M}\cdot\sum_{\nu=1}^{M}\sum_{\mu=1}^{M}\sum_{i=1}^{N} g_{i,\mu}\cdot g'_{i,\nu} =$$

$$= \frac{1}{M}\cdot\frac{1}{M}\cdot\sum_{i=1}^{N}\sum_{\nu=1}^{M}\sum_{\mu=1}^{M} g_{i,\mu}\cdot g'_{i,\nu} = \frac{1}{M}\cdot\frac{1}{M}\cdot\sum_{i=1}^{N}\left(\sum_{\nu=1}^{M} g'_{i,\nu}\right)\cdot\left(\sum_{\mu=1}^{M} g_{i,\mu}\right) =$$

$$= \frac{1}{M}\cdot\frac{1}{M}\cdot\sum_{i=1}^{N} m'_i\cdot m_i = Xu\,(T, T').$$

Теперь становится понятным, почему в разд. 16.1 значения для *Kanna* с 9.44% и 11.67% (случайно) оказываются довольно хорошо сравнимыми со средними значениями 6.60 и 7.69 из разд. 16.2.

16.4. Теорема Каппа-Фи

Перейдем к случаю $T' = T$. Здесь мы сразу же замечаем, что среди всех значений $Kanna\,(T^{(r)}, T')$ выделяется «особый случай» $Kanna\,(T^{(0)}, T') = 1$. Этот случай нетипичен, поэтому в процессе усреднения его целесообразно исключить, находя среднее лишь для $M-1$ оставшихся случаев с $r \neq 0$:

$$\frac{1}{M-1}\sum_{\rho=1}^{M-1} Kanna\,(T^{(\rho)}, T).$$

16.4.1. Теперь получаем

$$\frac{1}{M-1}\cdot\sum_{\rho=1}^{M-1} Kanna\,(T^{(\rho)}, T) = \frac{1}{M-1}\cdot\left(\sum_{\rho=0}^{M-1} Kanna\,(T^{(\rho)}, T) - 1\right) =$$

$$= \frac{1}{M-1}\cdot(M\cdot \text{Пси}\,(T) - 1) = \frac{1}{M-1}\cdot\left(\sum_{i=1}^{N}(m_i^2/M - 1)\right) =$$

$$= \frac{1}{M-1}\cdot\frac{1}{M}\cdot\left(\sum_{i=1}^{N}(m_i^2 - M)\right) = \frac{1}{M-1}\cdot\frac{1}{M}\cdot\left(\sum_{i=1}^{N}(m_i^2 - m_i)\right) =$$

$$= \frac{1}{M-1}\cdot\frac{1}{M}\cdot\left(\sum_{i=1}^{N} m_i\cdot(m_i - 1)\right).$$

Таким образом, вводя новую величину

$$Фи\,(T) = \left(\sum_{i=1}^{N} m_i\cdot(m_i - 1)\right)\Big/(M\cdot(M-1)),$$

приходим к новой теореме.

Теорема Каппа-Фи.

$$\frac{1}{M-1}\sum_{\rho=1}^{M-1} Каппа(T^{(\rho)}, T) = Фи(T).$$

Вычисление $Фи(T)$ немного проще по сравнению с Пси(T), так как не только для случая $m_i = 0$, но и для $m_i = 1$ слагаемые в сумме нулевые. Это удобно для редких букв в коротких текстах. Заметим, что равенство $Фи(T) = 0$ выполняется тогда и только тогда, когда $m_i \in \{0, 1\}$. Но существует еще одна причина, почему люди в своей работе предпочитают иметь дело с $Фи$, а не с Пси: Кульбак был первым, кто предложил, используя подходящие вероятностные аргументы, применять $Фи$ вместо Пси.

Пример 3. Для криптотекста T ($M = 280$) из разд. 15.8.1 с частотами, установленными там, имеем:

$$280^2 \cdot \text{Пси}\,(T) = 289 + 16 + 169 + 0 + 49 + 289 + 529 + 676 + 25 + 144 + 9 + 4 + \\ + 4 + 1296 + 625 + 1 + 25 + 0 + 0 + 529 + 400 + 9 + 36 + 81 + 169 + 64 = 5438,$$

$$280 \cdot 279 \cdot \text{Пси}\,(T) = 272 + 12 + 156 + 0 + 42 + 272 + 506 + 650 + 20 + 132 + 6 + \\ + 2 + 2 + 1260 + 600 + 0 + 20 + 0 + 0 + 506 + 380 + 6 + 30 + 72 + 156 + 56 = 5158.$$

Таким образом, Пси(T)=5438/78400=6.936%, $Фи(T)$=5158/78120=6.603%.

Кроме того, для последовательности биграмм, т. е. для текста $T^{**} = T \times T^{(1)}$ получаем: Пси$(T^{**}) = 871/77841 = 1.119\%$, $Фи(T^{**}) = = 592/77562 = 0.763\%$.

16.4.2. $Фи(T)$ не очень сильно отличается от Пси (T):

$$\text{Пси}\,(T) = \frac{M-1}{M} Фи(T) + \frac{1}{M} = \text{Пси}\,(T) + \frac{1}{M}(1 - Фи(T)),$$

$$Фи(T) = \frac{M}{M-1}\text{Пси}\,(T) - \frac{1}{M-1} = \text{Пси}\,(T) + \Big(\frac{M}{M-1} - 1\Big)\text{Пси}\,(T) - \frac{1}{M-1},$$

$$\text{Пси}\,(T) - Фи\,(T) = \frac{1 - Фи(T)}{M} = \frac{1 - \text{Пси}\,(T)}{M-1},$$

так что

$$Фи(T) \leqslant \text{Пси}\,(T).$$

16.4.3. $Фи$ имеет те же свойства инвариантности, что и Пси.

Теорема инвариантности $7^{(Фи)}$. Для всех одноалфавитных, функциональных простых подстановок, в частности, для всех одноалфавитных линейных простых подстановок (включая сложение ВИЖЕНЕРА и вычитание БОФОРТА), *$Фи$ текста является инвариантом.*

Теорема инвариантности $8^{(\Phi и)}$. *Для всех перестановок Фи текста является инвариантом.*

16.4.4. Ожидаемое значение *Фи* для текста T длины M подсчитывается из вероятностей появления i-го символа χ_i в стохастическом источнике Q этого текста: значение $\Phi u(T)$ в зависимости от длины текста M равно

$$\langle \Phi u(T)\rangle_Q^{(M)} = \frac{M}{M-1}\cdot\left(\sum_{i=1}^{N} p_i\cdot\left(p_i-\frac{1}{M}\right)\right),$$

при этом

$$\langle \Phi u(T)\rangle_Q^{(M)} \geqslant \begin{cases} \dfrac{M}{M-1}\cdot\left(\dfrac{1}{N}-\dfrac{1}{M}\right)=\dfrac{1}{N}\cdot\dfrac{M-N}{M-1}, & \text{если} \quad M\geqslant N,\\ 0, & \text{если} \quad M\leqslant N. \end{cases}$$

Когда M возрастает все больше и больше, значение $\Phi u(T)$ приближается к значению Пси (T), а именно,

$$\langle \Phi u(T)\rangle_Q^{(\infty)} = \sum_{i=1}^{N} p_i^2.$$

16.5. Симметрические функции частот символов

Инвариантность для Пси, установленная в теоремах 7 и 8, выполняется также для всех симметрических функций частот символов m_i. Такой простейшей непостоянной полиномиальной функцией является $\sum_{i=1}^{N} m_i^2$. Эта функция является членом следующего интересного семейства[1]:

$$\text{Пси}_a(T) = \begin{cases} \left(\sum\limits_{i=1}^{N}(m_i/M)^a\right)^{1/(a-1)}, & \text{если} \quad 1<a<\infty,\\ \exp\left(\sum\limits_{i=1}^{N}(m_i/M)\cdot\ln(m_i/M)\right), & \text{если} \quad a=1,\\ \max\limits_{i=1}^{N}(m_i/M), & \text{если} \quad a=\infty \end{cases}$$

с нормализацией $\sum_{i=1}^{N}(m_i/M)=1$. $\text{Пси}_2(T)=\text{Пси}\,(T)$. Обобщая результат из разд. 16.2.2., видим, что для всех a, $1\leqslant a\leqslant\infty$, $\text{Пси}_a(T)=1/N=\varkappa_{\text{R}}$ тогда и только тогда, когда все m_i равны.

Интересны функции Пси_1 и Пси_∞, которые являются непрерывными предельными функциями семейства. Пси_1 имеет также представление

$$\text{Пси}_1(T) = \prod_{i=1}^{N}(m_i/M)^{m_i/M}.$$

Логарифмическая величина $-\operatorname{ld}\text{Пси}_a(T)$ называется a-энтропией Реньи текста T (1960)[2]. Это семейство имеет следующее представление:

[1] Где $x\cdot\ln x\nearrow 0$ при $x\searrow 0$; $x^x\nearrow 1$ при $x\searrow 0$.

[2] Альфред Реньи — венгерский математик, 1921–1970 гг.

$$-\operatorname{ld}\Пси_a(T)=\begin{cases}-\dfrac{1}{a-1}\cdot\operatorname{ld}\left(\sum\limits_{i=1}^{N}(m_i/M)^a\right), & \text{если} \quad 1<a<\infty,\\ -\left(\sum\limits_{i=1}^{N}(m_i/M)\cdot\operatorname{ld}(m_i/M)\right), & \text{если} \quad a=1,\\ -\max\limits_{i=1}^{N}\operatorname{ld}(m_i/M), & \text{если} \quad a=\infty.\end{cases}$$

1-энтропия Реньи $-\operatorname{ld}\Пси_1$ есть энтропия Шеннона (1945 г.)[3)].
2-энтропия Реньи $-\operatorname{ld}\Пси_2$ могла бы называться энтропией Кульбака.

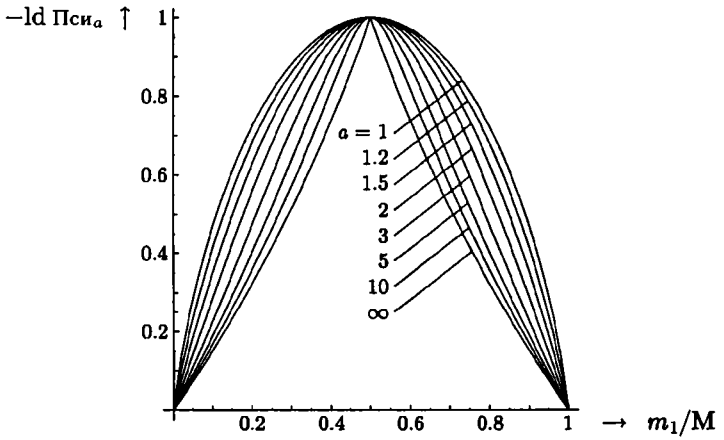


Рис. 115. График a-энтропии Реньи для $N=2$

На рис. 114 показан график $-\operatorname{ld}\Пси_a$ для $N=2$ и некоторых значений величины a.

Для английского текста T ($M=280$) из разд. 15.8.1 для отдельных символов

$$\begin{aligned}\Пси_1(T)&=5.852\%, & -\operatorname{ld}\Пси_1(T)&=4.095,\\ \Пси_2(T)&=6.936\%, & -\operatorname{ld}\Пси_2(T)&=3.850 \quad \text{(разд. 16.4.1)},\\ \Пси_\infty(T)&=12.857\%, & -\operatorname{ld}\Пси_\infty(T)&=2.959.\end{aligned}$$

[3)]Клод Шеннон (1916–2001 г.) — американский математик, инженер и создатель теории информации, известность получил в 1937 г. после публикации работы по релейным схемам и булевой алгебре (A Symbolic Analysis of Relay and Switching Circuits. Trans. AIEE 57, 713–723, 1938). В 1941 г. он работал в Лаборатории Белла над математическими проблемами секретных сообщений. Это привело его к созданию теории информации и передачи информации при наличии помех.

Для биграмм значения энтропии немного меньше:

$$\sqrt{\text{Пси}_1(T \times T^{(1)})} = 9.37\%, \quad -\frac{1}{2}\ \text{ld}\ \textit{Пси}_1(T \times T^{(1)}) = 3.42,$$

$$\sqrt{\text{Пси}_2(T \times T^{(1)})} = 10.58\%, \quad -\frac{1}{2}\ \text{ld}\ \text{Пси}_2(T \times T^{(1)}) = 3.24 \quad \text{(разд. 16.4.1)},$$

$$\sqrt{\text{Пси}_\infty(T \times T^{(1)})} = 17.96\%, \quad -\frac{1}{2}\ \text{ld}\ \text{Пси}_\infty(T \times T^{(1)}) = 2.48.$$

Глава 17

Исследование периодичности

> Можно принять в качестве принципа, что никогда не стоит биться над дешифрованием какого-либо непостижимого шифра, пока его автор сам не расшифрует какой-нибудь очень трудный шифр.
>
> *Чарльз Бэббидж*, 1854 *г.*

> Правило Бэббиджа лишило бы криптологов некоторых из важнейших характерных черт современной криптографии, вроде механизма Вернама, ротора, машины Хагелина.
>
> *Дэвид Кан*, 1967 *г.*

Даже если используется множество независимых алфавитов, периодический многоалфавитный шифр содержит один элемент, который трудно спрятать: число ключей в периоде шифра. Это основано на следующем свойстве стационарности стохастического источника: если P — открытый текст (длины M) порождается источником Q, то $P^{(s)}$ — тот же открытый текст P, но сдвинутый циклически на s позиций вправо, порождается тем же источником.

Теорема 1. Пусть p_i — вероятность порождения i-го символа χ_i источником Q. Пусть d — период периодического многоалфавитного функционального простого и однодольного шифра (для простоты предположим, что d делит M). Тогда шифротекст C открытого текста P и $C^{(k\cdot d)}$ — тот же шифротекст C, но сдвинутый на $k \cdot d$ позиций, будут порождаться тем же источником, и потому

$$\langle Kanna\,(C^{(k\cdot d)},\, C)\rangle_Q = \sum_{i=1}^{N} p_i^2 \quad \text{для всех } k.$$

Доказательство. Шифротекст открытого текста $P^{(k\cdot d)}$ (который получается из P циклическим сдвигом на $k \cdot d$) совпадает с $C^{(k\cdot d)}$, который получается из C циклическим сдвигом на $k \cdot d$. Согласно разд. 16.1.3 (*),

$$\langle Kanna\,(C^{(k\cdot d)},\, C)\rangle_Q = \langle Kanna\,(P^{(k\cdot d)},\, P)\rangle_Q = \sum_{i=1}^{N} p_i^2.$$

С другой стороны, аналогичное утверждение не может быть сделано для $\langle Kanna(C^{(u)}, C)\rangle_Q$, где d не делит n. Как правило, $C^{(u)} = C'$ и C порождаются стохастическими источниками Q' и Q, которые независимы друг от друга. Таким образом, если u не кратно d, то величина

$$\langle Kanna(C^{(u)}, C)\rangle_{Q'Q} = \sum_{i=1}^{N} p'_i p_i$$

может колебаться около $1/N$. По крайней мере, это имеет место в случае, достаточно большого числа алфавитов, которые выбраны так, что достигается полное смешивание вероятностей символов.

```
GEIEI  ASGDX  VZIJQ  LMWLA  AMXZY  ZMLWH
FZEKE  JLVDX  WKWKE  TXLBR  ATQHL  BMXAA
NUBAI  VSMUK  HSSPW  NVLWK  AGHGN  UMKWD
LNRWE  QJNXX  VVOAE  GEUWB  ZWMQY  MOMLW
XNBXM  WALPN  FDCFP  XHWZK  EXHSS  FXKIY
AHULM  KNUMY  EXDMW  BXZSB  CHVWZ  XPHWL
GNAMI  UK
```

Рис. 116. Криптотекст Калпа (G. W. Kulp)

17.1. Тест Каппа Фридмана

17.1.1. Уильям Фридман предложил построить диаграмму функции $Kanna(C^{(u)}, C)$ — индекса совпадения между $C^{(u)}$ и C. Для криптотекста на рис. 115 полученные точки изображены на рис. 116 (кроме случая $u = 0$, который исключается из построения). Для значений u, кратных 12, получаются высокие значения индекса совпадения, указывающие на то, что 12 может быть периодом.

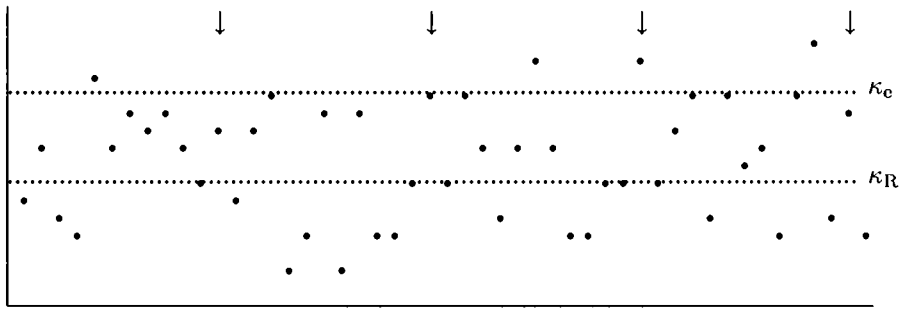


Рис. 117. Диаграмма Каппа для (английского) криптотекста Калпа

17.1.2. Криптотекст, изображенный на рис. 115, был помещен м-ром Калпом в Филадельфийской газете Alexander's Weekly Messinger в ответ на просьбу Эдгара Аллана По прислать шифротексты, зашифрованные одноалфавитно с сохранением пробелов между словами. Он был опубликован 26 февраля 1840 г. (рис. 118). Эдгар По в одном из следующих выпусков газеты наглядно показал, что эта криптограмма не удовлетворяет установленным правилам — он сделал это сведением каждой одноалфавитной подстановки собственно английских слов, приводящих к /mw/, /laam/, /mlw/ — к противоречию, и констатировал, что эта криптограмма — «тарабарщина из случайных символов, не имеющих никакого смысла» не что иное как «мошенничество». Один взгляд на распределение частот, приведенное на рис. 117, показывает его уравновешенную природу. Таким образом, частотное соответствие не могло работать.

12	7	2	5	10	4	6	9	6	3	10	12	14	9	2	4	4	2	7	2	7	7	16	15	4	8
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Рис. 118. Распределение частот криптотекста Калпа

17.1.3. Это наводит на мысль, что шифр мог быть многоалфавитным. Действительно, величина $1586/(187\cdot 186) = 4.56\%$ для *Фи*, или $1773/187^2 = 5.07\%$ для *Пси* близки к $\varkappa_R = 1/N$ и слишком низки для одноалфавитного шифра английского текста. Биграммная подстановка была также запрещена правилами, а ПЛЕЙФЕЙР был изобретен лишь в 1851 г. Так или иначе, как уже отмечалось, Эдгар По был одноалфавитно мыслящим человеком.

"Ge Jeasgdxv.
Zij gl mw. laam. xzy zmlwhfzek ejlvdxw kwke tx lbr atgh lbmx aanu bai Vsmukkss pwn vlwk agh gnumk wdlnzweg jnbxvv oaeg enwb zwmgy mo mlw wnbx mw al pnfdcfpkh wzkex hssf xkiyahul. Mk num yexdm wbxy sbc hv wyx Phwkgnamcuk?"

Рис. 119. Факсимиле криптотекста Калпа (1840 г.) (позже было обнаружено, что наборщик сделал несколько ошибок, например, читая q как g и пропустив одну букву)

17.1.4. Рис. 116 показывает, что мало значений Каппа подходят близко к $\varkappa_e$, зато большинство из них чуть выше или ниже $\varkappa_R$. Большие значения Каппа, даваемые периодом, должны также давать большие значения для всех кратных периода; это до некоторой степени исключает 5 и 15, тогда как 12 не может быть отвергнуто. Для криптотекста Калпа односимвольное шифрование, многоалфавитное с периодом 12 является правдоподобной гипотезой, но не более.

Криптотекст Калпа с его 187 символами довольно короткий; для более длинных текстов кратные периода выделяются намного лучше. Это можно видеть на рис. 119 для текста из 300 символов и на рис. 120 для текста из 3000 символов, где период определяется невооруженным взглядом.

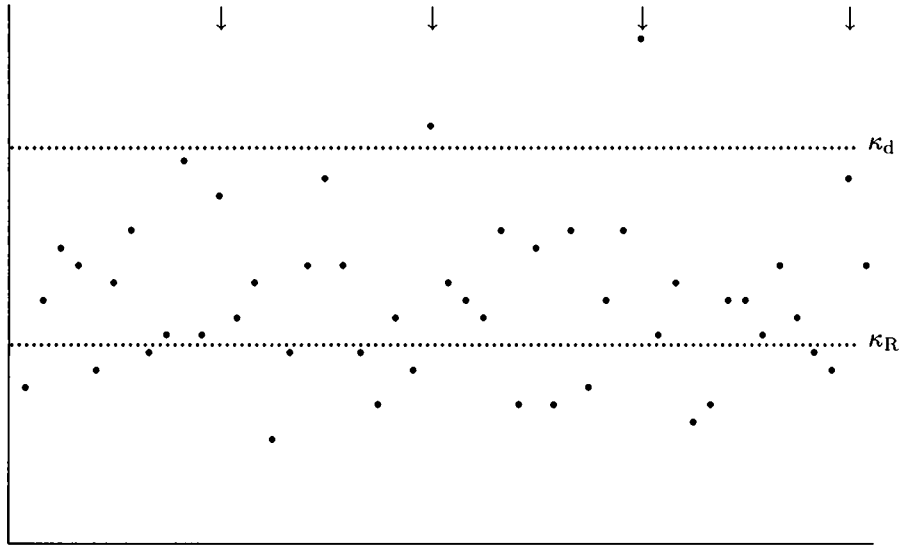


Рис. 120. Диаграмма *Kappa* для (немецкого) текста из 300 символов

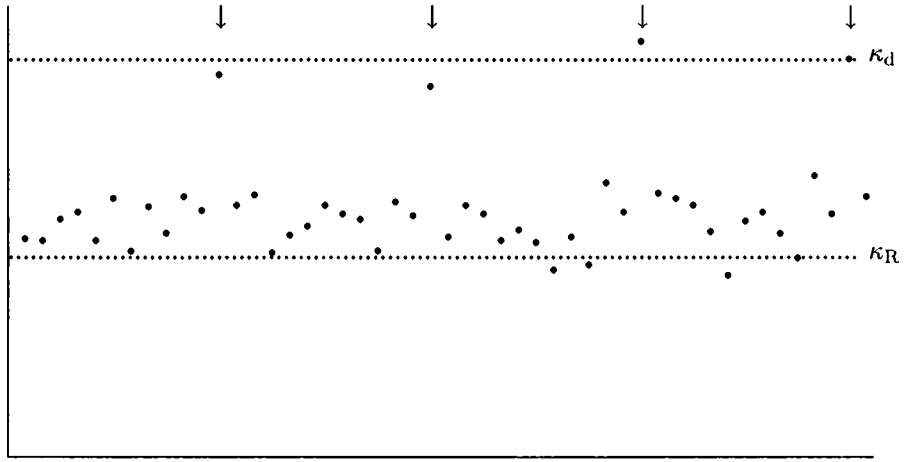


Рис. 121. Диаграмма *Kappa* для (немецкого) текста из 3000 символов

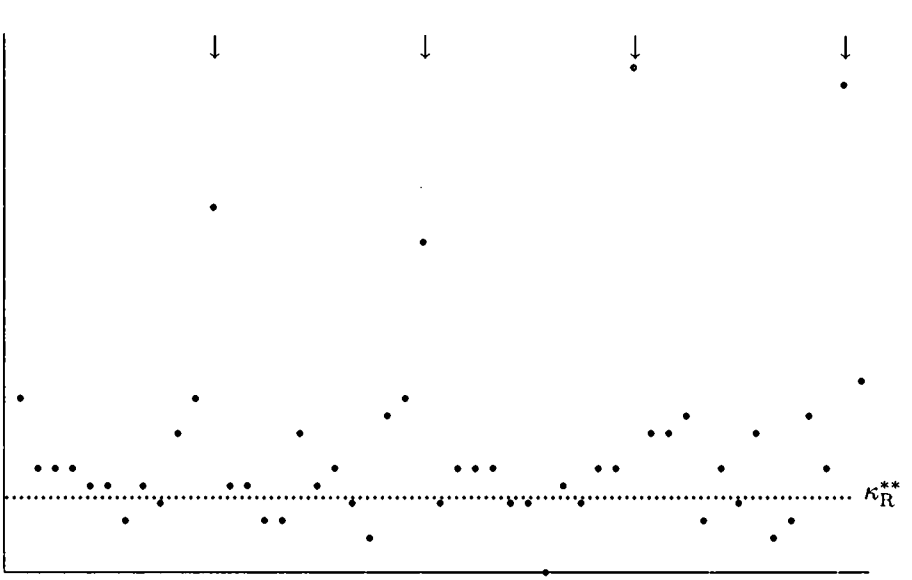


Рис. 122. Диаграмма *Каппа* для биграмм (немецкого) текста из 3000 символов

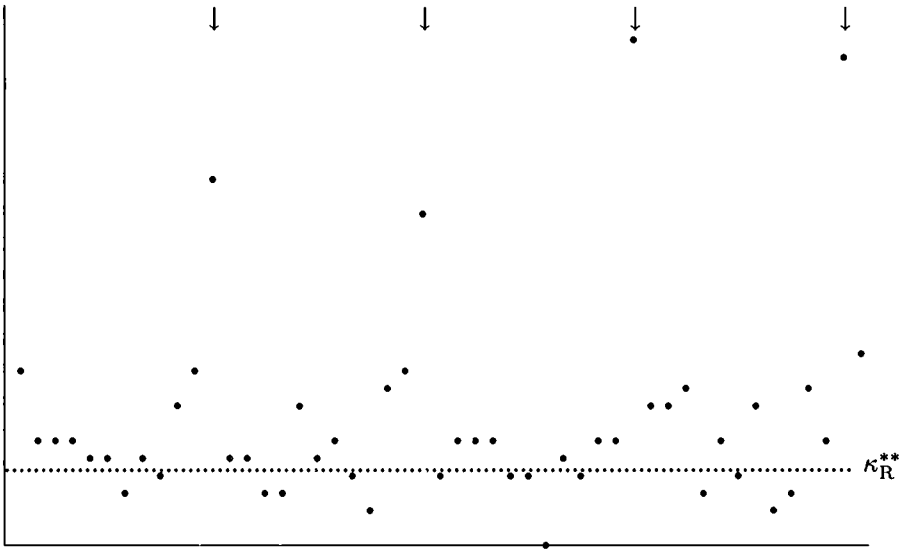


Рис. 123. Диаграмма *Каппа* для триграмм (немецкого) текста из 3000 символов

17.2. Тест Каппа для мультиграмм

Диаграмма *Каппа* не ограничивается отдельными символами. Биграммы и вообще мультиграммы тоже могут рассматриваться как символы, которые, однако, значительно увеличивают словарь.

При $N = 26$ для биграмм $\kappa_{\mathrm{R}}^{**} = 1/N^2 = 14.8\%\%$, для триграмм $\kappa_{\mathrm{R}}^{***} = 1/N^3 = 0.569\%\%$. Важно только, во сколько раз κ_S^{**} больше, чем κ_{R}^{**}, и оказывается, что коэффициент пропорциональности, равный примерно двум, для односимвольного случая заменяется множителем 4.5–7.5 для биграмм (рис. 121). Для английского языка, согласно Кульбаку, κ_{e}^{**} близко к 69%%, для немецкого языка, согласно Кульбаку и Бауэру, κ_{d}^{**} близко к 112%%. Это означает более явное отделение этих уровней. Для триграмм множитель будет равен примерно 40, но даже при 3000 символов колебание остается заметным (рис. 122). Для всего имеется определенный предел.

17.3. Криптоанализ с помощью машин

17.3.1. Использование перфокарт. Можно спокойно допустить, что в США методы Фридмана и Кульбака применялись во время Второй мировой войны, и делалось это с помощью машин. Примерно в 1932 г. Дайер из американских военно-морских сил для ускорения работы использовал машины IBM, данные в которые вводились при помощи перфокарт. Армия США стала применять такие машины в 1936 г. В 1941 г., в год Перл Харбора, разведывательная служба связи американской армии использовала 13 счетных машин, а в 1945 г. — уже 407. За их использование компания IBM получала в год \$ 750 000 арендной платы.

В Германии тоже использовались счетные машины. Они были, как и везде, нужны для выделения перешифрованных данных из кодов (разд. 9.2). Помимо этого они были полезны и для выполнения теста Каппа. С этой целью машины применялись также японцами (Кан, Такаги).

Предшественниками такой автоматической обработки с помощью перфокарточных машин были машины, использующие для ввода информации перфорированные бумажные листы. При подсчете совпадений криптотекст записывался бинарным 26-ти разрядным кодом с помощью прокалывания отверстий в листе. В то время как для подсчета однократного совпадения тексты достаточно записать один под другим, как это сделано в разд. 16.1, для теста Каппа подсчет должен делаться повторно для сдвинутых текстов. В этом случае оказываются оправданными сверхусилия по приготовлению листов с набитым текстом, так как с их помощью совпадения видны визуально: через совпавшие отверстия виден свет. Такое совпадение определяется не только намного быстрее, но также намного надежней. На рис. 123a изображены листы такими, какими они использовались в Блэтчли Парк, где они назывались «простыни Бэнбери», потому что перфорирование производилось в Бэнбери, небольшом соседнем городке. Таким способом можно было найти и подсчитать как совпадения отдельных символов, так и совпадения мультиграмм (рис. 123 b). Перфорированные бэнберские простыни можно было

делать вручную. Используя совсем несложную механизацию, можно было осуществить и более совершенное кодирование, которое сберегало бумагу, например, 5-битный код для кодирования десятичных цифр или 10-битный код, достаточный для кодирования как букв, так и цифр, который использовался в немецком OKW Йенсеном. Однако эти кодирования, исключая 5-битный код, требовали более сложных средств для обнаружения и регистрации совпадений.

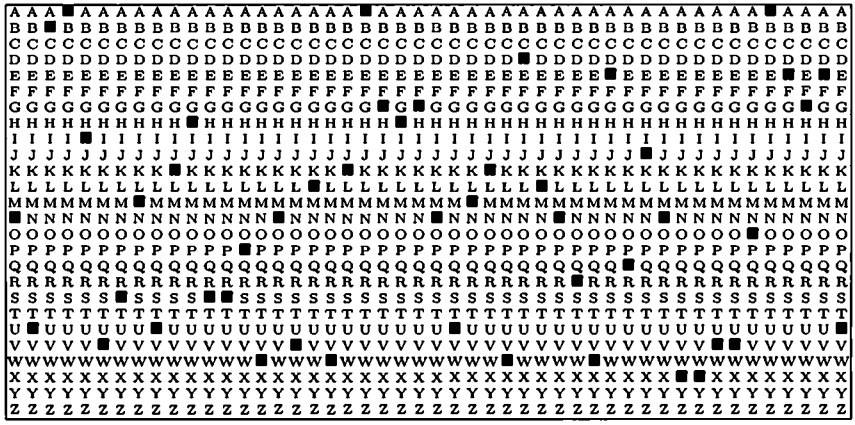


Рис. 123 a. Перфорированный лист для шифра из рис. 115:
NUBAIVSMUKHSSPWNVLWKAGHGNUMKWDLNRWEQJNXXVVOAEGEU

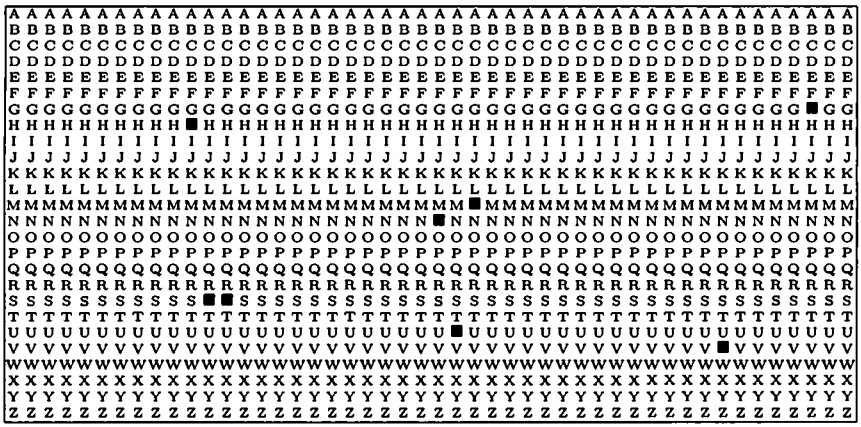


Рис. 123 b. Наложение двух перфорированных листов с отрезками криптотекста из рис. 115, сдвинутого на 72 символа

NUBA I V SMUKHSSPWNVLWKAGHGNUMKWDLN R WEQJNXXV V OAEG E U
CFPXHWZKEXHSSF X K I Y A HULMKNUMY E XDMW B XZSBCHVWZXPHWL
*** * * * *

17.3.2. Пилка дров. В отделе шифрования немецкого OKW, группа IV аналитического криптоанализа, возглавляемая Хюттенхайном, имела специальный аппарат, построенный Йенсеном, для определения совпадений и расстояний. Он назывался периодо- и фазоискателем (рис. 124) и работал с двумя идентичными 5-канальными телетайпными перфолентами, закольцованными в петли. Одна из петель содержала дополнительную позицию для перфорации. С каждым полным циклом прохождения перфолент через пару сканеров

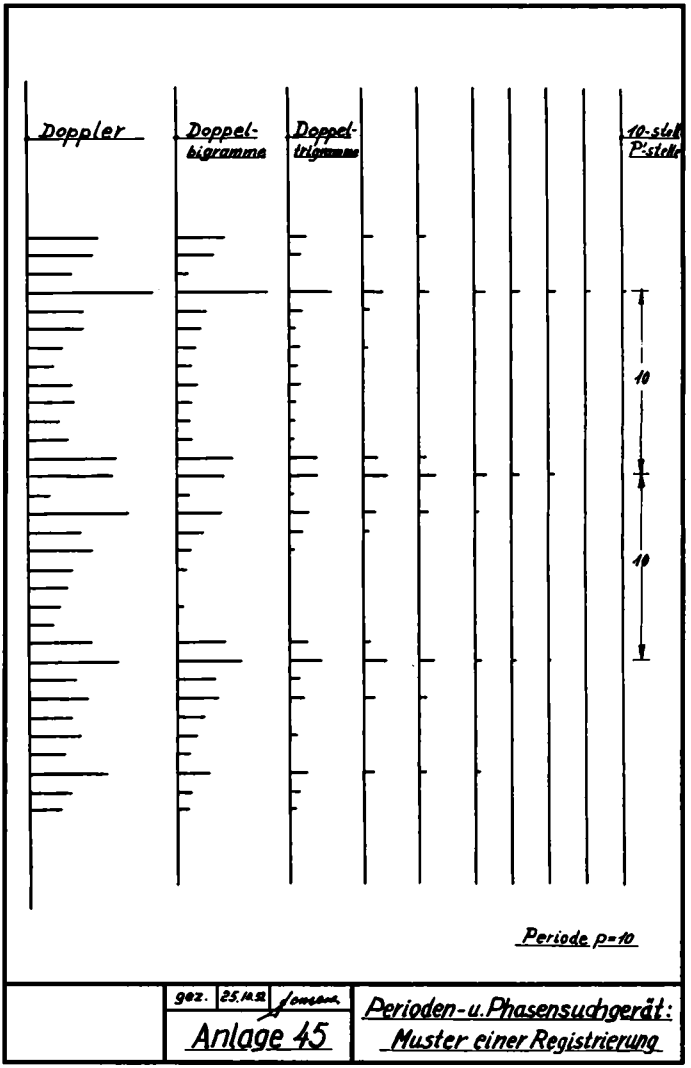


Рис. 124. Регистрация совпадений периодо- и фазоискателем (Willi Jensen, Hilfsgeräte der Kryptographie. Черновой вариант диссертации, 1953)

фаза между двумя сообщениями сдвигалась на одну позицию. Этот принцип «пилки дров», повидимому, был давно известен и применялся в разных других местах. Сканеры были фотоэлектрическими и для сравнения использовали релейную схему. Регистрация происходила автоматически, и для данного сдвига длина штриха (см. рис. 124) была пропорциональна числу совпадений. После полного цикла вторая лента сдвигалась вперед на одну позицию.

Одновременно, вторая релейная схема считала совпадающие биграммы, третья — совпадающие триграммы, и так вплоть до совпадающих 10-грамм («параллели»). Эти схемы автоматически строили диаграмму Каппа для отдельных символов, биграмм и т. д. При скорости сканирования 50 символов в секунду на текст из 600 символов требовалось 2 часа, что было в 100 раз быстрее, чем вручную. Этот аппарат был уничтожен в конце войны.

Материалы, доступные по работе Йенсена, не содержат никаких ссылок на Фридмана, но можно легко допустить, что по крайней мере его ранняя опубликованная работа была известна Хюттенхайну. Однако, он мог знать из разведывательных данных и о главной засекреченной работе[1] 1938–1941 гг.

17.3.3. ROBINSON'ы. В Великобритании ручная работа с перфокартами была механизирована при помощи HEATH ROBINSON[2] в мае 1943 г. Этот аппарат, сконструированный Винном-Вильямсом, имел компаратор, счетные схемы и мог фотоэлектрически читать две петли 5-канальных телетайпных перфолент со скоростью до 2000 символов в секунду (благодаря специальной электронной схеме быстрого счета). Согласно Мичи, HEATH ROBINSON тоже использовал принцип пилки дров, и успешно служил как для проверки совпадения, так и для нахождения повторений. Он был достаточно гибок, чтобы служить также для выделения перешифрований и для составления разностных таблиц (разд. 19.3). Сообщалось, что по методу Татта была исследована внутренняя структура шифровальной телетайпной машины SZ 40 (разд. 19.2.6); при этом, главным образом, использовались сложение по модулю 2 ключевого текста с криптотекстом, сдвигаемым до тех пор, пока не встретится верная фаза. Аппарат SUPER ROBINSON имел 4 ленты; DRAGON (так как текст протягивался (dragging)) имел аналогичные характеристики. В то время как версия Блэтчли Парк (т. е. английская) была электронной, американский DRAGON использовал реле.

Усовершенствованием стал COLOSSUS, который имел одну петлю, хранящуюся внутри, и был способен обрабатывать до 5000 символов в секунду без ошибок. COLOSSUS был построен для работы с шифрами телетайпного шиф-

---

[1] Фридман, военный криптоаналитик, Военный департамент, Служба начальника связи, Вашингтон, Офис правительственной печати. Vol. I: *Monoalphabetic Substitution Systems* 1938, 1942. Vol. II: *Simpler Varieties of Polyalphabetic Substitution Systems* 1938, 1943. Vol. III: *Simpler Varieties of Aperiodic Substitution Systems* 1938, 1939. Vol. IV: *Transposition and Fractionating Systems* 1941.

[2] Робинсон (Robinson W. H.) был английским карикатуристом, который рисовал великолепные и красивые, но непрактичные машины для всех возможных и невозможных задач. Имеются экземпляры HEATH ROBINSON, названные PETER ROBINSON и ROBINSON AND CLEAVER — по названиям Лондонских окружных универмагов. Всего к концу 1943 г. было заказано 12 Робинсонов.

рования и использовал сложные внутренние операции, которые выполнялись электронным способом.

17.3.4. Comparator. Недавно из работы Бурке стало больше известно об американских специальных аппаратах для исследования периодичности теста *Kanna* («I. C.»). Буш (1890–1974 гг.), хорошо известный уже по своей первой работе об аналоговых компьютерах (Дифференциальный Анализатор) для решения дифференциальных уравнений, в 1937 г. начал строить аппарат для подсчета совпадений, названный COMPARATOR, для отдела дешифрования американского военно-морского флота, следуя инструкциям его главы Венгера. Аппарат также работал в 26-значном коде. В отличие от аналогичного английского аппарата, применявшего испытанные инженерные технологии, Буш возлагал большие надежды на очень быстрое фотоэлектрическое сканирование и электронные счетчики (в 10-значном коде). Но в 1937 г. было рискованно строить электронный аппарат с более чем 100 электронными лампами, работающими вместе. Этот проект провалился в том числе и по организационным причинам. Несмотря на это проект был продолжен, что, возможно, объясняется ролью, которую во время войны играл Буш, будучи директором Национального комитета оборонных исследований (позднее Министерство научных исследований и развития). Медленный и недостаточный прогресс в реализации проекта не только обокрал (лишив успеха) адмирала Хупера, шефа связи военно-морских сил, и его помощника Венгера, сторонника чистого криптоанализа, — но также значительно отсрочил применение военно-морским флотом криптоаналитических машин, что привело уже в 1941 г., если не раньше, к значительному отставанию США от Великобритании в работах по дешифрованию машинных шифров. Положение изменилось только к 1946 г.

Однако чистый криптоанализ имел сильных сторонников среди математически мыслящих криптологов. Одна маленькая группа под руководством опытной Дрисколл при поддержке математика Энгстрома атаковала машину ENIGMA методами чистой математики. В результате этой работы была построена в 1942 г. и в конце 1943 г. принята на вооружение микрофильмовая машина HYPO (т. е. «гипотетическая»). Примерно в 1943 г. были сконструированы релейные машины VIPER и PYTHON, предназначенные для работы против японских роторных машин, и их электронный потомок RATTLER.

17.3.5. RAM. Как мы увидим в разд. 17.5, не только тест *Kanna*, но и тест *Xu* можно использовать для исследования периодичности. Тест *Xu*, предложенный Кульбаком в 1935 г., не был популярен в 1937 г., потому что помимо подсчета включал также сложение и даже умножение; однако к нему вернулись в 1940 г., и в машинах RAM («Rapid Analytical Machines [быстродействующие аналитические машины (*англ.*)]») в 1944 г. он добился заслуженного успеха. В принципе, машины COLOSSUS тоже могли выполнять исследование Кульбака, но неясно, было ли это на самом деле.

Как только появились универсальные электронные компьютеры, они стали использоваться в криптоаналитической работе. Первые специализированные для этой цели модели DEMON, OMALLEY, HECATE, WARLOCK использо-

вались в конце 1940-х гг., более совершенные компьютеры COMPARATOR, GOLDBERG[3], ATLAS I и ATLAS II стали браться на вооружение в 1950-х гг. Все больше и больше усилия криптоаналитиков перемещаются в программирование универсальных компьютеров, снабженных быстрым специальным и часто секретным дополнительным оборудованием. Этот процесс достиг высшей точки в настоящее время в архитектуре суперкомпьютеров CRAY, развиваемых с 1976 г. (Вклейка Q).

17.4. Анализ Касиски

В предельном случае диаграммы Каппа для мультиграмм, определяются лишь те длинные мультиграммы, которые встречаются повторно, и только для них регистрируются расстояния между повторениями. Такой поиск «параллелей» (по-немецки, Parallelstellensuchen) был опубликован в 1863 г. Касиски; еще задолго до эпохи Фридмана и Кульбака он отдавал предпочтение систематическому способу атаки профессиональных дешифровальщиков против многоалфавитного шифра и подрывал широко распространенное убеждение в его нераскрываемости (разд. 8.4.1) по крайней мере для периодического случая.

17.4.1. Ранние шаги. Несистематические атаки на многоалфавитные шифры начались вскоре после их изобретения. Порта был иногда удачлив: выражение OMNIA VINCIT AMOR [любовь преодолевает все (*лат.*)] было (чересчур) коротким и не раз использовалось в качестве тайного ключа некомпетентными шифровальщиками; оно потребовало у Порта всего несколько минут для разгадки, и шифр был взломан. Сам он использовал только длинные ключи и пропагандировал использование ключей, далеких от повседневного использования. И Джованни Ардженти, получив от своего хозяина Бонкампаньи, герцога Сора — племянника папы Григория XIII следующую криптограмму для испытания его способностей:

Q A E T E P E E E A C S Z M D D F I C T Z A D Q G B P P L E A Q T A I U I,

решил ее быстро, как он писал, 8 октября 1581 г.; он угадал ключ

I N P R I N C I P I O E R A T V E R B U M

и, рассчитывая на тот факт, что герцог всегда использовал десять инволютивных алфавитов типа ПОРТО, описанных в 1563 г. (разд. 7.4.4, рис. 53), подумал, а почему герцог должен был изобрести что-нибудь самостоятельно? Открытый текст был началом «Энеиды» Вергилия:

arma virumque cano troiae qui primus ab oris
[битвы и мужа пою, кто в Италию первым из Трои (*лат.*)].

[3]Говорят, что эта машина названа в честь Рубе Голдберга, американского двойника HEATH ROBINSON. Но возможен намек на Эммануэля Голдберга, изобретателя фотоэлектрического датчика.

Порта рано напал на методическую идею: если с некоторым диском Альберти криптоалфавит сдвигать в каждом шаге на одну позицию, то некоторые часто встречающиеся биграммы типа /ab/, /hi/, /op/ или триграммы типа /def/ (в *deficio*) или /stu/ (в *studium*) порождают буквы, повторяющиеся в криптотексте. Найдя в криптотексте MMM и через 51 позицию снова MMM, Порта заключил, что период ключа равен 17 и что ключ был повторен трижды, так как период 51 был бы чересчур длинным, а период 3 чересчур коротким для умного шифровальщика.

Порта остановился буквально в шаге от метода, найденного Касиски. Ему оставалось только понять, что имеет значение не шаблон 111, а его повторение в нескольких фрагментах криптотекста, вызываемое совпадением частого фрагмента открытого текста с одним и тем же куском повторяющегося ключа, которое должно естественно случаться лишь на расстоянии, кратном периоду. Если бы Порта заметил это и опубликовал свое открытие, многоалфавитный шифр был бы уязвим уже во времена Эдгара Аллана По.

Следующий упрощенный пример Кана иллюстрирует анализ Касиски: предположим, что метод ВИЖЕНЕРА в алфавите $\mathbb{Z}_{26}$ работает с ключом RUN (чересчур) малой длины 3:

```
t o b e o r n o t t o b e t h a t i s t h e q u e s t i o n
R U N R U N R U N R U N R U N R U N R U N R U N R U N R U N
K I O V I E E I G K I O V N U R N V J N U V K H V M G Z I A
```

Тогда ключевой фрагмент RUNR встречается с фрагментом открытого текста /tobe/ дважды на расстоянии 9, что приводит к повторению фрагмента KIOV, кроме того, ключевой фрагмент UN дважды встречается с фрагментом /th/ открытого текста на расстоянии 6. Расстояния 9 и 6 должны быть кратны периоду, который может быть лишь равным 3 (или 1).

Аналогичный пример с ключом COMET длины 5 такой:

```
t h e r e i s a n o t h e r f a m o u s p i a n o p l a y
C O M E T C O M E T C O M E T C O M E T C O M E T C O M E
V V Q V X K G M R H V V Q V Y C A A Y L R W M R H R Z M C
```

Здесь расстояния между повторяющимися фрагментами текста равны 10 и 15, так что период ключевой последовательности может быть только 5 (или 1).

17.4.2. Бэббидж и дешифрование. За десять лет до Касиски Чарльз Бэббидж мог иметь лишь слабое подозрение о важности повторений. Он не только любил прочитывать одноалфавитные шифрсообщения в колонках объявлений викторианских лондонских газет, но также любил разгадывать секретные многоалфавитные шифры с разделением на слова. Его работа с линейными простыми схемами шифрования привела его в 1846 г. к дешифрованию шифров ВИЖЕНЕРА и БОФОРТА с помощью математических уравнений (разд. 7.4.1), и таким образом он мог находить решение, применяя вероятные слова как в открытом тексте, так и в ключе. В результате,

как показывают бумаги Бэббиджа в Британском музее, он добился глубокого понимания тонкостей периодических шифров, хотя даже если он и обнаружил важность повторений Касиски, он все равно не написал об этом, как это предположил Франксен в 1984 г.

Таким образом, честь первооткрывателя систематических средств атаки против многоалфавитного шифра, не ограниченного только линейными подстановками, и следовательно, основателя современной криптологии принадлежит прусскому пехотному офицеру.

Фридрих Касиски родился 29 ноября 1805 г. в Шлохау, Восточная Пруссия (теперь г. Глухов, Польша). В 1822 г. он вступил в 33-й полк графа Роона, где прослужил 30 лет. Во времена досуга он обращался к криптографии. В 1863 г. уважаемым Берлинским издательством Миттлер и Сын был опубликован его 95-страничный манускрипт «Die Geheimschriften und Dechiffrirkunst [Секретное письмо и искусство дешифрования (*нем.*)]».

Прежде всего скажем, что его публикация не сделала сенсации, — и разочарованный Касиски обратился к естественной истории, став местной знаменитостью. Революция в криптологии, которую он вызвал, наступила после его смерти, случившейся 22 мая 1881 г. Керкхоффс комментировал работу Касиски в важной статье 1883 г., на ней основывались книги Де Виари 1893 г. и Деластеля 1902 г. В начале двадцатого века эта революция была уже в движении, и уязвимость многоалфавитных шифров была в общем принята среди профессионалов.

В свете открытия Фридманом в 1925 г. индекса совпадения анализ Касиски кажется довольно грубым методом — повторения биграмм проигнорированы, «потому что они слишком часты», — повторения отдельных символов тем более, в то же время *Kanna* считает все повторения и спрашивает лишь, имеется ли их больше среднего числа или нет. Игнорирование повторения биграмм у Касиски оправдывается тем фактом, что в редких случаях они могли появляться случайно. Даже с триграммами это бывает, и это нарушает анализ, тогда как индекс совпадения из-за своей стохастической природы лишен подобных недостатков.

Анализ Касиски устанавливает в качестве периода наибольший общий делитель расстояний записанных повторений, прагматично исключая те из них, которые квалифицируются как надоедливые и предполагаются случайными. В этом отношении анализ Касиски является интуитивным и мало пригодным для автоматизации. Кроме того, для определения периода анализ Касиски нуждается в более длинных текстах, чем анализ Фридмана.

Случайные повторения часто наблюдаются и легко объяснимы в случае линейных подстановок. Причиной является закон коммутативности, который имеет место для сложении по mod N: слово /anton/ с ключом BERTA и /berta/ с ключом ANTON дают одно и то же. Такой эффект встречается с частотой больше средней, если как открытый текст, так и ключ берутся из одного и того же естественного языка, в частности, из одного и того же жанра. Повторения в *ключевом тексте* тоже могут привести к «неправильным» повторениям в криптотексте — ключи со словами вроде DANSEUSECANCAN,

VIERUNDVIERZIG могут раздражать «незаконного» дешифровальщика. Мы вернемся к этому в разд. 18.5.

17.4.3. Пример. Встречающиеся в литературе примеры на анализ Касиски почти всегда являются очковтирательством: они дают больше повторений, чем можно ожидать в среднем. О следующем примере (Кан) этого сказать нельзя. Открытый текст даже заслужил положительный отзыв от американского офицера-связиста А. Майера (1866 г.). Криптотекст выглядит так:

```
ANYVG YSTYN RPLWH RDTKX RNYPV QTGHP
HZKFE YUMUS AYWVK ZYEZM EZUDL JKTUL
JLKQB JUQVU ECKBN RCTHP KESXM AZOEN
SXGOL PGNLE EBMMT GCSSV MRSEZ MXHLP
KJEJH TUPZU EDWKN NNRWA GEEXS LKZUD
LJKFI XHTKP IAZMX FACWC TQIDU WBRRL
TTKVN AJWVB REAWT NSEZM OECSS VMRSL
JMLEE BMMTG AYVIY GHPEM YFARW AOAEL
UPIUA YYMGE EMJQK SFCGU GYBPJ BPZYP
JASNN FSTUS STYVG YS
```

Подсчет частот символов показан на рис. 125; он слишком однороден, чтобы криптотекст мог быть получен одноалфавитной подстановкой или перестановкой. Таким образом, подозрение падает на многоалфавитную подстановку. Это подтверждается богатством повторений фрагментов длины 3 и большей, среди повторяющихся фрагментов есть очень длинные, вроде LEEBMMTG или CSSVMRS. Расстояния между повторениями приведены на рис. 126 вместе с их разложениями на простые множители. Наибольший общий делитель расстояний равен 2, но этот очень малый период лишь кажущийся, он получен из-за случайных повторений.

Дословно следуя Касиски, расстояния разлагаются на множители, и найденный множитель наибольшей частоты будет периодом. Литература интерпретирует это, следуя Охаверу обычно так, что должны быть выписаны все множители расстояний, а не только простые (рис. 127). Это, возможно, приведет к двум множителям, встречающимся одинаково часто; в таком случае берется больший из них, если он кратен меньшему, в противном случае лучше следовать двум возможностям.

На рис. 127, кроме множителя 2, который отвергается как чересчур маленький, имеется два множителя 3 и 6, которые встречаются 7 раз. Это делает кандидатом множитель 6. Хорошо, когда оказывается, что это верно, однако мы должны положиться на удачу. Откорректированное правило — надо

```
14 8 7 5 22 6 12 8 5 11 14 13 16 13 4 13 5 11 18 15 14 10 9 7 16 11
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
```

Рис. 125. Подсчет частот в криптотексте Кана

Фрагмент	Расстояние	Разложение на простые множители
YVGYS	280	$2^3 \cdot 5 \cdot 7$
STY	274	$2 \cdot 137$
GHP	198	$2 \cdot 3^2 \cdot 11$
ZUDLJK	96	$2^5 \cdot 3$
LEEBMMTG	114	$2 \cdot 3 \cdot 19$
CSSVMRS	96	$2^5 \cdot 3$
SEZM	84	$2^2 \cdot 3 \cdot 7$
ZMX	48	$2^4 \cdot 3$
GEE	108	$2^2 \cdot 3^3$

Рис. 126. Разложение на простые сомножители расстояний между повторениями Касиски

брать наибольший общий делитель расстояний между неслучайными повторениями, — страдает от того дефекта, что мы лишь впоследствии узнаем, какие повторения являются случайными, а какие нет. Интуитивно мы должны склоняться к пропуску тех «надоедливых» повторений, чьи расстояния не содержат некоторого часто встречающегося простого множителя, — на рис. 126 расстояния между фрагментами YVGYSБ, а также STY не содержат частого множителя 3. Поскольку YVGYSB — довольно длинный фрагмент, трудно поверить, что он является случайным, но в противном случае периодом должно было бы стать 2, во что поверить еще труднее. Если пренебречь повторениями GHP и LEEBMMTG, то кандидатом на период мог бы быть множитель 12, но LEEBMMTG очень длинно для случайного повторения, так что это тоже маловероятно. Таким образом, мы склоняемся к мысли, что периодом является 6.

Безусловно, это показывает другую слабую сторону анализа Касиски. Слишком малая величина заявленного периода, вызванная случайным повторением, разрушает последовательный процесс восстановления алфавитов.

Фрагмент	Расстояние	2	3	4	5	6	7	8	9	10	11	12	14	16	18	19	20	21	22	24
YVGYS	280	√	√√		√√		√				√					√				(?)
STY	274	√																		(?)
GHP	198	√√		√		√		√		√				√				√		
ZUDLJK	96	√√√		√		√				√		√							√	
LEEBMMTG	114	√√		√											√					
CSSVMRS	96	√√√		√		√				√		√							√	
SEZM	84	√√√		√√						√√						√				
ZMX	48	√√√		√		√				√		√							√	
GEE	108	√√√		√		√				√		√								

Рис. 127. Множители расстояний повторений Касиски

С другой стороны, может случиться, что наибольший общий делитель всех расстояний кратен истинному периоду. Это не только приводит к увеличению последовательной рабочей нагрузки, но также делает восстановление алфавитов менее надежным.

При всех обстоятельствах анализ Фридмана более достоверен, чем анализ Касиски.

Опережая дальнейшее дешифрование (разд. 18.1), мы заметим, что повторения STY и YVGYS оказываются случайными, порожденными линейной подстановкой над $\mathbb{Z}_{26}$. Ключом является SIGNAL, и YVGYS возникает первый раз из /signa/+GNALS, а второй из /gnals/+SIGNA; STY первый раз появляется из /als/+SIG, а второй из /sig/+ALS. Такое действие оказывает использование ключевого слова, взятого из жанра открытого текста. Оно усложняет дешифрование и очень желанно для криптографа. Кстати, оно также притупляет агрессивность тестов *Kanna* и *Xu*.

Случайные повторения могут появляться не только в результате коммутативности. Великий французский криптолог Этьен Базерье однажды потерпел неудачу с шифром БОФОРТ: в 1898 г. в телеграмме от мятежного гецога Орлеанского:

GNJLN RBEOR PFCLS OKYNX TNDBI LJNZE OIGSS HBFZN ETNDB...

он нашел повторение Касиски TNDB длины 4 с расстоянием 21, но это оказалось случайностью, вызванной /lesd/+ERVE и /prou/+IERV (настоящим ключом было: VENDREDIDIXSEPTFEVRIER). Еще два повторения EO и AQ длины 2 встречаются с расстояниями 22 и 13. Что делать? Базерье понравилось более длинное повторение TNDB, и он предположил, что период равен 21, но это было неверное предположение, проверка которого отняла у него много времени. В конце концов оказалось, что только короткое повторение EO длины 2 с периодом 22 было неслучайным. Базерье язвительно заметил: «en cryptographie aucun règle n'est absolue [в криптографии едва ли каждое правило абсолютно (*фр.*)]».

17.4.4. Машины. Несмотря на свою слабость, анализ Касиски еще применялся во время Второй мировой войны в качестве вспомогательного метода. Шифровальный отдел немецкого OKW разработал и применял специальный аппарат поиска параллельных мест, созданный Йенсеном помимо упоминавшегося периодо- и фазоискателя (разд. 17.3.2), пригодный для анализа Фридмана. Криптотекст перфорировался в 10-битный код (разд. 17.3.1) на ленту из пленки в 2 копиях. Одна копия (A) была замкнута в петлю и двигалась непрерывно через сканер, тогда как другая (B) продвигалась на одну позицию в своем сканере при каждом полном обороте петли (A). В случае совпадения два отверстия встречаются в сканере, который распознает совпадение с помощью фотоэлектрического элемента. Используя диафрагму изменяемой ширины, можно последовательно определить повторения биграмм, триграмм и т. д. В пределах доступной точности измерений можно найти повторения вплоть до десяти символов. Запись результатов производилась проблеском на

алюминиевой пластинке, движущейся в двух направлениях — в одном для (A) и другом для (B). Аппарат был довольно быстродействующим: чтобы обработать набор текстов, содержащих в сумме до 10000 символов, необходимо произвести 10^8 сравнений, аппарату на это требовалось меньше 3 часов. Аппарат служил, главным образом, для получения быстрой информации о текстах, зашифрованных одним и тем же ключом. Затем для детального исследования использовался периодо- и фазоискатель. Аппарат был уничтожен в конце войны, после продолжительного использования.

В США Буш построил в 1943 г. для OP-20-G аппарат ICKY, который находил длинные повторения или шаблоны одинаковых подгрупп, и при этом позволял выбирать различные комбинации при помощи штепсельного коммутатора. Примерно в середине 1944 г. под руководством Фридмана начала разрабатываться использующая микропленки универсальная криптоаналитическая машина Eastman 5202.

17.4.5. Фотоэлектрические датчики, подобные используемым немецкими, американскими и английскими криптоаналитиками, пытались применять ранее для поиска документов. В 1927 г. Маул из Берлина получил американские патенты 2 000 403 и 2 000 404, права на которые имеет IBM. Работа Голдберга «Статистическая машина» (американский патент 1 838 389, 29 декабря 1931 г., представлена 5 апреля 1928 г.), предназначенная для поиска документов, предшествовала обоим проектам Буша 1937 г. (COMPARATOR и первый RAPID SELECTOR, позднее названный «Memex»).

17.5. Создание глубины и Фи тест Кульбака

Как только угадан период d многоалфавитного шифра, начинается простой процесс определения вручную числа совпадений для сдвигов на $k \cdot d$ позиций, состоящий в подписывании криптотекста в строки длины d одна под другой, так что образуется d столбцов T_1, T_2, T_3, ..., T_d. На жаргоне криптоаналитиков это называется «выписать в глубину».

На рис. 128 показан результат такого выписывания в глубину для криптотекста Калпа (разд. 17.1.2) с угаданным периодом $d = 12$. Глаз сразу ловит совпадение отдельных символов (допплеры) с минимальным расстоянием d. Например, Z в 12-м столбце на первой и второй строках. Но легко видны также повторения на расстоянии $k \cdot d$, например, еще одно Z в 12-м столбце на предпоследней строке. Биграммные повторения (биграммные допплеры) тоже просматриваются, например, биграмма WK в 4-й и 7-й строках, биграмма MW во второй и 11-й строках, биграмма WZ в 12-й и предпоследней строках, биграмма NU в 6-й, 8-й и 2-й от конца строках.

17.5.1. Формирование столбцов. Небольшое усилие по расположению криптотекста в u столбцов дает гораздо больше, чем нахождение нескольких повторений Касиски. Если правильно угадан период ($u = d$), то тогда (и только тогда) в каждом столбце T_ρ шифрование одноалфавитно. Оно может быть проверено составлением $\Phi u(T_\rho)$ для $\rho = 1, 2, \ldots, u-1, u$. В положительном случае ($u = d$ — период) для каждого ρ должны ожидаться

G	E	I	E	I	A	S	G	D	X	V	Z	
I	J	Q	L	M	W	L	A	A	M	X	Z	
Y	Z	M	L	W	H	F	Z	E	K	E	J	
L	V	D	X	W	K	W	K	E	T	X	L	
B	R	A	T	Q	H	L	B	M	X	A	A	
N	U	B	A	I	V	S	M	U	K	H	S	
S	P	W	N	V	L	W	K	A	G	H	G	
N	U	M	K	W	D	L	N	R	W	E	Q	
J	N	X	X	V	V	O	A	E	G	E	U	
W	B	Z	W	M	Q	Y	M	O	M	L	W	
X	N	B	X	M	W	A	L	P	N	F	D	
C	F	P	X	H	W	Z	K	E	X	H	S	
S	F	X	K	I	Y	A	H	U	L	M	K	
N	U	M	Y	E	X	D	M	W	B	X	Z	
S	B	C	H	V	W	Z	X	P	H	W	L	
G	N	A	M	I	U	K						
ϕ_ρ	14	16	12	16	30	16	14	14	18	12	18	10 $\Sigma = 190$

Рис. 128. Криптотекст Калпа, выписанный в глубину из 12 столбцов

значения $Фи(T_\rho)$, близкие к $\varkappa_S$, т. е. гораздо большие, чем $\varkappa_R = 1/N$. В отрицательном случае (т. е. если $u \neq d$) значения $Фи(T_\rho)$ должны колебаться. Это $Фи$-тест Кульбака — очень точный критерий для исследования периода.

17.5.2. Фи-тест лучше, чем тест Каппа. Можно найти средние величины для значений $Фи\ (T_\rho)$, $\rho = 1, 2, \ldots, u-1, u$. Получаем:

$$\phi_\rho = \sum_{i=1}^{N} m_i^{(\rho)} \cdot (m_i^{(\rho)} - 1) \quad \text{для} \quad \rho = 1, 2, \ldots, u-1, u,$$

где $m_i^{(\rho)}$ — частота i-го символа χ_i в ρ-м столбце.

Мы должны рассмотреть правильно нормализованное значение ϕ_ρ:

$$Фи^{(u)}(T) = u \cdot \sum_{\rho=1}^{u} \phi_\rho / M \cdot (M-1).$$

Если $u \mid M$, то, как и в разд. 16.4.1, имеет место следующая теорема.

Теорема Каппа-Фи$^{(u)}$.

$$\frac{1}{M-1} \sum_{\rho=1}^{M-1} Kappa\,(T^{(u\cdot\rho)}, T) = Фи^{(u)}(T).$$

Таким образом, $Фи^{(u)}(T)$ — среднее арифметическое значений функции $Kappa\,(T^{(u\cdot\rho)}, T)$, т. е. от числа всех совпадений текстов $T^{(u\cdot\rho)}$ и T, расстояния между которыми кратны u. Оказывается, что это очень точный инструмент.

17.5.3. Пример. Для $u = 12$ имеется 12 алфавитов, каждый из которых имеет 16 или 15 символов в столбце (рис. 128). Подсчет ϕ_ρ дает:

для первого столбца с тремя S, тремя N и двумя G:

$$\phi_1 = 6 + 6 + 2 = 14;$$

для второго столбца с тремя N, тремя U, двумя B и двумя F:

$$\phi_2 = 6 + 6 + 2 + 2 = 16;$$

для третьего столбца с тремя M, двумя A, двумя B и двумя X:

$$\phi_3 = 6 + 2 + 2 + 2 = 12;$$

для четвертого столбца с четырьмя X, двумя K и двумя L:

$$\phi_4 = 12 + 2 + 2 = 16;$$

для пятого столбца с четырьмя I, тремя M, тремя V, тремя W:

$$\phi_5 = 12 + 6 + 6 + 6 = 30;$$

для шестого столбца с четырьмя W, двумя H и двумя V:

$$\phi_6 = 12 + 2 + 2 = 16;$$

и т. д.

Таким образом, $\sum\limits_{\rho=1}^{12} \phi_\rho = 190;$

$$\Phi u^{(12)} = 12 \cdot 190/(187 \cdot 186) = 6.56\%.$$

Заметим, что $\kappa_e = 6.58\%$. Таким образом, $u = 12$ вполне может быть периодом.

Для $u = 11$ существует 11 алфавитов, и для каждого имеется 17 символов в столбце. Рис. 129 показывает существование *допплера* W на расстоянии 11 во 2-й и 3-й строках и *допплера* V в первой, шестой и седьмой строках 11-го столбца. Никаких *биграммных допплеров* больше нет.

Подсчет ϕ_ρ дает (рис. 129):

$$\sum_{\rho=1}^{11} \phi_\rho = 110, \qquad \Phi u^{(11)} = 11 \cdot 110/(187 \cdot 186) = 3.48\%.$$

Заметим, что $\kappa_R = 3.85\%$. Мы видим, что величина $\Phi u^{(11)}$ значительно меньше, чем $\Phi u^{(12)}$, и $u = 11$ имеет малые шансы быть периодом.

	G	E	I	E	I	A	S	G	D	X	V	
	Z	I	J	Q	L	M	W	L	A	A	M	
	X	Z	Y	Z	M	L	W	H	F	Z	E	
	K	E	J	L	V	D	X	W	K	W	K	
	E	T	X	L	B	R	A	T	Q	H	L	
	B	M	X	A	A	N	U	B	A	I	V	
	S	M	U	K	H	S	S	P	W	N	V	
	L	W	K	A	G	H	G	N	U	M	K	
	W	D	L	N	R	W	E	Q	J	N	X	
	X	V	V	O	A	E	G	E	U	W	B	
	Z	W	M	Q	Y	M	O	M	L	W	X	
	N	B	X	M	W	A	L	P	N	F	D	
	C	F	P	X	H	W	Z	K	E	X	H	
	S	S	F	X	K	I	Y	A	H	U	L	
	M	K	N	U	M	Y	E	X	D	M	W	
	B	X	Z	S	B	C	H	V	W	Z	X	
	P	H	W	L	G	N	A	M	I	U	K	
ϕ_ρ	8	6	8	12	10	8	10	4	8	16	20	$\Sigma = 110$

Рис. 129. Криптотекст Калпа из 11 столбцов

	G	E	I	E	I	A	S	G	D	X	V	Z	I
	J	Q	L	M	W	L	A	A	M	X	Z	Y	Z
	M	L	W	H	F	Z	E	K	E	J	L	V	D
	X	W	K	W	K	E	T	X	L	B	R	A	T
	Q	H	L	B	M	X	A	A	N	U	B	A	I
	V	S	M	U	K	H	S	S	P	W	N	V	L
	W	K	A	G	H	G	N	U	M	K	W	D	L
	N	R	W	E	Q	J	N	X	X	V	V	O	A
	E	G	E	U	W	B	Z	W	M	Q	Y	M	O
	M	L	W	X	N	B	X	M	W	A	L	P	N
	F	D	C	F	P	X	H	W	Z	K	E	X	H
	S	S	F	X	K	I	Y	A	H	U	L	M	K
	N	U	M	Y	E	X	D	M	W	B	X	Z	S
	B	C	H	V	W	Z	X	P	H	W	L	G	N
	A	M	I	U	K								
ϕ_ρ	4	4	12	10	18	10	8	12	10	10	14	8	6	$\Sigma = 126$

Рис. 130. Криптотекст Калпа из 13 столбцов

Для $u = 13$ имеется 13 алфавитов, и для каждого 14 или 15 символов в столбце. В соответствии с рис. 130, производим подсчет ϕ_ρ:

$$\sum_{\rho=1}^{13} \phi_\rho = 126, \qquad \Phi u^{(13)} = 13 \cdot 126/(187 \cdot 186) = 4.71\%.$$

Таким образом, $u = 13$ тоже не является кандидатом на период.

Итак, $\Phi u^{(u)}$ можно подсчитать для $u = 2, 3, \ldots$, и построить диаграмму (рис. 131). Значение для $d = 12$ выделяется более заметно, чем на рис. 116. Можно сделать вывод, что анализ Кульбака эффективнее, чем анализ Фридмана.

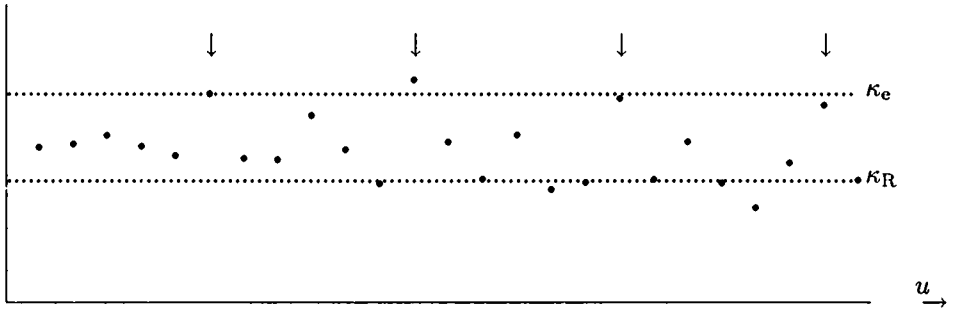


Рис. 131. Диаграмма $\Phi u^{(u)}$ для криптотекста Калпа

На рис. 131 мы видим, что кроме $d = 12$ имеется еще один пик для $d = 24$, который можно было ожидать. Но существуют еще чуть меньшие пики при $d = 6$ и при $d = 18$. Поэтому нельзя исключать и периода $d = 6$.

17.6. Оценка длины периода

Из теоремы *Kanna-Фи* (разд. 16.4.1) мы получаем величину

$$\langle \Phi u(T) \rangle_Q^{(M)} = \frac{1}{M-1} \sum_{\rho=1}^{M-1} \langle Kanna(T, T^{\rho}) \rangle_Q .$$

Из замечания, сделанного в начале этой главы мы можем вывести, что

$$\langle Kanna(T, T^{k \cdot d}) \rangle_Q \approx \kappa_S ,$$

более того, мы можем найти

$$\langle Kanna(T, T^{u}) \rangle_Q \approx \kappa_R = \frac{1}{N} \quad \text{при } u \text{ не кратном } d.$$

Предполагая в дальнейшем для простоты, что M кратно периоду d, получим, что в сумме слагаемое κ_S появляется $M/d - 1$ раз, тогда как κ_R появляется $M - M/d$ раз; таким образом, $\langle \Phi u(T) \rangle_Q^{(M)}$ лежит между κ_S и κ_R, точнее,

$$(M-1) \cdot \langle \Phi u(T) \rangle_Q^{(M)} \approx \left(\frac{M}{d} - 1\right) \cdot \kappa_S + \left(\left(M-1\right) - \left(\frac{M}{d} - 1\right)\right) \cdot \kappa_R .$$

Предполагая, что наблюдаемое значение $\Phi u(T)$ приближается ее ожидаемым значением, получаем (Синков, около 1935 г.):

$$(M-1)\cdot\Phi u(T)\approx\Big(\frac{M}{d}-1\Big)\cdot\varkappa_{\mathrm{S}}+\Big(\Big(M-1\Big)-\Big(\frac{M}{d}-1\Big)\Big)\cdot\varkappa_{\mathrm{R}}.$$

Хотя это всего лишь оценка, такое фундаментальное равенство показывает на качественном уровне, как при постоянном стохастическом источнике изменяется значение Φu при увеличении периода многоалфавитного шифра.

Для большого M и $d\ll M$ хорошим приближением является формула

$$\Phi u(T)\approx\frac{1}{d}\cdot\varkappa_{\mathrm{S}}+\Big(1-\frac{1}{d}\Big)\cdot\varkappa_{\mathrm{R}}.$$

Соотношение Синкова можно разрешить относительно d:

$$\Big(\frac{M}{d}-1\Big)\approx\frac{(M-1)(\Phi u(T)-\varkappa_{\mathrm{R}}}{\varkappa_{\mathrm{S}}-\varkappa_{\mathrm{R}}},$$

т. е.

$$d\approx\frac{\varkappa_{\mathrm{S}}-\varkappa_{\mathrm{R}}}{(\varkappa_{\mathrm{S}}-\Phi u(T))/M+(\Phi u(T)-\varkappa_{\mathrm{R}})}.$$

Для большого M и $d\ll M$ можно использовать формулу

$$d\approx\frac{\varkappa_{\mathrm{S}}-\varkappa_{\mathrm{R}}}{\Phi u(T)-\varkappa_{\mathrm{R}}}.$$

Например, для криптотекста Калпа с $M=187$, согласно разд. 17.1.1, $\Phi u=4.56\%$, так что при $\varkappa_{\mathrm{S}}=\varkappa_{\mathrm{e}}=6.58\%$, $\varkappa_{\mathrm{R}}=1/N=3.85\%$, мы получаем

$$d\approx\frac{2.73\%}{(2.02\%/187)+0.71\%}=3.79,$$

или, упрощая,

$$d\approx\frac{2.75\%}{0.71\%}=3.85.$$

Это значение является низким в сравнении с предположительным периодом $d=12$, и лучше соответствует $d=6$. Но оценка является довольно неустойчивой и не должна рассматриваться слишком серьезно. Оценка Синкова может служить вспомогательным средством серьезным анализам Касиски, Фридмана или Кульбака, если период маленький.

Глава 18

Выравнивание сопутствующих алфавитов

При известном периоде d многоалфавитный шифротекст дешифруется достаточно достоверно, а с помощью создания глубины дешифрование может быть сведено к решению d одноалфавитных шифров, и можно даже попытаться — если повезет — свести сопутствующие алфавиты к первичному алфавиту. Для случая шифра ВИЖЕНЕРА перебор всех сопутствующих алфавитов с помощью соответствующих профилей (разд. 18.1) довольно несложен. Общее выравнивание можно провести в случае шифра АЛЬБЕРТИ, если один из стандартных алфавитов известен или обнаружен (разд. 18.2). Однако в общем случае требуется взаимное выравнивание всех алфавитов (разд. 18.3), чтобы восстановить неизвестный первичный алфавит (разд. 18.4). Для этой цели идеально подходит анализ Кульбака. Но к случаю, когда неизвестны неродственные алфавиты, и каждый из них должен быть определен самостоятельно, такой подход неприменим.

18.1. Выравнивание по профилю

Учитывая широкое распространение шифра ВИЖЕНЕРА, часто бывает целесообразно попробовать именно этот вход, не требующий больших усилий.

В случае, если известен период d, строятся d профилей. Ленточный метод для перебора (гл. 12) и нахождение шаблонов (гл. 13) для отдельных одноалфавитных сложений ЦЕЗАРЯ здесь не работают, так как тексты разорваны на куски (по-немецки, zerrissen).

2 0 0 0 1 0 2 3 0 6 3 3 3 0 0 0 0 0 6 1 5 2 5 0 2 4
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Рис. 132. Частотное распределение в первом столбце криптотекста Кана

18.1.1. Использование глубины. Таким образом, для текста Кана (см. разд. 17.4.3) в предположении, что период d равен 6, мы рекомендуем создавать глубину и подсчитывать частоты для каждого из 6 столбцов. Тогда для первого столбца, т. е. для подтекста, состоящего из 1-го, 7-го, 13-го, ... символов исходного текста, результат можно увидеть на рис. 132. Даже без графика английский профиль сразу узнается: NOPQR — это низменность vwxyz, слева от нее JKLM — хребет rstu. Затем DEFGH на том расстоянии, где должен быть хребет lmno, который показан неясно. Но криптоаналитик должен быть готов к таким колебаниям, особенно если глубина невелика. Это замечание вполне приложимо также к наблюдению, что W не имеет той частоты, какая ожидалась бы от е-пика.

При $S \hat{=} \mathrm{a}$ первый столбец выравнен (найден сдвиг его алфавита относительно стандартного алфавита), и можно ожидать, что система в целом является системой ВИЖЕНЕРА. А пока найдена первая ключевая буква S.

Проводя аналогичную процедуру с другими столбцами, мы шаг за шагом получаем ключ

SIGNAL,

который подтверждается последующим дешифрованием всего шифротекста. С подчеркнутыми фрагментами основных неслучайных повторений открытый текст такой:

ifsig	nalsa	retob	edisp	layed	inthe
prese	nceof	anene	mythe	ymust	begua
rdedb	yciph	ersth	eciph	ersmu	stbec
apabl	eoffr	equen	tchan	gesth	erule
sbywh	ichth	esech	anges	arema	demus
tbesi	mplec	ipher	sareu	ndisc	overa
blein	propo	rtion	asthe	ircha	ngesa
refre	quent	andas	theme	ssage	sinea
chcha	ngear	ebrie	ffrom	alber	tjmye
rsman	ualof	signa	ls		

Теперь мы даже можем видеть, как были получены неслучайные повторения: длиннейшее LEEBMMTG произошло от повторяющейся комбинации из /frequent/ и GNALSIGN; другое, ZUDLJK — из повторной комбинации /must-be/ и ALSIGNA. Странно, но частая встречаемость слова /cipher/ не привела к повторениям. Повторения SEZM, GHP, ZMX, GEE являются неслучайными, возникшими от встречи /sthe/, /the/, /her/ и /are/ с ALSI, NAL, SIG и GNA. Повторения же YVGYS и STY являются случайными.

18.1.2. Составление профилей. В случае относительно длинных ключей глубина мала, и становится затруднительно выявить частотные различия. В таком случае существует еще одна простая возможность графического расстояния. Для криптотекста Калпа (разд. 17.1.2, рис. 115) подготовительная работа по формированию столбцов уже сделана в разд. 17.4, и выравнивание можно проводить после анализа Кульбака периода этого криптотекста.

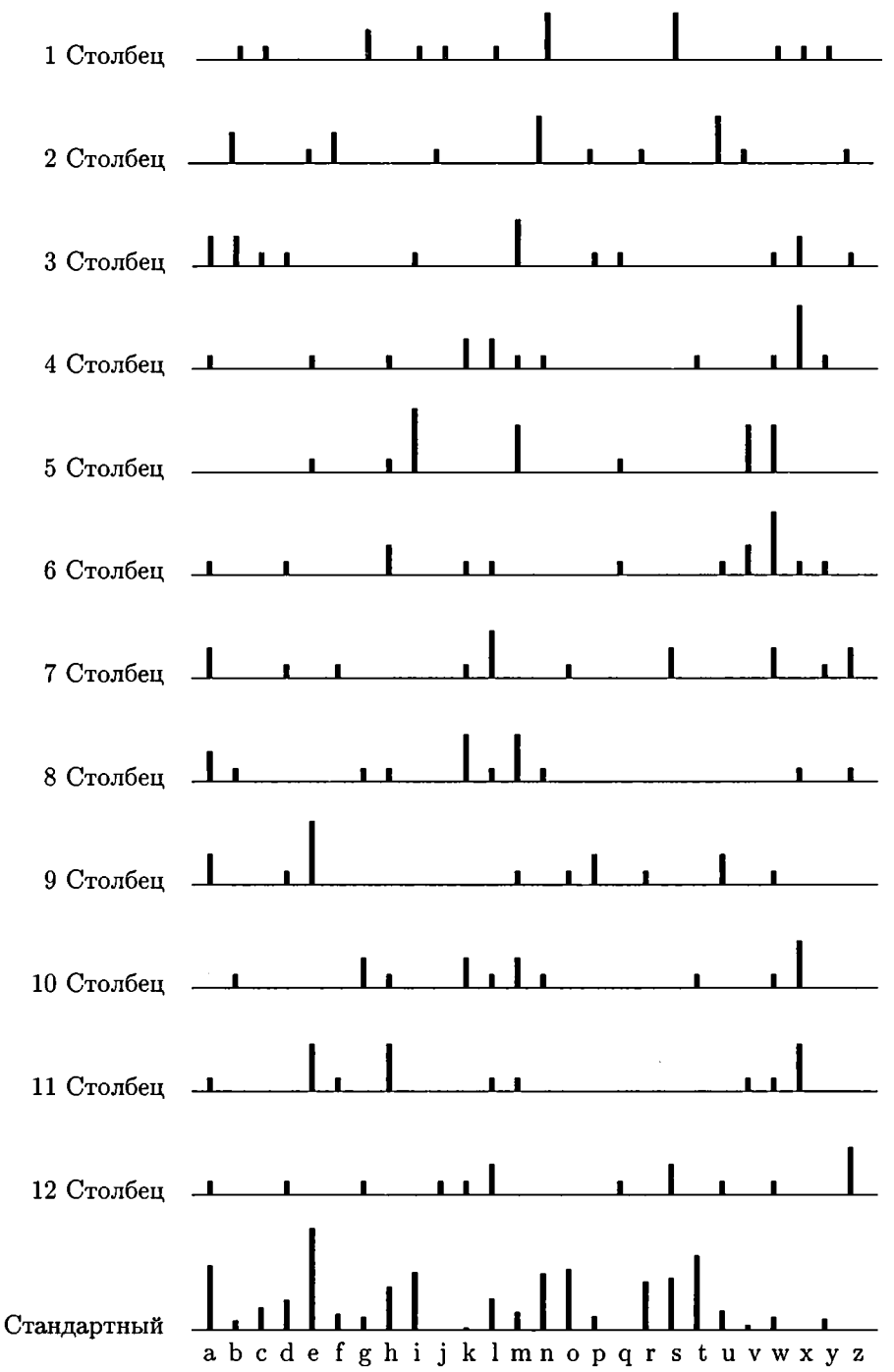


Рис. 133. Профили для криптотекста Калпа

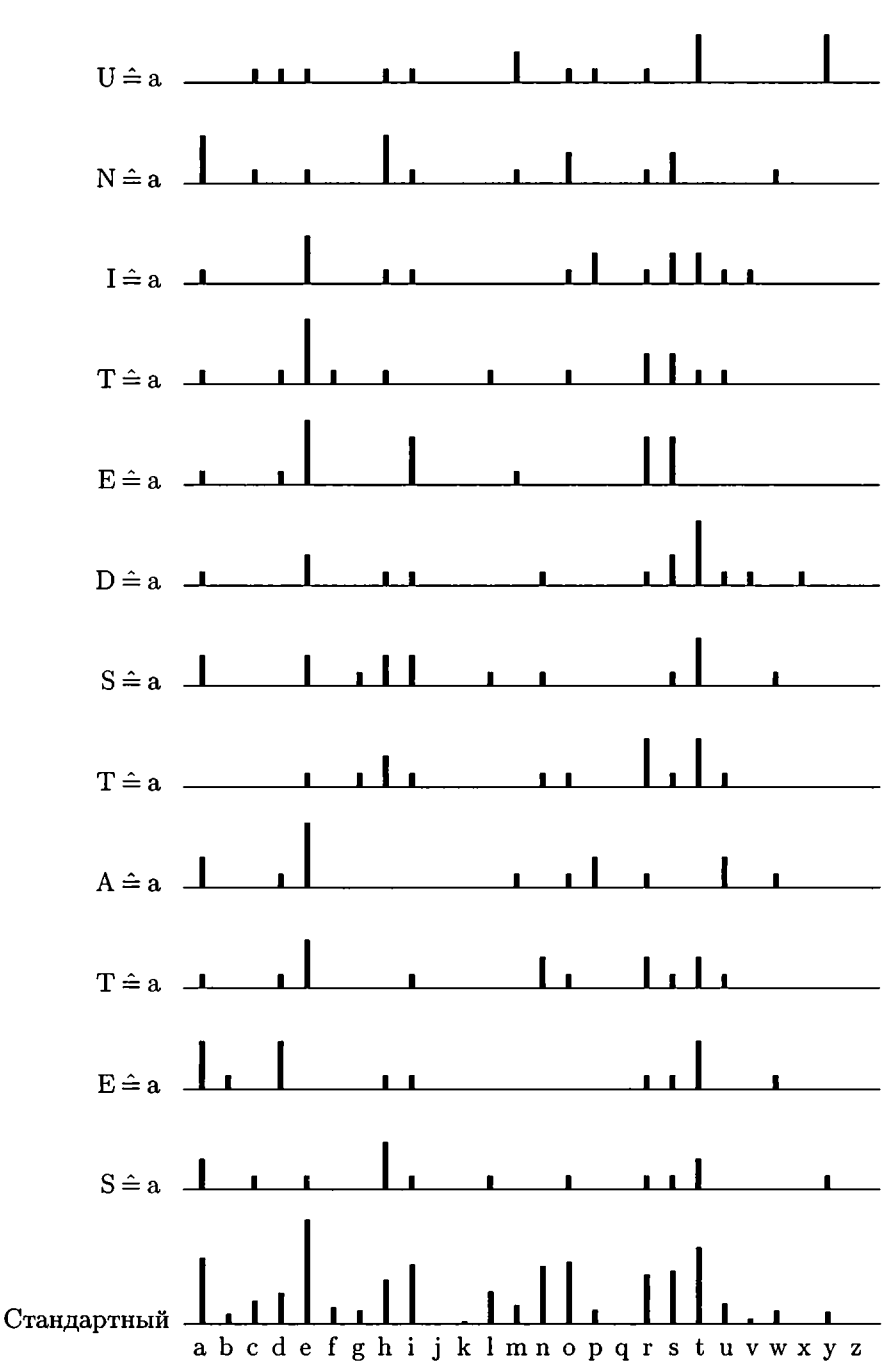


Рис. 134. Выравненные профили для криптотекста Калпа

Из рис. 128 получаем двенадцать профилей, представленных на рис. 133. На рис. 134 они упорядочены в соответствии с частотным профилем английского языка. Здесь испытание тоже оказалось успешным.

Одиннадцатый столбец вводит в заблуждение: самая частая буква /e/ английского языка здесь вообще не встречается. Действительно, можно сказать, что в случае малого множества символов предпочтительно сначала выстроить редкие символы, особенно самые редкие, которые отсутствуют. И тогда некоторые /e/ из других столбцов — 3-го, 4-го, 9-го, 10-го — дают хороший ключ к разгадке.

В конце концов оказывается, что буква /a/ открытого текста соответствует в первом алфавите букве U, во втором — T, в пятом и одиннадцатом алфавитах — E, в шестом алфавите — D, в седьмом и двенадцатом алфавитах — S. В девятом столбце букве /a/ открытого текста соответствует A, и таким образом, соответствующая подстановка тождественна. Подавить такую возможность в случае шифра ВИЖЕНЕРА было бы технической ошибкой, так как это открыло бы возможность атаке на несовпадение (разд. 14.1).

Выравнивание здесь, безусловно, облегчается тем фактом, что один алфавит встречается трижды. Такие шаблоны в ключе тоже являются технической ошибкой.

Ключевое слово теперь обнаруживается как

UNITEDSTATES,

что имеет определенный смысл, учитывая ситуацию в Филадельфии в 1840 г. В духе афоризма Рорбаха, проведенное дешифрование является вполне убедительным. Таким образом, получается, что Эдгар Аллан По был немного несправедлив, говоря, что этот текст был жульничеством. Дешифрованный текст приводится на рис. 135. Дешифрование было успешно проведено Винкелем в 1975 г. и опубликовано в колонке Мартина Гарднера в «Scientific American» в августе 1977 г.

```
mrale xande rhowi sitth atthe messe
ngera rrive shere atthe samet imewi
ththe satur dayco urier andot hersa
turda ypape rswhe nacco rding tothe
datei tispu blish edthr eeday sprev
iousi sthef aultw ithyo uorth epost
maste rs
```

Рис. 135. Открытый текст сообщения Калпа

[М-р Александр, как могло случиться, что посыльный прибыл сюда одновременно с субботним Курьером и другими субботними газетами, если согласно дате они опубликованы на три дня раньше? Является ли это ошибкой ваших почтовых чиновников? (*англ.*).]

18.2. Выравнивание относительно известного алфавита

Выравнивание алфавитов «на глаз» может оказаться затруднительным, например, на рис. 133, скажем, с первым или восьмым столбцами.

18.2.1. Применение *Xu*. Оказывается, что вычисление средних является достаточно точным инструментом. Естественна идея, определить выравнивающие сдвиги подсчетом *Xu* частот рассматриваемого и первичного алфавитов. Это показано на рис. 136 для несдвинутого стандартного алфавита, и на рис. 137 для подходящим образом сдвинутого алфавита, где /a/ соответствует U.

0	1	1	0	0	0	2	0	1	1	0	1	0	
A	B	C	D	E	F	G	H	I	J	K	L	M	
8.04	1.54	3.06	3.99	12.51	2.30	1.96	5.49	7.26	0.16	0.67	4.14	2.53	
a	b	c	d	e	f	g	h	i	j	k	l	m	
3	0	0	0	0	3	0	0	0	1	1	1	0	
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
7.09	7.60	2.00	0.11	6.12	6.54	9.25	2.71	0.99	1.92	0.19	1.73	0.09	64.81
n	o	p	q	r	s	t	u	v	w	x	y	z	

Рис. 136. *Xu* для стандартного алфавита относительно первого столбца

0	0	1	1	1	0	0	1	1	0	0	0	2	
U	V	W	X	Y	Z	A	B	C	D	E	F	G	
8.04	1.54	3.06	3.99	12.51	2.30	1.96	5.49	7.26	0.16	0.67	4.14	2.53	
a	b	c	d	e	f	g	h	i	j	k	l	m	
0	1	1	0	1	0	3	0	0	0	0	3	0	
H	I	J	K	L	M	N	O	P	Q	R	S	T	
7.09	7.60	2.00	0.11	6.12	6.54	9.25	2.71	0.99	1.92	0.19	1.73	0.09	86.03
n	o	p	q	r	s	t	u	v	w	x	y	z	

Рис. 137. *Xu* для стандартного алфавита относительно первого столбца, сдвинутого на a $\hat{=}$ U

В первом случае (a $\hat{=}$ A) дает для *Xu* значение $64.81/16\% = 4.05\%$; во втором случае (a $\hat{=}$ U) получается значительно большее значение $86.03/16\% = 5.37\%$. Таблица 20 дает список значений *Xu* для всех сдвигов. Оказывается, что кроме сдвига a $\hat{=}$ U выделяются также a $\hat{=}$ J и a $\hat{=}$ F. Эти три варианта нуждаются в последующей переборной обработке.

Это показывает, что периодическая система ВИЖЕНЕРА может быть дешифрована механически при разумных благоприятных условиях. Для текста столь же длинного, как в случае «Калп против Эдгара По», дешифрование будет успешным, а при помощи персонального компьютера и просто выполнимым.

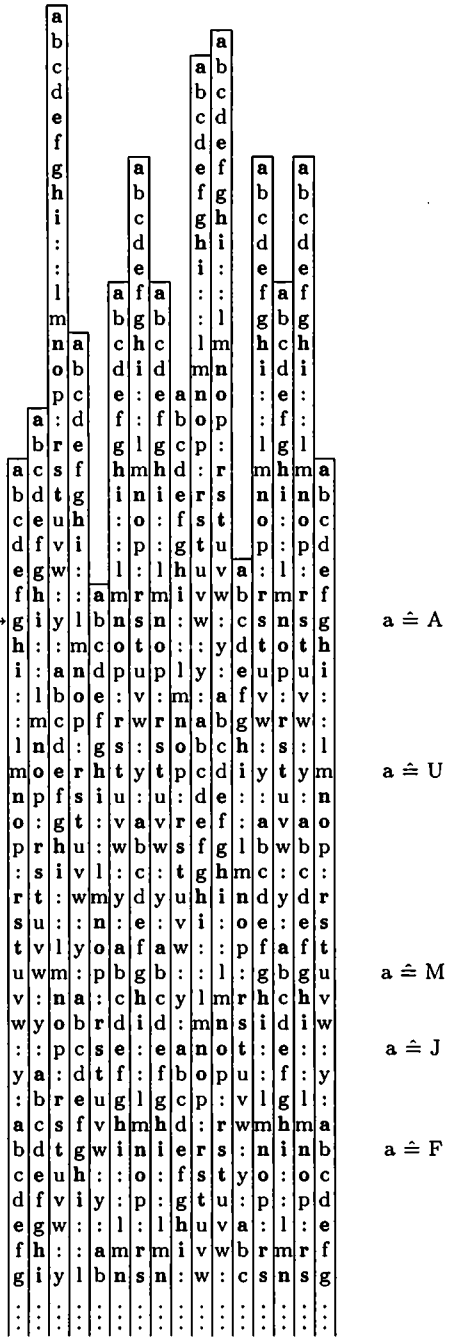


Рис. 138. Подсчет значений *Xu* для стандартного алфавита относительно первого столбца

18.2.2. Ленточный метод. Основная идея выравнивания относительно первичного алфавита (в случае системы ВИЖЕНЕРА — стандартного алфавита, а в случае системы АЛЬБЕРТИ — смешанного алфавита, попавшего в незаконные руки) появляется в литературе, хотя и без подсчета *Xu*, довольно рано. Общий способ использования лент (см. 12.8.1) состоял в печати наиболее частых символов первичного алфавита (в английском языке это 9 символов: etaonirsh) жирным шрифтом или красным цветом — в механических решениях во время Второй мировой войны (Витт, Рорбах) использовалась полупрозрачная бумага; а также в выбрасывании самых редких символов (в английском языке это 5 символов: jkqxz). Используя столбцы криптотекста в качестве строк, соответствующий открытый текст следует искать в какой-то другой строке. Наиболее естественно брать строку с макси-

Выравнивание	*Xu*	
a ≙ A	4.05%	
a ≙ B	3.54%	
a ≙ C	3.70%	
a ≙ D	2.64%	
a ≙ E	4.38%	
a ≙ F	5.54%	←
a ≙ G	4.07%	
a ≙ H	2.97%	
a ≙ I	2.98%	
a ≙ J	5.13%	←
a ≙ K	4.43%	
a ≙ L	3.60%	
a ≙ M	1.72%	
a ≙ N	4.30%	
a ≙ O	4.85%	
a ≙ P	4.11%	
a ≙ Q	3.01%	
a ≙ R	2.77%	
a ≙ S	4.71%	
a ≙ T	3.59%	
a ≙ U	5.38%	←
a ≙ V	3.71%	
a ≙ W	3.37%	
a ≙ X	2.65%	
a ≙ Y	4.18%	
a ≙ Z	4.62%	

Таблица 20. Подсчет значений *Xu* для стандартного алфавита по сравнению с первым столбцом

мальным числом символов, при условии, что строка не имеет или имеет мало отсутствующих символов. Это показано на рис. 138 для первого столбца

G I Y L B N S N J W X C S N S G

из примера Калпа (рис. 128). Ясно выделяется строка, маркированная a $\hat{=}$ U. Строка, маркированная a $\hat{=}$ F, имеет на один жирный символ больше, но в то же время имеет небольшой недостаток — присутствует :. Как видно из табл. 20, для первого столбца проблема возникает с выбором между двумя ключами U и F. Остальные строки, например, a $\hat{=}$ J имеют более низкие значения, минимум достигается при a $\hat{=}$ M.

18.2.3. Дополнительная помощь. Как только определяется сдвиг второго столбца, мы можем надеяться, что далее нам помогут частоты биграмм (в нашем случае — в выборе между конкурирующими ключами U и F). Таким образом, определение одних индивидуальных букв ключа помогает при определении других букв.

18.2.4. Метод скользящей планки. Родственный (ленточному) метод использует скользящую планку или диск, в точности соответствующие оригиналу, на которые нанесен стандартный или смешанный алфавит. Наиболее частые буквы со стороны открытого текста выделяются (например, жирным шрифтом), а наиболее редкие опускаются. Со стороны криптотекста отмечаются наблюдаемые частоты (например, штрихами). Так, для букв столбца

G I Y L B N S N J W X C S N S G

частоты маркируются так (рис. 139):

$$\mathrm{A\overline{B}\overline{C}DEF\overline{\overline{G}}H\overline{I}\overline{J}K\overline{L}M\overline{\overline{\overline{N}}}OPQR\overline{\overline{S}}TUV\overline{W}\overline{X}\overline{Y}Z}$$

Затем две планки (или два диска) начинают двигаться в противоположных направлениях до тех пор, пока не произойдет самая «жирная» встреча.

18.2.5. Применимость метода. Такой метод «управляемого перебора» можно использовать во всех случаях, когда известно, как из первичной подстановки получить все другие подстановки. В частности, это можно сделать

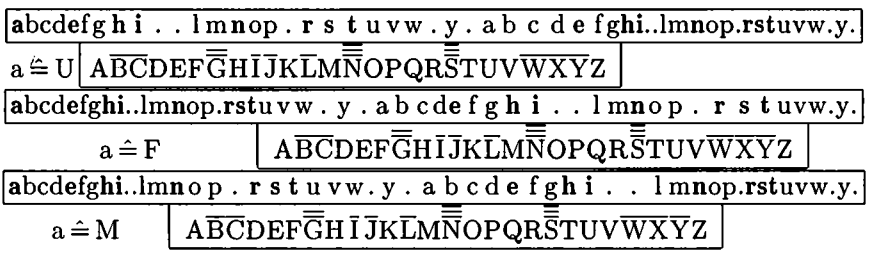


Рис. 139. Планка для дешифрования столбца разорванного текста. (a $\hat{=}$ U и a $\hat{=}$ F — хорошие соответствия, a $\hat{=}$ M — плохое соответствие)

```
11 8 12 9 12 8 3 4 10 21 5 7 19 9 20 10 8 20 12 4 8 15 22 13 7 26
A B C D E F G H I  J K L M N O P Q R S T U V W X Y Z
```

Рис. 140. Частотное распределение в криптотексте из разд. 18.3.1

в схеме шифрования РОТОР. Глубина не менее 6–9 приводит к цели, т. е. для успеха обычно требуется 6–9 букв криптотекста на букву ключа

Другими словами, для предотвращения подобной атаки длина открытого текста не должна превышать шестикратную длину ключа.

Кроме того, этот метод можно также применять и для многоалфавитных шифров с произвольными неродственными алфавитами. В частности, это возможно, если неуполномоченному читателю доступен цилиндр M-94, или ленточный аппарат CSP-642, или результат группировки по семействам (разд. 14.3.6). В таком случае остается лишь сопоставить столбец разорванного открытого текста с каждым отдельным алфавитом. Правда, глубина столбцов при этом будет часто чересчур короткой для успешного применения этого метода. Обычно требуется не менее 40–50 букв криптотекста на букву ключа, чтобы добиться успеха.

Другими словами, для предотвращения такой атаки открытый текст не должен быть длиннее нескольких длин ключа.

18.3. Взаимное выравнивание сопутствующих алфавитов

Если исходный алфавит неизвестен, то все же существует возможность взаимного выравнивания сопутствующих алфавитов и замены многоалфавитного криптотекста одноалфавитно зашифрованным промежуточным криптотекстом, который можно обработать методами глав 12–15. Эта процедура полезна также, если исходный алфавит является стандартным алфавитом, но этот факт не известен, скажем, потому, что криптотекст чересчур короткий.

18.3.1. Пример. Следующий криптотекст Синкова (1968 г.) из 303 символов имеет хорошо сбалансированное частотное распределение, показанное на рис. 140, и наводящее на мысль, что шифр не является одноалфавитным.

```
SWWJR GPRDN FMWJE XEWGR ZJQDN VJZRV
SZXOJ VWWRO VBHRM MOFDL IPAXV EZWUT
CZOZA AQQJL UPKZZ XUMJA PCZOE BAWZR
ZYKZI POFOL UOCRE NYKRI CAMOX IOORR
ZJKOL VWWJN VPKZA AFOCA MZOMR CJZDY
EJXEL XRFQI ZJCMA RJVWI DSWZX ASOTR
BJBZO QPXMI PDJVZ ZXHGQ SZFDQ FJZJR
BMWIC EZMWL MECVY VWZOX TWHSR UUBMT
NSJDW SSOOW CUNJY VJEWI PPFSL MOQVY
CVWRI SMMHW XMEJY NUZMV MXWCR NBRDE
SNB
```

Значение Xu составляет 4.58%, и это подтверждает наше предположение. Имеется 9 повторений длины 3, но более длинных нет; расстояния между повторениями таковы:

$$\begin{array}{lcl}
\text{WWJ} & : & 125 = 5\cdot 5\cdot 5 \\
\text{RZJ} & : & 100 = 2\cdot 2\cdot 5\cdot 5 \\
\text{JVW} & : & 132 = 2\cdot 2\cdot 3\cdot 11 \\
\text{VWW} & : & 90 = 2\cdot 3\cdot 3\cdot 5 \\
\text{CZO} & : & 21 = 3\cdot 7 \\
\text{ZAA} & : & 70 = 2\cdot 5\cdot 7 \\
\text{PKZ} & : & 60 = 2\cdot 2\cdot 3\cdot 5 \\
\text{ZZX} & : & 121 = 11\cdot 11 \\
\text{CAM} & : & 28 = 2\cdot 2\cdot 7.
\end{array}$$

Анализ Касиски бессилен различить два возможных периода 5 и 7. Однако этого можно добиться с помощью анализа Кульбака. Записывая глубину для 7 столбцов, мы получаем довольно низкое значение $Xu^{(7)} = 4.44\%$, тогда как глубина для 5 столбцов дает значения ϕ_r, приведенные на рис. 141 a, откуда находим намного большее значение

$$Xu^{(5)} = 5\cdot(190+243+258+262+240)/(303\cdot 302) = 6.51\%.$$

Поскольку частоты к тому же сильно несбалансированы, то можно предположить, что каждый столбец в действительности зашифрован одноалфавитно. Однако ни один из столбцов не имеет распределения частот, похожего на сдвинутый профиль, принадлежащий английскому, немецкому или французскому языку, что очевидно при взгляде на рис. 141 b. Таким образом, это напоминает не систему ВИЖЕНЕРА, а скорее более общую систему АЛЬБЕРТИ, и мы должны определить ее исходный алфавит.

18.3.2. Получение промежуточного криптотекста. Если бы текст был в 10 раз длиннее, мы могли бы попытаться дешифровать по отдельности каждый столбец, — и в конце, возможно, были бы удивлены, найдя что все алфавиты имеют общий первичный алфавит. Но при 61 или 60 символах в столбце текстовая основа для этого слишком мала. Поэтому мы можем лишь надеяться, предположив использование системы АЛЬБЕРТИ, что эти 5 алфавитов можно так взаимно выровнять, что получится одноалфавитно зашифрованный промежуточный криптотекст из 303 символов, вполне достаточный для применения стандартных методов.

Для взаимного выравнивания i-го и k-го столбцов, подсчитывается Xu для k-го столбца и циклически сдвинутого на q позиций i-го столбца при $q = 0, 1, \ldots, N-1$. Обычно в этой последовательности все значения колеблются в окрестности $\varkappa_{\mathrm{R}}$, за одним исключением, которое должно находиться в окрестности $\varkappa_{\mathrm{S}}$, и соответствующее значение q дает выравнивающий сдвиг. Таблица 21 показывает это для первого и второго столбцов с результирующим выравниванием $A^{(1)} \doteq X^{(2)}$.

1-й столбец

3 3 5 1 3 2 1 0 2 0 0 0 5 4 0 4 1 1 6 1 3 7 0 4 0 5
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z $\phi_1 = 196$

2-й столбец

2 2 1 1 2 1 0 0 0 1 0 0 0 4 1 5 6 1 1 4 0 4 1 5 2 2 6
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z $\phi_2 = 236$

3-й столбец

1 3 3 0 2 5 0 3 0 2 5 0 4 1 6 0 3 2 0 0 0 1 1 1 3 0 6
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z $\phi_3 = 258$

4-й столбец

0 0 2 7 1 0 2 1 1 8 0 0 5 0 7 0 1 7 2 1 1 3 3 1 0 7
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z $\phi_4 = 262$

5-й столбец

5 0 1 0 4 0 0 0 7 1 0 7 1 3 2 0 2 9 0 2 0 3 3 3 5 2
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z $\phi_5 = 240$

Рис. 141 a. Частотное распределение для 5 столбцов и значения ϕ_r

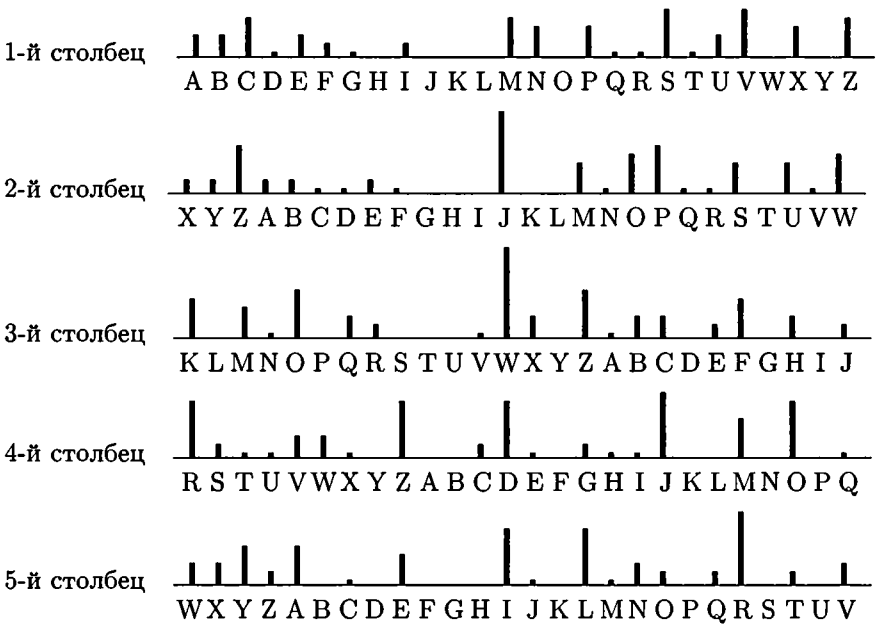


Рис. 141 b. Выровненные профили для пяти столбцов

Выравнивание	Xu
$A^{(1)} \triangleq A^{(2)}$	$157/61^2 = 4.22\%$
$A^{(1)} \triangleq B^{(2)}$	$133/61^2 = 3.57\%$
$A^{(1)} \triangleq C^{(2)}$	$162/61^2 = 4.35\%$
$A^{(1)} \triangleq D^{(2)}$	$122/61^2 = 3.28\%$
$A^{(1)} \triangleq E^{(2)}$	$144/61^2 = 3.87\%$
$A^{(1)} \triangleq F^{(2)}$	$138/61^2 = 3.71\%$
$A^{(1)} \triangleq G^{(2)}$	$102/61^2 = 2.74\%$
$A^{(1)} \triangleq H^{(2)}$	$170/61^2 = 4.57\%$
$A^{(1)} \triangleq I^{(2)}$	$119/61^2 = 3.20\%$
$A^{(1)} \triangleq J^{(2)}$	$126/61^2 = 3.39\%$
$A^{(1)} \triangleq K^{(2)}$	$188/61^2 = 5.05\%$
$A^{(1)} \triangleq L^{(2)}$	$83/61^2 = 2.23\%$
$A^{(1)} \triangleq M^{(2)}$	$160/61^2 = 4.30\%$
$A^{(1)} \triangleq N^{(2)}$	$133/61^2 = 3.57\%$
$A^{(1)} \triangleq O^{(2)}$	$165/61^2 = 4.43\%$
$A^{(1)} \triangleq P^{(2)}$	$137/61^2 = 3.68\%$
$A^{(1)} \triangleq Q^{(2)}$	$106/61^2 = 2.85\%$
$A^{(1)} \triangleq R^{(2)}$	$172/61^2 = 4.62\%$
$A^{(1)} \triangleq S^{(2)}$	$130/61^2 = 3.49\%$
$A^{(1)} \triangleq T^{(2)}$	$123/61^2 = 3.31\%$
$A^{(1)} \triangleq U^{(2)}$	$190/61^2 = 5.11\%$
$A^{(1)} \triangleq V^{(2)}$	$132/61^2 = 3.55\%$
$A^{(1)} \triangleq W^{(2)}$	$148/61^2 = 3.98\%$
$A^{(1)} \triangleq X^{(2)}$	$236/61^2 = 6.34\%$ ⟵
$A^{(1)} \triangleq Y^{(2)}$	$91/61^2 = 2.45\%$
$A^{(1)} \triangleq Z^{(2)}$	$154/61^2 = 4.14\%$

Таблица 21. Подсчитанные значения Xu для первого столбца относительно второго столбца

Синков указал разные стратегии выравнивания. «Цепная» стратегия: выравнивание 2-го столбца относительно 1-го, 3-го столбца относительно 2-го, 4-го относительно 3-го и т. д. «Звездная» стратегия: выравнивание 2-го столбца относительно 1-го, 3-го столбца относительно 1-го, 4-го относительно 1-го и т. д. Цепная стратегия имеет тот недостаток, что результат не может быть лучше, чем действие слабейшего выравнивания в цепи. Недостаток звездной стратегии в том, что выбор любого столбца может оказаться неудачным. Таким образом, иногда необходимо идти окольными путями.

Слабое выравнивание встречается, если более чем одно значение Xu выделяется на общем фоне. Такой случай возникает в нашем примере при

Выравнивание	*Xu*
$A^{(3)} \triangleq A^{(4)}$	$187/61 \cdot 60 = 5.11\%$
$A^{(3)} \triangleq B^{(4)}$	$86/61 \cdot 60 = 2.35\%$
$A^{(3)} \triangleq C^{(4)}$	$148/61 \cdot 60 = 4.04\%$
$A^{(3)} \triangleq D^{(4)}$	$164/61 \cdot 60 = 4.48\%$
$A^{(3)} \triangleq E^{(4)}$	$165/61 \cdot 60 = 4.51\%$
$A^{(3)} \triangleq F^{(4)}$	$117/61 \cdot 60 = 3.20\%$
$A^{(3)} \triangleq G^{(4)}$	$82/61 \cdot 60 = 2.24\%$
$A^{(3)} \triangleq H^{(4)}$	$231/61 \cdot 60 = 6.31\%$ ⟵
$A^{(3)} \triangleq I^{(4)}$	$122/61 \cdot 60 = 3.33\%$
$A^{(3)} \triangleq J^{(4)}$	$110/61 \cdot 60 = 3.01\%$
$A^{(3)} \triangleq K^{(4)}$	$143/61 \cdot 60 = 3.91\%$
$A^{(3)} \triangleq L^{(4)}$	$109/61 \cdot 60 = 2.98\%$
$A^{(3)} \triangleq M^{(4)}$	$150/61 \cdot 60 = 4.10\%$
$A^{(3)} \triangleq N^{(4)}$	$229/61 \cdot 60 = 6.26\%$ ⟵
$A^{(3)} \triangleq O^{(4)}$	$53/61 \cdot 60 = 1.45\%$
$A^{(3)} \triangleq P^{(4)}$	$180/61 \cdot 60 = 4.92\%$
$A^{(3)} \triangleq Q^{(4)}$	$146/61 \cdot 60 = 3.99\%$
$A^{(3)} \triangleq R^{(4)}$	$103/61 \cdot 60 = 2.77\%$
$A^{(3)} \triangleq S^{(4)}$	$204/61 \cdot 60 = 5.57\%$
$A^{(3)} \triangleq T^{(4)}$	$108/61 \cdot 60 = 2.95\%$
$A^{(3)} \triangleq U^{(4)}$	$126/61 \cdot 60 = 3.44\%$
$A^{(3)} \triangleq V^{(4)}$	$190/61 \cdot 60 = 5.19\%$
$A^{(3)} \triangleq W^{(4)}$	$114/61 \cdot 60 = 3.11\%$
$A^{(3)} \triangleq X^{(4)}$	$124/61 \cdot 60 = 3.39\%$
$A^{(3)} \triangleq Y^{(4)}$	$145/61 \cdot 60 = 3.96\%$
$A^{(3)} \triangleq Z^{(4)}$	$124/61 \cdot 60 = 3.39\%$

Таблица 22. Подсчитанные значения *Xu* для третьего столбца относительно четвертого

подсчете *Xu* для 3-го и 4-го столбцов. Как видно из табл. 22, почти нет разницы между $A^{(3)} \triangleq H^{(4)}$ и $A^{(3)} \triangleq N^{(4)}$. Можно исследовать несколько вариантов, а можно обойти слабые выравнивания. В нашем случае оказывается, что как цепная, так и звездная стратегии выбирают выравнивание $A^{(3)} \triangleq H^{(4)}$. Это в конце концов приводит к выравниванию, показанному на рис. 141 b.

Таким образом, этот пример показывает, как периодическое шифрование АЛЬБЕРТИ при умеренно благоприятных условиях можно механически свести к промежуточному шифротексту, который, вероятнее всего, оказывается одноалфавитно зашифрованным.

В данном примере Синкова 60 букв криптотекста на одну букву ключа оказалось более чем достаточно. При этом все необходимые расчеты можно сделать на персональном компьютере.

18.3.3. Побочный результат. На рис. 141 b среди слов, читаемых вертикально, таких как AXKRW, BYLSX, CZMTY, ..., имеется слово ROBIN. Это, вероятно, и есть ключевое слово.

18.4. Восстановление исходного алфавита

Одноалфавитно зашифрованный промежуточный текст получается систематической заменой букв, как показано на рис. 141 b. Например, фрагмент SWWJR трактуется следующим образом:

$$\mathrm{SWWJR} = \mathrm{S}^{(1)}\mathrm{W}^{(2)}\mathrm{W}^{(3)}\mathrm{J}^{(4)}\mathrm{R}^{(5)} = \mathrm{S}^{(1)}\mathrm{Z}^{(1)}\mathrm{M}^{(1)}\mathrm{S}^{(1)}\mathrm{V}^{(1)} = \mathrm{SZMSV}^{(1)}$$

В общем, получается следующий промежуточный криптотекст, записанный в алфавите первого столбца:

```
SZMSV GSHMR FPMSI XHMPV ZMGMR VMPAZ
SCNXN VZMAS VEXAQ MRVMP ISQGZ ECMDX
CCEIE ATGSP USAID XXCSE PFPXI BDMIV
ZBAIM PRVXP URSAI NBAAM CDCXB IREAV
ZMAXP VZMSR VSAIE AIELE MCEVV CMPMC
EMNNP XUVZM ZMSVE RMLFM DVMIB AVECV
BMRIS QSNVM PGZED ZAXPU SCVMU FMPSV
BPMRG ECCFP MHSEC VZPXB TZXBV UXRVX
NVZMA SVEXA CXDSC VMUFM PSVBP MRGEC
CYMAM SPCQA XPUSC NXPVZ MAMLV NEHMI
SQR
```

Его частотное распределение таково:

```
20 10 21 7 19 6 7 4 14 0 0 3 40 9 0 23 5 13 24 2 8 31 0 20 1 16
A  B  C  D  E  F G H I  J K L M  N  O P  Q R S  T  U V W X Y Z
```

Поэтому входом является

$$\mathrm{M}^{(1)} \hat{=} \mathrm{e}, \quad \mathrm{V}^{(1)} \hat{=} \mathrm{t}, \quad \mathrm{S}^{(1)} \hat{=} \mathrm{a}.$$

Из частой встречаемости триграммы VZM $\hat{=}$ tZe получаем Z $\hat{=}$ h.

Работая в свободном стиле, можно из встречаемости /heat/ в порядке рабочей гипотезы предположить, что встречается и /temperature/; действительно, слово VMUFMPSVBPM $\hat{=}$ teUFePatBPe встречается даже дважды: в конце седьмой и в девятой строках. Это уже дает

$$\mathrm{U}^{(1)} \hat{=} \mathrm{m}, \quad \mathrm{F}^{(1)} \hat{=} \mathrm{p}, \quad \mathrm{P}^{(1)} \hat{=} \mathrm{r}, \quad \mathrm{B}^{(1)} \hat{=} \mathrm{u}.$$

В начале пятой строки имеется

$$\mathrm{VZMAXPVZMSRV} \hat{=} \mathrm{theAXrtheaRT} \hat{=} \mathrm{thenorthethecast}.$$

Таким образом,

$$A^{(1)} \hat{=} n, \quad X^{(1)} \hat{=} o, \quad R^{(1)} \hat{=} s.$$

Теперь дешифрование становится совсем легким, так как самые частые буквы e t a o n r s h и некоторые редкие буквы уже определены. Из фрагментарного дешифрования

```
a h e a t   G a H e s   p r e a I   o H e r t   h e G e s   t e r n h
a C N o N   t h e n a   t E o n Q   e s t e r   I a Q G h   E C e D o
C C E I E   n T G a r   m a п I D   o o C a E   r p r o I   u D e I t
h u n I e   r s t o r   m s a n I   N u n n e   C D C o u   I s E n t
h e n o r   t h e a s   t a n I E   n I E L E   e C E t t   C e r e C
E e N N r   o m t h e   h e a t E   s e L p e   D t e I u   n t E C t
u e s I a   Q a N t e   r G h E D   h n o r m   a C t e m   p e r a t
u r e s G   E C C p r   e H a E C   t h r o u   T h o u t   m o s t o
N t h e n   a t E o n   C o D a C   t e m p e   r a t u r   e s G E C
C Y e n e   a r C Q n   o r m a C   N o r t h   e п e L t   N E H e I
a Q s
```

шаг за шагом выводим:

$$G^{(1)} \hat{=} w, \quad H^{(1)} \hat{=} v, \quad I^{(1)} \hat{=} d, \quad C^{(1)} \hat{=} l,$$
$$N^{(1)} \hat{=} f, \quad E^{(1)} \hat{=} i, \quad Q^{(1)} \hat{=} y, \quad mD^{(1)} \hat{=} c,$$
$$T^{(1)} \hat{=} g, \quad L^{(1)} \hat{=} x, \quad Y^{(1)} \hat{=} b$$

и заканчиваем открытым текстом, который, очевидно, имеет смысл:

```
a h e a t   w a v e s   p r e a d   o v e r t   h e w e s   t e r n h
a l f o f   t h e n a   t i o n y   e s t e r   d a y w h   i l e c o
l l i d i   n g w a r   m a n d c   o o l a i   r p r o d   u c e d t
h u n d e   r s t o r   m s a n d   f u n n e   l c l o u   d s i n t
h e n o r   t h e a s   t a n d i   n d i x i   e l i t t   l e r e l
i e f f r   o m t h e   h e a t i   s e x p e   c t e d u   n t i l t
u e s d a   y a f t e   r w h i c   h n o r m   a l t e m   p e r a t
u r e s w   i l l p r   e v a i l   t h r o u   g h o u t   m o s t o
f t h e n   a t i o n   l o c a l   t e m p e   r a t u r   e s w i l
l b e n e   a r l y n   o r m a l   f o r t h   e n e x t   f i v e d
a y s
```

К этому моменту исходный алфавит восстановлен вплоть до произвольного сдвига (с точностью до не встретившихся в тексте букв /j/, /k/, /q, /z/). Предполагая, что ROBIN действительно было ключевым словом, находим исходный алфавит, принадлежащий ключевой букве A, и соответствующую расшифровывающую подстановку:

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
A s a g m t * o b h n u l c i p w v d * * x e f * r y.
```

Теперь проявляется даже «второй» ключ, а именно пароль для построения алфавита: если алфавит последовательно записывается в пять столбцов построенного трафарета:

```
s o l v e
a b c d f
g h i * *
m n p * r
t u w x y
*
```

Внезапно появляется пароль /solve/. В результате мы можем пополнить алфавит:

```
s o l v e
a b c d f
g h i j k
m n p q r
t u w x y
z
```

Теперь можно привести полный исходный алфавит, принадлежащий ключевой букве A:

```
a b c d e f g h i j k l m n o p q r s t u v w x y z
A B H M R V W C I N S X L D J G O T Y A E K Q P U Z F.
```

Этим результатом дешифрование завершено в смысле Ропбаха.

Оказывается, что неслучайными повторениями являются RZJ и VWW, которые оба расшифровываются как /the/; PKZ = /and/; WWJ = /hea/ в hea(t) и (nort)hea(st); ZAA = /din/ в (colli)din(g) и в (an)d in. Совершенно невероятно появление четырех случайных повторений JVW, CZO, ZZX, CAN, они имеют расстояния 132, 21, 121 и 28, ни одно из которых не содержит множителя 5.

18.5. Симметрия позиции Керкхоффа

В разд. 18.4 по методическим причинам был проведен строго частотный анализ (без привлечения других методов). Однако часто существуют подсказки, позволяющие угадывать вероятные слова и приводящие к анализу шаблонов, которые в свою очередь обеспечивают дешифрование редких символов в некоторых сопутствующих алфавитах. В 1883 г. Керкхоффс открыл, что в подходящих случаях бывает возможно из дешифрования символов одного столбца получить дешифрование символов другого столбца. Он назвал это свойство сопутствующих алфавитов (основанное на коммутативности сложения, гл. 5) *symétrie de position* (симметрией позиции, *фр.*). Метод обычно называемый именем Керкхоффса, объясняется в следующем оригинальном примере Керкхоффса.

18.5.1. Пример. Пусть задан криптотекст

```
RBNBJ  JHGTS  PTABG  JXZBG  JICEM  QAMUW
IVGAG  NEIMW  REZKZ  SUABR  RBPBJ  CGYBG
JJMHE  NPMUZ  CHGWO  UDCKO  JKKBC  PVPMJ
NPGKW  PWADW  CPBVM  RBZBH  JWZDN  MEUAO
JFBMN  KEXHZ  AWMWK  AQMTG  LVGHC  QBMWE
```

и предположим, что анализ Касиски повторений биграмм RB, BJ, BG, RE, MJ, PQ вызвал подозрение, что шифр является многоалфавитным с периодом 5, и возможно, со сдвинутыми сопутствующими алфавитами. Из телеграммы от 2 сентября, посланной из Лондона в Каир агентству Гавас, Керкхоффс взял в качестве *mots probables* (вероятных слов, *фр.*) следующий список: *Arabie*, *Wolseley*[1], *Suez*, *Ismaïlia*, *canal*, *général*, *soldats*. Сначала он сделал частотный анализ первых пяти столбцов и выяснил, что в первом столбце $J^{(1)} \hat{=} c$, во втором и четвертом $B^{(2)} \hat{=} B^{(4)} \hat{=} e$, в третьем $M^{(3)} \hat{=} e$ и в пятом $Z^{(5)} \hat{=} c$. Это дает частичное дешифрование

```
RBNBJ  JHGTS  PTABG  JXZBG  JICEM    QAMUW
*e*e*  e****  ****e*  e***e*  e***  *  **e**
```

и он попытался в этих обстоятельствах выдвинуть гипотезу, что криптотекст направлен *le général Wolseley...* Это дало вход:

```
RBNBJ  JHGTS  PTABG  JXZBG  JICEM    QAMUW
legen  eralw  olsel  ey*e*  e***  *  **e**
```

Таким образом, $G^{(5)} \hat{=} l$, и тогда мы пытаемся продолжить словом *télégraphie*:

```
RBNBJ  JHGTS  PTABG  JXZBG  JICEM    QAMUW
legen  eralw  olsel  eytel  egra p    hie**
```

Такова предыстория. Теперь начинается специальный метод. До сих пор мы имели

```
     a b c d e f g h i j k l m n o p q r s t u v w x y z
(1)        J     Q     R     P
(2)        B  I  A  T           H              X
(3)  G     M  N                 C A Z
(4)  E     B           T
(5)        Z           G  J  M           S
```

Вводится симметрия позиции. Начнем с того, что вторая и четвертая строки должны быть идентичны, так как $B^{(2)} \hat{=} B^{(4)} \hat{=} e$ и $T^{(2)} \hat{=} T^{(4)} \hat{=} l$. Поэтому

```
     a b c d e f g h i j k l m n o p q r s t u v w x y z
(2)  E     B  I  A  T           H              X
(4)  E     B  I  A  T           H              X
```

[1] Лорд Уолсли, главнокомандующий Британской армии.

Поскольку J встречается в первой и пятой строках, причем $J^{(1)} \hat{=} c$, $J^{(5)} \hat{=} n = e + 9$, мы заключаем, что вся пятая строка сдвинута на 9 позиций вправо. Это приводит к определению восьми символов криптотекста:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
(1)				G		J		M	Q			R		S	P						Z					
(5)						Z						G		J		M	Q			R		S	P			

Но третья и пятая строки тоже связаны между собой, помимо прочих, условиями: $A^{(3)} \hat{=} e$, $A^{(5)} \hat{=} p = e + 11$. Это приводит к фиксации одиннадцати символов криптотекста:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
(1)				G		J		M	Q	N		R		S	P					C	A	Z				
(3)	G		J		M	Q	N			R		S	P					C	A	Z						
(5)				C	A	Z						G		J		M	Q	N			R		S	P		

Наконец, вторая и третья строки связаны, кроме всего прочего, условиями $A^{(2)} \hat{=} i$, $A^{(3)} \hat{=} s = i + 11$. Это теперь дает связи со всеми пятью алфавитами и фиксацию семнадцати символов криптотекста:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
(1)			G	H	J		M	Q	N		X	R	E	S	P		B		I	C	A	Z		T		
(2)	E	S	P		B		I	C	A	Z		T					G	H	J		M	Q	N		X	R
(3)	G	H	J		M	Q	N		X	R	E	S	P		B		I	C	A	Z		T				
(4)	E	S	P		B		I	C	A	Z		T					G	H	J		M	Q	N		X	R
(5)		I	C	A	Z		T				G	H	J		M	Q	N		X	R	E	S	P		B	

Еще не произведено дешифрование девяти символов криптотекста: D, F, R, L, O, U, V, W, Y. Однако можно ожидать, что с 17 расшифрованными символами из 26, дальнейшее дешифрование является пустяком. Действительно, фрагментарная расшифровка первых трех строк телеграммы

RBNBJ	JHGTS	PTABG	JXZBG	JICEM	QAMUW
legen	eralw	olsel	eytel	egrap	hie**
IVGAG	NEIMW	REZKZ	SUABR	RBPBJ	CGYBG
s*ail	iaqu*	latte	n*seu	lemen	tq*el
JJMHE	NPMUZ	CHGWO	UDCKO	JKKBC	PVPMJ
eserv	ice*e	tra**	**r**	e***ec	ommun

связывает V и W с вероятным словом *Ismailia*. Очевидное заполнение брешей дает U и Y, и если в третьей строке узнается *transport*, то определяются D, K и O. Символы F и L встречаются по одному разу (в пятой строке телеграммы), и их определение труднее.

Но уже существует лучший путь привести дешифрование к концу: должен появиться пароль, используемый в построении алфавита, и им является, очевидно, RESPUBLICA. Таким образом, мы можем выписать все пять используемых алфавитов:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
(1)	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W
(2)	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R
(3)	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F
(4)	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R
(5)	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B

В пополненном столбце под буквой /a/ открытого текста появляется ключевое слово DEGEL (оттепель, *фр.*). Наконец, приводится tabula recta (целая шифровальная таблица, *лат.*) — табл. 23. Остается только выяснить, какая строка является первой. Следуя и здесь Керкхоффсу, мы выбираем строку с паролем в конце. Если строки, т. е. алфавиты нумеруются буквами от A до Z, то ключевым словом длины 5 является FRHRW. Но верный ключ DEGEL использует столбец под буквой /a/ открытого текста.

Открытый текст гласит: «*Le général Wolseley télégraphie et de communication soit complement organisé pour faire une nouvelle marche en v...*».

Суммируя, можно сказать, что «симметрия позиции» позволяет «вырвать больше открытого текста из малого количества криптотекста» (Дэвид Кан).

18.5.2. Volapük (Воляпюк). Керкхоффс (Fleming Auguste Kerckhoffs) (полный список данных ему имен: Jean-Guillaume-Hubert-Victor-François-Alexandre, его дворянское имя было von Nieuwenhof) родился 19 января 1835 г. в Нуте, герцогство Лимбург (теперь в Бельгии). Он ходил в школу около Аахена, учился после пребывания в Англии в университете Luik (Льеж), преподавал в высшей школе современные языки и долго работал разъездным секретарем, пока наконец не занял академическую должность в Мелуне (юго-восток Парижа). Как учитель он был довольно эксцентричным, но очень активным в учебных контактах. В 1873 г. он стал французским гражданином, в 1873–1876 гг. он учился в университетах Бонна и Тюбингена и стал доктором филологии. В 1878 г. Керкхоффс получил кафедру немецкого языка в двух высших учебных заведениях Парижа. Его первым вкладом в криптологию была написанная 1882 г. работа «Военная криптография». Эта 69-страничная статья, опубликованная в «Журнале военных наук» в январе и феврале 1883 г. вместе с работой Касиски (1863 г.) заложила основы научной криптографии XIX в.

Однако для большинства людей основа славы Керкхоффса — в его ревностной и трагической поддержке интернационального языка Воляпюк (Volapük), предложенного в 1879 г. Шлейером. В 1887 г. Керкхоффс был назначен Директором международной академии Воляпюк. Подобно эсперанто (1887 г.) и другим подобным предложениям, Воляпюк не смог утвердить себя. Керкхоффс жил достаточно долго, чтобы увидеть упадок Воляпюка, и умер от разрыва сердца в 1903 г.

18.5.3. Пример с сюрпризом. Симметрия позиции, разумеется, полезна также и когда речь идет о системе ВИЖЕНЕР. Мы покажем это для криптотекста Калпа (рис. 115), предположив сначала, что использована систе-

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A
B	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z
C	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y
D	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T
E	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V
F	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W
G	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D
H	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F
I	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G
J	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H
K	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J
L	M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K
M	Q	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M
N	N	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q
O	O	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N
P	X	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O
Q	R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X
R	E	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R
S	S	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E
T	P	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S
U	U	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P
V	B	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U
W	L	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B
X	I	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L
Y	C	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I
Z	A	Z	Y	T	V	W	D	F	G	H	J	K	M	Q	N	O	X	R	E	S	P	U	B	L	I	C

Таблица 23. Таблица алфавитов (tabula recta) для примера Керкхоффса

ма АЛЬБЕРТИ и что существуют причины ожидать периода 12. Сначала выпишем глубину двенадцать для столбцов, как на рис. 128 и обнаружим, что символы одного и того же столбца часто обнаруживают расстояние 4, 7 или 11:

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
G	E	I	E	I	A	S	G	D	X	V	Z
I	J	Q	L	M	W	L	A	A	M	X	Z
Y	Z	M	L	W	H	F	Z	E	K	E	J
L	V	D	X	W	K	W	K	E	T	X	L
B	R	A	T	Q	H	L	B	M	X	A	A
N	U	B	A	I	V	S	M	U	K	H	S
S	P	W	N	V	L	W	K	A	G	H	G
N	U	M	K	W	D	L	N	R	W	E	Q
J	N	X	X	V	V	O	A	E	G	E	U
W	B	Z	W	M	Q	Y	M	O	M	L	W
X	N	B	X	M	W	A	L	P	N	F	D
C	F	P	X	H	W	Z	K	E	X	H	S
S	F	X	K	I	Y	A	H	U	L	M	K
N	U	M	Y	E	X	D	M	W	B	X	Z
S	B	C	H	V	W	Z	X	P	H	W	L
G	N	A	M	I	U	K					

Фактически, мы находим шесть троек с этими расстояниями:

	(3)	(4)	(6)	(7)	(10)	(12)
	B	M	W	L	M	L
+7						
	I	T	D	S	T	S
+4						
	M	X	H	W	X	W

Отсюда видно также, что как 4-й и 11-й столбцы, так и 7-й и 12-й подчинены одному и тому же ключу.

Применяя систематическую процедуру, образуем для каждого столбца парные разности между символами криптотекста и выпишем их в таблицу. Этот «разностный метод» лишь один из вариантов *симметрии позиции*. В нашем случае она показывает превосходство разностей 7, 4, 11 (или дополнительных разностей 22, 19, 15) и, следовательно, приводит к новым случаям появления разностей 4, 7 или 11:

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	N		B	M		W	L	M		M	X	L
+7												
		N	I	T	E	D	S		A	T	E	S
+4												
	Y	R	M	X	I	H	W	X	E	X		W
Сдвиг	0	19	14	25	10	9	24	25	6	25	10	24

Заметим, что в средней строке проявляется слово UNITEDSTATES. Сдвиги, требуемые для выравнивания столбцов относительно первого столбца, указаны в нижней строке. Они выявляют ключ; дешифрование относительно этого

ключа сводит криптотекст к одноалфавитно зашифрованному промежуточному криптотексту. Вычитание соответствующих ключевых букв дает начало этого текста

G L U F Y R U H X Y L B
I Q C M C N N B U N N B
Y G Y M M Y H A Y L U L
L C P Y M B Y L Y U N N.

Дешифрование при помощи частотного анализа не встречает никаких проблем, и промежуточное шифрование оказывается сложением ЦЕЗАРЯ — мы этого вовсе не использовали — с U $\hat{=}$ a; т. е. дешифрование выполняется путем отсчета вперед шести мест в алфавитном порядке. Начало открытого текста поэтому таково (ср. с разд. 18.1.2):

m r a l e x a n d e r h
o w i s i t t h a t t h
e m e s s e n g e r a r
r i v e s h e r e a t t.

Возникший здесь «разностный метод» окажется полезным и при отслаивании перешифрованного кода в следующем параграфе, поскольку он не использует частотного анализа. Если бы в метод симметрии позиции были включены и частотные соображения, то можно было бы догадаться, что в отмеченных выше тройках букве последней строки, как наиболее частой, соответствует /e/, и, значит, букве первой строки соответствует /t/, а букве второй строки /a/. А поскольку /a/ является нулевым элементом в $\mathbb{Z}_{26}$, это наблюдение проясняет, почему проявляется ключевое слово UNITEDSTATES. (Имеются также и «плохие» разности, например, в 3-м столбце M и X имеют разность +11.)

18.6. Отслаивание перешифрования. Разностный метод

Подводя итог обсуждению разд. 18.3, мы видим, что (взаимное) выравнивание сопутствующих алфавитов не имеет отношения к открытому тексту, который появляется лишь из одноалфавитно зашифрованного промежуточного текста (разд. 18.4). Это отражает тот факт, что схемы шифрования АЛЬБЕРТИ являются композициями: одноалфавитная функциональная подстановка, следующая за многоалфавитным сложением ВИЖЕНЕРа.

18.6.1. Отслаивание. Таким образом, техника разд. 18.3, 18.5 также применима и для перешифрованного кода, т. е. для композиции (разд. 9.2.2) кодирования с последующим шифрованием ВИЖЕНЕРа либо над $\mathbb{Z}_{26}$ (буквенный код), либо над $\mathbb{Z}_{10}$ (числовой код). Последний случай называется на английском жаргоне «отслаиванием числовой добавки от зашифрованного кода». Зашифрованный код будем называть для краткости в дальнейшем «шифрокодом», а код, взятый из кодовой книги, для краткости будем называть в дальнейшем «простокодом».

Предполагая определенную ширину простокода (обычно известную, например, 5) и определенный период, скажем, 15, образуем столбцы одинаково перешифрованных слов (столбцы шифрокодовых групп). Часто встречающиеся слова или фразы открытого текста приводят к частым разностям в каждом столбце шифрокодовых групп. Если материал является достаточно объемным, может оказаться возможным вычислить *Хи* двух столбцов, и тем самым облегчить выравнивание. Однако часто материал не настолько богат, чтобы взаимным выравниванием можно было установить упомянутый шифрокод.

18.6.2. Симметрия позиции. Но существует еще *симметрия позиции*. Две шифрокодовые группы, которые выделяются в двух столбцах, принадлежат одному простокоду тогда и только тогда, когда их разности равны разностям добавок, принадлежащих этим столбцам. Заметим, что эти разности, согласно Шеннону, являются классами вычетов линейных многосимвольных подстановок.

Для примера предположим, что имеется три столбца, и в каждом находятся три выделяющихся шифрокодовых группы.

(1)	(2)	(3)
47965	60597	27904
69451	34689	41537
11057	10056	26443

Если число 47965 из первого столбца шифрокодовых групп и число 60597 из второго их столбца принадлежат одному и тому же простокоду, то число 11057 из первого столбца шифрокодовых групп и число 34689 их второго их столбца тоже принадлежат одному и тому же простокоду, так как[2)]

$$47965 - 11057 = 60597 - 34689 = \mathbf{36918}.$$

Чтобы систематически находить такие совпадения для каждого столбца шифрокодовых групп, в этом примере все взаимные разности этого столбца подсчитываются в виде 3×3-матрицы; получим три таких матрицы:

(1)			(2)			(3)		
00000	88514	**36918**	00000	**36918**	50541	00000	**86477**	01561
22596	00000	58404	**74192**	00000	**24633**	**24633**	00000	25194
74192	52606	00000	50569	**86477**	00000	09549	85916	00000

С этой информацией мы находим также соотношения между шифрокодовыми группами в разных столбцах.

Внутри первого и внутри второго столбца мы имеем

$$47965 - 11057 = 60597 - 34689 = 36918.$$

[2)] Сложение (и вычитание) шифрокодовых групп поразрядное по модулю 10. — *Прим. перев.*

Поэтому между первым и вторым столбцами имеется соотношение:

$$47905 - 60597 = 11057 - 34689 = 87478.$$

Внутри второго и внутри третьего столбцов находим

$$34689 - 10056 = 41537 - 27904 = 24633.$$

Поэтому между вторым и третьим столбцами имеем соотношение:

$$34689 - 41537 = 10056 - 27904 = 93152.$$

Сводя второй столбец шифрокодовых групп (2) относительно первого столбца (1) прибавлением всюду в (2) разности 87478, и сводя третий столбец шифрокодовых групп (3) относительно первого столбца (1) прибавлением всюду в (3) разности 87478 (для (2) относительно (1)) и разности 93152 (для (3) относительно (2)) — поэтому вместе разность 70520, — мы получаем:

	(1′)	(2′)	(3′)
	47965	47965	97424
	69451	11057	11057
	11057	97424	96963
Сдвиг	0	87478	70520

Наконец следует посмотреть в первом столбце шифрокодовых групп, не встречается ли еще (хоть и более редко) группа 97424, а также не встречается ли еще в (приведенном) третьем столбце шифрокодовых групп (хоть и более редко) группа 47965; кроме того, нужно рассмотреть вопрос о встречаемости групп 69451, 97424, 96963. Если повезет, в дальнейшем можно обнаружить используемые простокоды.

18.6.3. Использование машин. Указанная выше процедура при всей своей логической простоте требует громоздких расчетов, и неудивительно, что уже в 1920-х гг. криптоаналитики искали пути для осуществления автоматизации выравнивания. Перфокарточное оборудование было доступно и подходило для этой работы. Во время Второй мировой войны англичане (Дж. Тилтмэн, сентябрь 1939 г.), американцы (Р. Дж. Фабиан, ноябрь 1940 г.) и немцы такое оборудование использовали.

Затем были построены специальные устройства. В шифровальном отделе немецкого Вермахта был спроектирован «аппарат подсчета разностей», который обрабатывал шифрокодовые группы, перфорированные на ленте с помощью механических сканеров и релейных схем. Он обрабатывал семь различных пятизначных групп в секунду и записывал результат при помощи пишущей машинки, что было в 10-15 раз быстрее расчетов проводимых человеком с максимальной скоростью.

В противоположность этим цифровыми методам, аналоговые устройства с фотоэлектрическими измерителями тетраграмм были использованы для определения наиболее часто встречающихся простокодов — как в отделе *Xu* OKW,

так и в Спецслужбе Далема Министерства иностранных дел. Для такого приведения, т. е. для вычитания разности из столбца шифрокодовых групп с известным относительным базисом, математиком Виттом было спроектировано специальное оптическое аналоговое устройство.

Меньше известно о специальных устройствах, применявшихся Союзниками во Второй мировой войне. Машины HEATH ROBINSON и COLOSSUS были построены в Блетчли Парк, оперировали двоичными числами и использовались, главным образом, против телетайпных шифромашин вроде машины Лоренца SZ 42. В США машины COPPERHEAD 1943 г., аналогичные HEATH ROBINSON, также работали с оптическим сканированием и применялись против японских перешифрованных кодов. Сравнимой, возможно, с немецкими разработками, была машина TESSIE 1942 г., построенная компанией Истмен для военно-морского флота. Эта машина работала с фотоэлектрическими измерителями и использовалась для нахождения четырехзначных шифрокодовых групп, необходимых для отслаивания перешифрования. Она была направлена как против японского военно-морского кода высшего уровня, так и против немецких «кратких ручных сигналов» подводных лодок, применявшихся для сообщений вспышками о дислокации и пеленговавшихся Союзниками.

18.7. Дешифрование кода

В самом конце, после отслаивания перешифрования, остается дешифрование промежуточного шифрокода, т. е. восстановление кодовой книги («книгостроительство»). Промежуточный шифрокод сдвинут относительно простокода на константу, но это не имеет никакого отношения к той работе, которая должна быть сделана, и которая большей частью является лингвистической по своей природе. Скорее здесь уместны методы гл. 15. Эта работа сильно упрощается, если код является одночастным кодом (разд. 4.4.2); тогда кодовая группа, лежащая между двумя группами с уже известными эквивалентами открытого текста, имеет эквивалент открытого текста «в промежутках». В этой точке, видимой или воображаемой, находят богатую возможность для приложения ассоциация и комбинация. Книгостроительство — это такая часть криптоанализа, где одна математика бесполезна. Систематическая трактовка лингвистической стороны криптоанализа была впервые предпринята в 1892 г. Валерио.

18.8. Восстановление пароля

Преимущество, предлагаемое сопутствующими алфавитами, состоит в том, что они порождаются из одного первичного алфавита. Это может помочь спешащему криптографу, если он легко вспомнит или построит первичный алфавит. Популярным средством для этого являются пароли (разд. 3.2.5). Восстановленные пароли не только дают незаконному дешифровальщику до-

полнительную уверенность, но также позволяют проводить систематические атаки. Поэтому применение осмысленных паролей является слабостью.

18.8.1. Фридман. На первый взгляд, то, что представил в 1917 г. Фридман, выглядело подобно трюку фокусника. Допустим, что первичный алфавит таков:

a b c d e f g h i j k l m n o p q r s t u v w x y z
N T U V P W X J F Y Z D K Q C A B O G R L I S H M E

Оказывается, он является одноцикловым с циклом

(a n q b t r o c u l d v i f w s g x h j y m k z e p).

Теперь, отправляясь от произвольного символа, например, /a/, применим эту подстановку повторно и результаты выпишем внизу с расстояниями 1, 3, 5:

N Q B T R O C U L D V I F W S G X . . . ,
N * * Q * * B * * T * * R * * O * . . . ,
N * * * * Q * * * * B * * * * T * . . . ,

и так далее. Тогда получим

1	N Q B T R O C U L D V I F W S G X H J Y M K Z E P A
3	N D J Q V Y B I M T F K R W Z O S E C G P U X A L H
5	N K X I C Q Z H F U B E J W L T P Y S D R A M G V O
7	N G R Y L E F Q X O M D P W B H C K V A S T J U Z I
9	**N T** C D F G J K P Q R U V W X Y Z **A B O L I S H M E**
⋮	⋮ ⋮

Оказывается, что с расстоянием 9 производится последовательность, содержащая пароль ABOLISHMENT.

Теперь, взяв на стороне открытого текста ту же последовательность, только сдвинутую на 9 мест вправо, получим следующую последовательность:

a b o l i s h m e n t c d f g j k p q r u v w x y z
9 N T C D F G J K P Q R U V W X Y Z A B O L I S H M E

Переупорядочивая, получаем исходный алфавит, чье построение из ключевого слова теперь выяснено: это 9-я степень цикла

(a b o l i s h m e n t c d f g j k p q r u v w x y z)

с паролем abolishment.

Трюк фокусника становится понятнее, если представить себе, что для каждого расстояния получается некоторый алфавит, из которого переупорядочением снова получается исходный алфавит, например, для расстояния 7:

a s t j u z i n g r y l e f q x o m d p w b h c k v
7 N G R Y L E F Q X O M D P W B H C K V A S T J U Z I

хотя это и не дает осмысленного пароля. Или с расстоянием 1 получается, очевидно,

```
  a n q b t r o c u l d v i f w s g x h j y m k z e p
1 N Q B T R O C U L D V I F W S G X H J Y M K Z E P A
```

На самом деле, третья степень этого алфавита восстанавливает пароль, так как $3 \times 9 \equiv 1 \pmod{26}$.

18.8.2. Снова Фридман. Уильям Фридман указал также в 1918 г. процесс восстановления пароля в самом общем случае (разд. 3.5.2) системы АЛЬБЕРТИ с паролями как со стороны открытого текста, так и со стороны криптотекста. Мы вернемся к этому в разд. 19.5.3.

Глава 19

Компромиссы

> Качество машины зависит от способа ее использования.
>
> *Борис Хагелин*

Среди криптографических ошибок, перечисленных в гл. 11, худшей являются компромиссы, потому что они открывают линии атаки. Вслед за компромиссом ОТ-КТ, рассмотренным в разд. 14.6 (где ОТ — открытый текст, а КТ — криптотекст), в этой главе мы познакомимся с компромиссами ОТ-ОТ и КТ-КТ.

19.1. Наложение Керкхоффса

Многоалфавитное шифрование с периодическим ключевым текстом даже с неизвестными и неродственными алфавитами не дает гарантии против незаконного дешифрования. Как только определен период (гл. 17), построение глубины (гл. 18) приводит к одноалфавитно зашифрованному ОТ, однако этот ОТ разорван, что делает дешифрование очень коротких текстов трудным или невозможным (разд. 18.2.5, 18.3.2).

Но даже если ключ не является периодическим или если его период сравним с длиной ОТ, методы гл. 18 могут быть применены всякий раз, когда несколько открытых текстов зашифровано одним и тем же ключом. Если криптотексты могут быть установлены в фазе с ключевым текстом, то этот компромисс ОТ-ОТ ключа тоже позволяет построить глубину, т. е. построить столбцы символов криптотекста или столбцы шифрокодовых групп, каждый из которых состоит из одноалфавитно зашифрованного (но еще разорванного) открытого текста. Август Керкхоффс рассмотрел эту ситуацию в своей статье 1883 года. Установка «в фазе» разных текстов называется их наложением (нем., *Überlagerung*). Наложение будет работать только в случае, когда большое число открытых текстов доступно и может быть согласовано; в частности, это является очень вероятным, когда используются шифровальные машины,

которые имеют одну и ту же или лишь слегка измененную стартовую позицию ключа (механически порожденного). Именно таким способом Тилтман[1] преуспел во взломе старой машины ENIGMA без штепсельного коммутатора (английское кодовое название «rocket»), такие машины использовались на немецких железных дорогах для передачи расписания движения транспортным составам во время Второй мировой войны.

Нетрудно видеть, что периодическое применение не слишком длинного ключа, которое приводит к нескольким повторениям, также означает компромисс ОТ-ОТ. Из приведенного замечания следует, что все методы, используемые для определения периода ключа, можно также использовать для теста: находятся ли в фазе различные открытые тексты, и если нет, то согласовать их. В этом случае будут как раз к месту тесты *Kanna* и *Xu*.

19.1.1. Пример. Следующий пример наложения, данный Керкхоффсом, предполагает, что все открытые тексты находятся в фазе. Тринадцать текстов дают простое упражнение (типографские ошибки исправлены):

	1 2 3 4 5	6 7 8 9 10	1112131415	1617181920	2122232425	26272829
(i)	U H Y B R	J I M B C	F A M M F	J H D M R	I Q	
(ii)	U H W P R	B Q L K I	B L W R E	J R B K L	H I X B Q	E X H M
(iii)	I E W H C	H Q K Q M	T M V G J	J E D Z V	A	
(iv)	U W V R R	H I K M C	W W R G H	D C X S R	Q H	
(v)	U H S H A	H K S V C	J W Z V X	J Y N D M	Q Q N	
(vi)	Y H V H M	A G Q K C	W X P V I	H H W L Z	V L T H V	
(vii)	L H V H A	A G R L P	F M S O H	I P W Z Z	J E L Q R	B W
(viii)	S W U I R	X I C J U	F S H G W	R S Z B A	A L	
(ix)	U H W H V	A Y U L C	J W O U K	D E B K Q		
(x)	Y W X H Y	H B A L G	B V P S W	I W W J R	R H	
(xi)	W Q R E X	B I E N H	M V Y M H	S I Y M		
(xii)	S W U H D	H P J J C	K X G M H	L		
(xiii)	G Q V Q R	V O T Q Q	S P W R			

Керкхоффс начинает с утверждения, что предположительно $H^{(2)} \hat{=} e$, $H^{(4)} \hat{=} e$, $R^{(5)} \hat{=} e$, $H^{(6)} \hat{=} e$, $I^{(7)} \hat{=} e$, $L^{(9)} \hat{=} e$, $C^{(10)} \hat{=} e$, и что из-за многих совпадений вторая, четвертая и шестая позиции подпадают под один и тот же ключ (он не стал предполагать, что $U^{(1)} \hat{=} e$).

Криптотекст (iv) тогда предварительно дешифруется как $****eee**\ldots$, что наводит на поиски слова, оканчивающегося на ée; подходит l'armée. Имеем (iv) larmeee$**\ldots$ Криптотекст (v) le$*e*e***$ наводит на (v) legeneral ... Криптотекст (vi) $*ere*v***\ldots$ оставляет выбор между (vi) serezvons ... или (vi) ferezvons ... Криптотекст (vii) $*erenvo*e*\ldots$ интерпретируется Керкхоффсом довольно уверенно как (vii) nerenvoyez ... Затем он продолжает работу с оставшимися текстами. Но вход он уже сделал.

[1] «...он был всегда обаятелен и интеллигентен, и хотя он выглядел военным, он конечно не вел себя подобно напыщенному ничтожеству» (Деннистон).

Наложение как методическая идея подходит также и для случая неродственных алфавитов. При условии, что используется не слишком много алфавитов, скажем, не более двух дюжин в случае односимвольных алфавитов, и что криптотекст достаточно длинный, чтобы большинство этих алфавитов использовалось по крайней мере несколько раз, тогда как только устанавливается тождество каких-нибудь алфавитов, эффективная глубина материала соответственно увеличивается. Если ключ имеет немецкое название, то в среднем каждый шестой его символ является буквой E, и таким образом, каждый шестой алфавит является одним и тем же. В случае английского языка положение лишь немногим лучше.

19.1.2. Симметрия позиции. В данном примере дальнейшее дешифрование должно было бы стать громоздким, если бы Керкхоффс не сделал предположения, что мы имеем дело с сопутствующими алфавитами, которые просто сдвинуты друг относительно друга, и следовательно, может быть использована симметрия позиции, кульминация его работы. В таком случае все идет подобно часовому механизму. Вот некоторые из расшифрованных сообщений (на языке французской северной Африки):

(i)	leprefetdepoliceestici	«*le préfet de police est ici*»
(ii)	lespertesdelennemisontgrandes	«*les pertes de l'ennemi sont grandes*»
(iii)	onsemetsurladefencive	«*on se met sur la défencive*»
(iv)	larmeeestentreeauCaire	«*l'armee est entrée au Caire*»
(v)	legeneralestaAlexandrie	«*le general est à Alexandrie*»
(vi)	serezvousenetatderesister	«*serez vous en état de résister*»
(vii)	nerenvoyezpaslesprisonnier	«*ne renvoyez pas les prisonnier.*»

Оказывается, что Керкхоффс использовал тот же самый шифр АЛЬБЕРТИ, что и в разд. 18.5.1. Ключ является периодическим, он может быть восстановлен таблицей шифрования (табл. 23) и имеет вид

JEMEMETSURLADEFENSIVE.

19.2. Наложение для шифров с ключевой группой

При благоприятных обстоятельствах даже экстремальный случай (не рассматриваемый Августом Керкхоффсом) наложения только двух криптотекстов, зашифрованных одним и тем же ключом, не безнадежен, если известны алфавиты.

19.2.1. Чистое шифрование. Предположим в этом пункте, что криптосистема не только (как обычно) инъективная и дефинальная, т. е. для каждой схемы шифрования $\chi_s\colon V^{(n)} \to W^{(m)}$ существует схема дешифрования (расшифрования): $\chi_s^{-1}\colon W^{(m)} \to V^{(n)}$, так что

$$\chi_s^{-1}(\chi_s(p)) = p \quad \text{для всех} \quad p \in V^{(n)}, \tag{*}$$

но также, что она функциональная и сюръективная, т. е.

$$\chi_s(\chi_s^{-1}(c)) = c \quad \text{для всех} \quad c \in W^{(m)}. \tag{**}$$

Тогда $|V^{(n)}| = |W^{(m)}|$. В этом случае удобно отождествлять символы открытого текста с символами криптотекста. Таким образом, $n = m$, $W \doteq V$ и предполагаем, что имеет место эндоморфизм χ_s: $V^{(n)} \leftrightarrow V^{(n)}$. Итак, пусть M — заданная криптосистема из $V^{(n)} \times V^{(n)}$, $|V| = N$. M является ключевым пространством.

Сделаем теперь важное предположение, что криптосистема M является чистой криптосистемой (разд. 9.1.1), т. е. она замкнута относительно композиции: композиция двух схем шифрования χ_s и χ_t из пространства M снова принадлежит M:

$$\chi_s(\chi_t(p)) = \chi_{s \bullet t}, \quad \text{где } s \bullet t \text{ однозначно определено.}$$

Эта композиция криптосистем из M ассоциативна:

$$\chi_{r \bullet s}(\chi_t(p)) = \chi_r(\chi_{s \bullet t}(p)).$$

Так как мы предположили, что каждая схема шифрования χ из M имеет обратную χ^{-1} из M, то схемы шифрования из M образуют группу относительно композиции, а именно ключевую группу M,

$$\chi_{s^{-1}}(p) \quad \text{определяется как} \quad \chi_s^{-1}(p).$$

Тривиальным образом ключевая группа может состоять из единственного элемента: $M = \{\mathrm{id}\}$ или может иметь N^n элементов, скажем,

$$M = \{\mathrm{id}, \chi, \chi^2, \chi^3, \ldots, \chi^{N^n - 1}\},$$

где χ является одноцикловой подстановкой; или может иметь максимальное возможное число элементов $(N^n)!$, $M \doteq V^n \leftrightarrow V^n$.

Теперь пусть $c' = (c'_1, c'_2, c'_3, \ldots)$ и $c'' = (c''_1, c''_2, c''_3, \ldots)$ — два криптотекста, которые являются зашифрованными одним и тем же ключом $k = (k_1, k_2, k_3, \ldots)$ открытыми текстами $p' = (p'_1, p'_2, p'_3, \ldots)$ и $p'' = (p''_1, p''_2, p''_3, \ldots)$:

$$c'_i = \chi_{k_i}(p'_i), \qquad c''_i = \chi_{k_i}(p''_i).$$

Далее мы предположим, что криптосистема транзитивна (разд. 14.3.4). Тогда ключевая группа M является транзитивной группой перестановок в классическом смысле, т. е. существует некоторый символ a из V^n такой, что для каждого символа y из $V^{(n)}$ существует схема шифрования χ_t из M такая, что $y = \chi_t(a)$. Отсюда получаем, что число ключей не меньше, чем мощность алфавита V^n, $|M| \geqslant N^n$, и каждый символ из $V^{(n)}$ может быть инъективно поставлен в соответствие некоторому ключу. Другими словами, символы являются классами эквивалентности.

Мы можем также предположить систему Шеннона (разд. 2.6.4), где ключ k_i однозначно определен парой, состоящей из символа открытого текста p_i и символа криптотекста c_i, что влечет $|M| \leqslant N^n$. В общем случае ключ k_i не обязательно однозначно определен символами p_i и c_i[2)].

[2)] Лишь в случае $(N^n)! = N^n$, т. е. для $N^n = 1$ или $N^n = 2$ ключ обязательно определен парой p_i, c_i. Это включает как интересный случай лишь $V \doteq Z_2$, $n = 1$; тогда существует лишь две схемы шифрования: тождественная O и инволютивная L (разд. 8.3.1); $p_i = c_i$ дает $k_i \doteq O$, $p_i \neq c_i$ дает $k_i \doteq L$.

Таким образом, мы имеем чистую транзитивную криптосистему с $|M| = N^n$, принадлежащую латинскому квадрату. Соотношение между символами и ключами взаимно однозначно; идентификация ключей и символов, соответствующая равенству $s = \chi_s(a)$, дает $s \bullet t = \chi_{s\bullet t}(a) = \chi_s(\chi_t(a)) = \chi_s(t)$ и, таким образом, $\chi_s(p) = s \bullet p$,

$$\chi_{s\bullet t}(p) = \chi_{\chi_s(t)}(p).$$

Более того,

$$\chi_{s^{-1}}(c) = s^{-1} \bullet c = \chi_s^{-1}(c).$$

Теперь имеет смысл говорить о $\chi_{c_i'}^{-1}(c_i'')$, т. е. о символе криптотекста c_i'', расшифрованным символом криптотекста c_i' как ключом в фазе. Простой расчет показывает, что при таких условиях ключ просто аннулируется, или, более точно,

$$\chi_{c_i'}^{-1}(c_i'') = \chi_{p_i'}^{-1}(p_i'').$$

19.2.2. Разности. Для (эндоморфной) криптосистемы Шеннона с транзитивной ключевой группой мы образуем разность $d_i \stackrel{\text{def}}{=} \chi_{c_i'}^{-1}(c_i'')$ двух наблюдаемых криптотекстов и найдем два открытых текста p_i' и p_i'' таких, что их разность $\chi_{p_i'}^{-1}(p_i'')$ равна d_i. Это можно попытаться сделать способом зигзага, очень похожим на указанный в разд. 14.4; чтобы решение было однозначным, два открытых текста должны быть такими, чтобы сумма их избыточностей (разд. 12.6, примечание 4) была не менее 100%.

Если схемы шифрования образуют к тому же коммутативную группу относительно композиции, то имеет место *симметрия позиции* Керкхоффса:

$$\chi_s(t) = \chi_t(s).$$

Для (изоморфной) криптосистемы с коммутативной ключевой группой ключ k однозначно определяется как парой p' и c', так и парой p'' и c'', так как при $c_i' = \chi_{k_i}(p_i')$ также $c_i' = \chi_{p_i'}(k_i)$ и таким образом $k_i = \chi_{p_i'}^{-1}(c_i')$. Такие криптосистемы обязательно являются шенноновскими; если выполняется неравенство $|M| > N^n$, то ключевая группа не коммутативна. С такими группами мы встретимся в разд. 19.2.4.

Кроме того, в коммутативном случае ввиду выполнения равенства $\chi_{p_i'}(d_i) = p_i''$ выполняется также равенство $\chi_{d_i}(p_i') = p_i''$. Таким образом, p_i'' получается из p_i' шифрованием с d_i в качестве ключа.

Для каждой ключевой группы существует множество *двойственных схем шифрования* $\{\check{\chi}_s\}$, для которых

$$\check{\chi}_s(p) = s \bullet p^{-1}, \quad \check{\chi}_s^{-1}(c) = c^{-1} \bullet s \quad \text{и} \quad \check{\chi}_{s\bullet t^{-1}}(p) = \check{\chi}_{\check{\chi}_s(t)}(p).$$

Теперь получаем $\check{\chi}_{c_i'}^{-1}(c_i'') = \check{\chi}_{p_i''}(p_i')$. Для случая коммутативной ключевой группы двойственное шифрование является взаимообратным: $\check{\chi}_s^{-1}(t) = \check{\chi}_s(t)$.

19.2.3. Циклические ключевые группы. Если (для $n=1$) сопутствующие алфавиты построены с помощью циклических сдвигов первичного алфавита с N символами, то число ключей совпадает с числом символов; фактически, ключевая группа является коммутативной и даже циклической группой порядка N. Таким образом, для шифрования ВИЖЕНЕР эту группу можно рассматривать как модель ключевой группы со сложением по модулю N, тогда как шифрование БОФОРТ (когда оно применялось в машине Hagelin M-209) является двойственным к шифрованию ВИЖЕНЕР. Здесь первичные алфавиты известны в любом случае.

Так же может быть рассмотрено и шифрование АЛЬБЕРТИ. Например, в шифровальной таблице из примера Керкхоффса (табл. 23) с порождающим циклом ρ и первичным алфавитом

$$P\colon \begin{array}{l} \mathtt{a\,b\,c\,d\,e\,f\,g\,h\,i\,j\,k\,l\,m\,n\,o\,p\,q\,r\,s\,t\,u\,v\,w\,x\,y\,z} \\ \mathtt{Z\,Y\,T\,V\,W\,D\,F\,G\,H\,J\,K\,M\,Q\,N\,O\,X\,R\,E\,S\,P\,U\,B\,L\,I\,C\,A,} \end{array}$$

мы имеем: $A=\rho^0 P$, $B=\rho^1 P$, $C=\rho^2 P$, $D=\rho^3 P$, ... , $Z=\rho^{25} P$.

Мы можем предположить, что P уже известна, скажем, потому, что известен используемый диск АЛЬБЕРТИ.

Теперь мы снова вернемся к криптотекстам (i) и (ii) из разд. 19.1.1, обозначив (i) и (ii) через (c') и (c''). Образуя из (c') и (c'') модифицированную разность

$$d_i = \chi^{-1}_{P^{-1}c'_i}(P^{-1}c''_i),$$

мы получим, что $d_i \mathrel{\hat{=}} \rho^{\delta_i} P$ тогда и только тогда, когда в заданном перестановочном алфавите P требуется пройти δ_i шагов для того, чтобы перейти от c''_i к c'_i. В результате имеем

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
c''	U	H	W	P	R	B	Q	L	K	I	B	L	W	R	E	J	R	B	K	L	H	I
c'	U	H	Y	B	R	J	I	M	B	C	F	A	M	M	F	J	H	D	M	R	I	Q
d	a	a	x	c	a	o	l	p	l	b	l	d	h	v	p	a	s	k	b	u	p	p
δ	0	0	23	2	0	14	11	15	11	1	11	3	7	21	15	0	18	10	1	20	15	15

Теперь можно получить

$$d_i = \chi^{-1}_{P^{-1}c'_i}(P^{-1}c''_i) = \chi^{-1}_{P^{-1}p'_i}(P^{-1}p''_i);$$

известную последовательность $d=(d_1, d_2, d_3, \ldots)$ можно интерпретировать как криптотекст, полученный при шифровании χ^{-1} из (неизвестного) открытого текста $P^{-1}p''$ с помощью (неизвестного) ключа $P^{-1}p'$. Эта перемена ролей позволяет для определения ключа использовать все возможные атаки с шаблонами и частотами.

Например, равенство $d_i=\mathrm{a}$, которое встречается для $i=1, 2, 5, 16$ означает совпадение p'_i и p''_i. Во французском языке это встречается с частотой около 30% для $p'_i=p''_i=/\mathrm{c}/$, тогда как каждое из $p'_i=p''_i=/\mathrm{a}/$ и $p'_i=p''_i=/\mathrm{r}/$

встречаются с частотой лишь 10%. (Фактически, смелое предположение относительно /e/ должно выполниться для $i = 2, 5, 16$, тогда как для $i = 1$ совпадающей буквой является /l/.) В случае, когда доступна информация о жанре сообщения, может быть рекомендован метод вероятного слова. Таким образом, предполагая, что мы имеем дело с французским языком, и учитывая известные обстоятельства, берем вероятное слово *ennemi* [враг]. Остается проверить (перебором), будет ли соответствовать одной из возможных позиций слова *ennemi* в p'' осмысленное французское слово в p' (или наоборот). Следующие последовательные испытания

	1	2	3	4	5	6	7	8	9
p''	e	n	n	e	m	i	*	*	*
d	a	a	x	c	a	o	l	p	l
δ	0	0	23	2	0	14	11	15	11
p'	e	n	k	g	m	w	*	*	*

	1	2	3	4	5	6	7	8	9
p''	*	e	n	n	e	m	i	*	*
d	a	a	x	c	a	o	l	p	l
δ	0	0	23	2	0	14	11	15	11
p'	*	e	k	p	e	a	t	*	*

	1	2	3	4	5	6	7	8	9
p''	*	*	e	n	n	e	m	i	*
d	a	a	x	c	a	o	l	p	l
δ	0	0	23	2	0	14	11	15	11
p'	*	*	b	p	п	s	x	x	*

	1	2	3	4	5	6	7	8	9
p''	*	*	*	e	п	n	e	m	i
d	a	a	x	c	a	o	l	p	l
δ	0	0	23	2	0	14	11	15	11
p'	*	*	*	g	п	b	p	b	t

не дают успеха, но в 13-й позиции (и нигде больше)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
p''	*	*	*	*	*	*	*	*	*	*	*	*	e	n	п	e	m	i	*	*	*	*
d	a	a	x	c	a	o	l	p	l	b	l	d	h	v	p	a	s	k	b	u	p	p
δ	0	0	23	2	0	14	11	15	11	1	11	3	7	21	15	0	18	10	1	20	15	15
p'	*	*	*	*	*	*	*	*	*	*	*	*	l	i	c	e	e	s	*	*	*	*

появляется фрагмент /licees/, который можно удлинить до /police-est/. Теперь перемена ролей p' и p'' дает

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
p''	*	*	*	*	*	*	*	*	*	*	e	l	e	n	п	e	m	i	s	*	*	*
d	a	a	x	c	a	o	l	p	l	b	l	d	h	v	p	a	s	k	b	u	p	p
δ	0	0	23	2	0	14	11	15	11	1	11	3	7	21	15	0	18	10	1	20	15	15
p'	*	*	*	*	*	*	*	*	*	*	p	o	l	i	c	e	e	s	t	*	*	*

наводящее на удлинение /de l'ennemi sont/. Снова идя в обратную сторону, получаем

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22
p''	* * * * *	* * * * d	e l e n n	e m i s o	n t
d	a a x c a	o l p l b	l d h v p	a s k b u	p p
δ	0 0 23 2 0	14 11 15 11 1	11 3 7 21 15	0 18 10 1 20	15 15
p'	* * * * *	* * * * e	p o l i c	e e s t i	c i

Таким образом, вероятное слово может служить в качестве семени, которое способом зигзага прорастает вправо и влево в обоих текстах. Этот метод позволяет, в частности, использовать бессодержательные слова, окончания и приставки, которые нередки; в английском языке /and/, /the/, /that/, /which/, /under/, /tion/, во французском /les/, /que/, /ion/, в немецком /und/, /ein/, /ung/, /bar/, /heit/, /unter/. В нашем примере новым удачным семенем является /les/:

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22
p''	l e s * *	* * * * d	e l e п п	e m i s o	n t
d	a a x c a	o l p l b	l d h v p	a s k b u	p p
δ	0 0 23 2 0	14 11 15 11 1	11 3 7 21 15	0 18 10 1 20	15 15
p'	l e p * *	* * * * e	p o l i c	e e s t i	c i

и мы с маленьким кусочком получаем счастливую возможность /leprefetd/ в p' и подтверждение дополнительным /lespertes/ в p'':

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22
p''	l e s p e	r t e s d	e l e n n	e m i s o	n t
d	a a x c a	o l p l b	l d h v p	a s k b u	p p
δ	0 0 23 2 0	14 11 15 11 1	11 3 7 21 15	0 18 10 1 20	15 15
p'	l e p r e	f e t d e	p o l i c	e e s t i	c i

Это заканчивает дешифрование рассматриваемых текстов. Правда, таким способом можно получить дешифрование только более короткого из двух текстов, но есть еще ключ, который мы до сих пор не использовали. Теперь он может быть восстановлен: наложение p' и c' дает, согласно табл. 23 (которую мы предполагали известной):

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22
c'	U H Y B R	J I M B C	F A M M F	J H D M R	I Q
p'	l e p r e	f e t d e	p o l i c	e e s t i	c i
k	*J E M E M*	*E T S S U*	*R L A D E*	*F E N S I*	*V E*

и новое подтверждение «осмысленным» ключевым предложением. Для успеха этого зигзагового метода было достаточно, чтобы каждый из двух открытых текстов имел более чем 50-процентную избыточность.

19.2.4. Другие ключевые группы. Ключевая группа, с которой мы только что имели дело, типична для шифра АЛЬБЕРТИ и особенно для шифра ВИЖЕНЕР (двойственно: БОФОРТ). Как было сказано выше, она является циклической группой C_N порядка N, где $V = W = \mathbb{Z}_N$. Это только один пример группы данного порядка. Для $\mathbb{Z}_{26}$ кроме циклической группы C_{26} существует еще одна коммутативная группа: прямое произведение $C_{13} \times C_2$ циклической группы порядка 13 и циклической группы порядка 2. Это группа, порожденная 13 шифрами ПОРТА, которые получаются промежуточным кодированием $\mathbb{Z}_{26} \to \mathbb{Z}_{13} \times \mathbb{Z}_2$. Существует также некоммутативная группа $\mathcal{D}_{13}$ прядка 26 с двумя образующими S и T; $S^{13} = T^2 = (ST)^2 = I$, но до сих пор она не имела отношения к криптографии.

Для $\mathbb{Z}_{25}$ кроме циклической группы C_{25} существует еще одна коммутативная группа — прямое произведение $C_5^2 \doteq C_5 \times C_5$ двух циклических групп 5-го порядка, получаемая промежуточным кодированием Полибия $\mathbb{Z}_{25} \to \mathbb{Z}_5 \times \mathbb{Z}_5$. Других групп 25-го порядка (даже некоммутативных) не существует. Для $\mathbb{Z}_{10}$ кроме циклической группы C_{10} существует коммутативная группа $C_5 \times C_2$ — прямое произведение циклических групп пятого и второго порядков. Она получается промежуточным биквинарным кодированием $\mathbb{Z}_{10} \to \mathbb{Z}_5 \times \mathbb{Z}_2$. Существует также некоммутативная группа $\mathcal{D}_5$ с образующими элементами S и T, $S^5 = T^2 = (ST)^2 = I$.

Ввиду популярности бинарного кодирования особенно интересны группы порядка 2^n. Для произвольного j наиболее известны группы C_{2^j} и $C_2^j = C_2 \times C_2 \times \ldots \times C_2$. Для $j = 2$ имеем циклическую группу 4-го порядка и группу Клейна: *Vierergruppe* [четверная группа (*нем.*)]. Для $j = 3$ существуют некоммутативные группы Q кватернионов и $\mathcal{D}_4$ с образующими S и T, $S^4 = T^2 = (ST)^2 = I$, обе не имеющие отношения к криптологии. Это замечание относится также к некоммутативным группам для $j = 4$ и $j = 5$.

О различии между $\mathbb{Z}_{2^n}$ и $\mathbb{Z}_2^n$, так же, как и о различии между $\mathbb{Z}_{10^n}$ и $\mathbb{Z}_{10}^n$, а именно об отсутствии механизма переноса, см. в разд. 8.3.3.

19.2.5. Частный случай C_2^5. Представление $\mathbb{Z}_{2^5}$ является словарем Международного телетайпного алфавита № 2 (CCITT 2). С 1929 г. произошел возврат

0	t	4	o	2	h	n	m	5	l	r	g	i	p	c	v	e	z	d	b	s	y	f	x	a	w	j	3	u	q	k	1	
O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	L	L	L	L	L	L	L	L	L	L	L	L	L	L	16
O	O	O	O	O	O	O	O	L	L	L	L	L	L	L	O	O	O	O	O	O	O	O	L	L	L	L	L	L	L	L	L	8
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
O	O	O	O	L	L	L	L	O	O	O	O	L	L	L	L	O	O	O	O	L	L	L	L	O	O	O	O	L	L	L	L	4
O	O	L	L	O	O	L	L	O	O	L	L	O	O	L	L	O	O	L	L	O	O	L	L	O	O	L	L	O	O	L	L	2
O	L	O	L	O	L	O	L	O	L	O	L	O	L	O	L	O	L	O	L	O	L	O	L	O	L	O	L	O	L	O	L	1
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	

Таблица 24. Бинарное кодирование Международного телетайпного алфавита № 2 (CCITT 2). 0: пусто, 1: верхний регистр буквы, 2: пробел, 3: верхний регистр цифры, 4: возврат каретки, 5: загрузка линии.

к алфавиту Мюррея, существовавшему до 1900 года. Пятиканальное представление предлагает шифрование, ключевой группой которого является C_2^5 (а не C_{2^5}), а именно шифр с 32 алфавитами, порожденный подстановками O или L (разд. 8.3.1) пяти бинарных символов. Действительное кодирование $\mathbb{Z}_{32} \to \mathbb{Z}_2^5$ CCITT 2 показано в табл. 24. Кроме 26 букв (нижний регистр), имеются еще шесть контрольных символов телетайпной машины, чьи функции не имеют отношения к криптографической защите сообщений, мы обозначим их 0, 1, 2, 3, 4, 5, и используем 2 как разделитель слов (в действительности бы-

	0	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	2	3	4	5	1
0	0	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	2	3	4	5	1
A	A	0	G	F	R	5	C	B	Q	S	4	N	Z	1	K	3	Y	H	D	I	W	2	X	T	V	P	L	U	O	J	E	M
B	B	G	0	Q	T	O	H	A	F	1	L	P	J	S	Y	E	K	C	W	M	D	V	U	R	2	N	4	X	5	Z	3	I
C	C	F	Q	0	U	K	A	H	G	4	S	E	M	L	5	P	O	B	2	J	V	D	T	X	W	3	1	R	Y	I	N	Z
D	D	R	T	U	0	4	2	W	X	K	5	I	3	Y	S	Z	1	V	A	N	B	C	Q	G	H	M	O	F	L	E	J	P
E	E	5	O	K	4	0	N	3	Y	U	R	C	W	X	F	B	Q	P	J	2	Z	I	1	L	M	H	T	S	G	D	A	V
F	F	C	H	A	2	N	0	Q	B	J	I	5	1	Z	E	Y	3	G	U	4	X	R	W	V	T	O	M	D	P	S	K	L
G	G	B	A	H	W	3	Q	0	C	M	Z	Y	4	I	P	5	N	F	T	1	R	X	2	D	U	K	J	V	E	L	O	S
H	H	Q	F	G	X	Y	B	C	0	L	1	3	I	4	O	N	5	A	V	Z	2	W	R	U	D	E	S	T	K	M	P	J
I	I	S	1	4	K	U	J	M	L	0	F	D	H	G	R	V	T	Z	N	A	P	E	O	Y	3	W	Q	5	X	C	2	B
J	J	4	L	S	5	R	I	Z	1	F	0	2	B	Q	U	W	X	M	E	C	3	N	Y	O	P	V	G	K	T	A	D	H
K	K	N	P	E	I	C	5	Y	3	D	2	0	X	W	A	Q	B	O	S	R	1	4	Z	M	L	G	V	J	H	U	F	T
L	L	Z	J	M	3	W	1	4	I	H	B	X	0	C	V	R	2	S	O	Q	5	Y	N	E	K	U	A	P	D	G	T	F
M	M	1	S	L	Y	X	Z	I	4	G	Q	W	C	0	T	2	R	J	P	B	N	3	5	K	E	D	F	O	U	H	V	A
N	N	K	Y	5	S	F	E	P	O	R	U	A	V	T	0	H	G	3	I	D	M	J	L	1	Z	B	X	4	Q	2	C	W
O	O	3	E	P	Z	B	Y	5	N	V	W	Q	R	2	H	0	C	K	L	X	4	1	I	J	S	F	D	M	A	T	G	U
P	P	Y	K	O	1	Q	3	N	5	T	X	B	2	R	G	C	0	E	M	W	I	Z	4	S	J	A	U	L	F	V	H	D
Q	Q	H	C	B	V	P	G	F	A	Z	M	O	S	J	3	K	E	0	X	L	U	T	D	2	R	5	I	W	N	1	Y	4
R	R	D	W	2	A	J	U	T	V	N	E	S	O	P	I	L	M	X	0	K	G	F	H	B	Q	1	3	C	Z	5	4	Y
S	S	I	M	J	N	2	4	1	Z	A	C	R	Q	B	D	X	W	L	K	0	Y	5	3	P	O	T	H	E	V	F	U	G
T	T	W	D	V	B	Z	X	R	2	P	3	1	5	N	M	4	I	U	G	Y	0	Q	C	A	F	S	E	H	J	O	L	K
U	U	2	V	D	C	I	R	X	W	E	N	4	Y	3	J	1	Z	T	F	5	Q	0	B	H	G	L	P	A	M	K	S	O
V	V	X	U	T	Q	1	W	2	R	O	Y	Z	N	5	L	I	4	D	H	3	C	B	0	F	A	J	K	G	S	P	M	E
W	W	T	R	X	G	L	V	D	U	Y	O	M	E	K	1	J	S	2	B	P	A	H	F	0	C	I	5	Q	4	3	Z	N
X	X	V	2	W	H	M	T	U	D	3	P	L	K	E	Z	S	J	R	Q	O	F	G	A	C	0	4	N	B	I	Y	1	5
Y	Y	P	N	3	M	H	O	K	E	W	V	G	U	D	B	F	A	5	1	T	S	L	J	I	4	0	2	Z	C	X	Q	R
Z	Z	L	4	1	O	T	M	J	S	Q	G	V	A	F	X	D	U	I	3	H	E	P	K	5	N	2	0	Y	R	B	W	C
2	2	U	X	R	F	S	D	V	T	5	K	J	P	O	4	M	L	W	C	E	H	A	G	Q	B	Z	Y	0	1	N	I	3
3	3	O	5	Y	L	G	P	E	K	X	T	H	D	U	Q	A	F	N	Z	V	J	M	S	4	I	C	R	1	0	W	B	2
4	4	J	Z	I	E	D	S	L	M	C	A	U	G	H	2	T	V	1	5	F	O	K	P	3	Y	X	B	N	W	0	R	Q
5	5	E	3	N	J	A	K	O	P	2	D	F	T	V	C	G	H	Y	4	U	L	S	M	Z	1	Q	W	I	B	R	0	X
1	1	M	I	Z	P	V	L	S	J	B	H	T	F	A	W	U	D	4	Y	G	K	O	E	N	5	R	C	3	2	Q	X	0

Таблица 25. Таблица шифрования (латинский квадрат) для телетайпных символов (сложение по модулю 2 в $\mathbb{Z}_2^5$)

ло использовано 12; это было слабостью разработки Бьюрлинга). Обозначая ключи соответственно 0, A, B, C, ..., Z, 2, 3, 4, 5, 1, приведем естественную таблицу шифрования, которую можно предположить известной.

Телетайпное кодирование было широко известно с начала XX в., и через Вернама с ним познакомились профессиональные криптологи; очевидно, естественная ключевая группа $\mathbb{Z}_2^5$ была хорошо известна. Таким образом, все предварительные условия для атаки, как в разд. 19.2.3, были выполнены, и в частности, мог быть восстановлен ключевой символ (ввиду коммутативности ключевой группы) из символа открытого текста и символа криптотекста.

Вымышленный пример взлома можно построить так: два криптотекста, состоящие примерно из 4000 символов, перехвачены англичанами во время немецкой атаки на Крите в середине мая 1941 г. Эти тексты содержат после совпадающих преамбул, вероятно, в фазе, следующие фрагменты:

c''	2WHNR	G1ATU	APLBV	RWOUF	YPBSX	ZNR4J	SR
c'	L0G2A	WGH2Z	KBVZV	QZWYK	YWJI0	KT5AZ	2K

Англичане образовали разность d (т. е. произвели сложение текстов по mod 2).

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22 23 24 25	26 27 28 29 30	31 32
c''	2WHNR	G1ATU	APLBV	RWOUF	YPBSX	ZNR4J	SR
c'	L0G2A	WGH2Z	KBVZV	QZWYK	YWJI0	KT5AZ	2K
d	fwcwd	dvqkp	nkn40	x5jl5	0slax	vm4jg	gs

и использовали вероятное слово /2kreta2/, чтобы найти осмысленное дополнение (в другом тексте). Оно нашлось в четвертой позиции.

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22 23 24 25	26 27 28 29 30	31 32
p''	***2k	reta2	*****	*****	*****	*****	**
d	fwcwd	dvqkp	nkn40	x5jl5	0slax	vm4jg	gs
p'	***ni	a2und	*****	*****	*****	*****	**

Краткий взгляд на карту Греции предлагает найти в p'' дополнение к /chania/ и дает:

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22 23 24 25	26 27 28 29 30	31 32
p''	auf2k	reta2	*****	*****	*****	*****	**
d	fwcwd	dvqkp	nkn40	x5jl5	0slax	vm4jg	gs
p'	chani	a2und	*****	*****	*****	*****	**

Теперь можно попытаться найти еще несколько географических названий, следующих за /und/. Другой путь — взять еще одно вероятное слово, скажем, /2angriff2/. Люди из Блетчли Парк добились успеха в позиции 19 для p'':

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22 23 24 25	26 27 28 29 30	31 32
p''	auf2k	reta2	*****	***2a	ngrif	f2***	**
d	fwcwd	dvqkp	nkn40	x5jl5	0slax	vm4jg	gs
p'	chani	a2und	*****	***fe	n2ost	wae**	**

Теперь все почти закончено: пропущенный кусок из p' должен читаться /2die-2haefen/, следующий за /ostwaerts2/. Это дает

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22 23 24 25	26 27 28 29 30	31 32
p''	a u f 2 k	r e t a 2	w i r d 2	d e r 2 a	n g r i f	f 2 d e r	2 g
d	**f w c w d**	**d v q k p**	**n k n 4 0**	**x 5 j l 5**	**0 s l a x**	**v m 4 j g**	**g s**
p'	c h a n i	a 2 u n d	2 d i e 2	h a e f e	n 2 o s t	w a e r t	s 2

Таким образом, расшифрованные тексты читаются так:

«auf kreta wird der angriff der g...»
«chania und die haefen ostwaerts...»

Криптоаналитики могли продолжать аналогично работать и с другими фрагментами текста. Англичане могли также восстановить и ключ, но так как в C_2^n вычитание совпадает со сложением, то для этого нужно найти лишь первое правильное согласование открытого текста с криптотекстом. Два возможных наложения дают в результате

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22 23 24 25	26 27 28 29 30	31 32
p'	c h a n i	a 2 u n d	2 d i e 2	h a e f e	n 2 o s t	w a e r t	s 2
k_1	M H B W S	T S W W O	T T O T E	A L L O C	B N W A T	M W A D E	G T
c'	L 0 G 2 A	W G H 2 Z	K B V Z V	Q Z W Y K	Y W J I 0	K T 5 A Z	2 K

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22 23 24 25	26 27 28 29 30	31 32
p''	a u f 2 k	r e t a 2	w i r d 2	d e r 2 a	n g r i f	f 2 d e r	2 g
k_2	Z U Q 0 N	B 3 6 M C	M 2 H O E	V T B R N	B D E 0 F	5 K J 5 3	0 Y
c'	L 0 G 2 A	W G H 2 Z	K B V Z V	Q Z W Y K	Y W J I 0	K T 5 A Z	2 K

Далее не трудно решить вопрос о том, какой ключевой текст менее нерегулярен (вроде k_1), предположив, что машина имеет довольно несложный механизм определения локальной нерегулярности.

19.2.6. Тунец. Используемые на высшем уровне немецкого Вермахта шифровальные телетайпные машины SZ 40 и SZ 42 были построены компанией Лоренца, которая порождала свои ключи полурегулярным способом, и поэтому было крайне рискованно, что может возникнуть OT-OT компромисс. К тому же Филби сообщил, что какой-то «М-р Лоренц» в 1932 г. предложил английскому министерству иностранных дел шифромашину и раскрыл ее. Вдобавок конкурирующее изделие, шифровальная телетайпная машина T 52, построенная компанией Сименс, была открыто описана в немецком Патенте № 615016 Иппом и Росбергом (разд. 9.1.3); соответствующий американский Патент № 1912983 Иппа, Росберга и Хеттлера был выдан 6 июня 1933 г. Для джентльменов из Блетчли Парк было не слишком трудно разобраться в этой ситуации реалистически. К счастью для англичан, аппарат SZ (английское кодовое название «тунец») использовал лишь 32 схемы шифрования, порождаемые подстановками O и L длины пять, тогда как T 52, как это объяснено в патенте, использовал также перестановки битов (разд. 9.1.3), и таким образом, имел ключевое множество с намного более чем 32 ключами.

Несчастье для немцев в действительности проявилось намного раньше, чем его можно было ожидать. Как сообщил в 1993 г. Гуд (а первые упоминания были сделаны в 1978 г. Джонсоном и в 1983 г. Ходжесом), первый OT-OT компромисс случился даже раньше, чем машина SZ начала регулярно использоваться: во время тестирования вновь устанавливаемой Вермахтом радиолинии между Веной и Афинами, из-за ошибки немецкого телеграфиста, два сообщения p' и p'' были посланы с одной и той же позиции ключевых колес, т. е. были зашифрованы в фазе одним и тем же ключом. Кто-то в Блетчли Парк обратил на это внимание; обнаруженный компромисс позволил полковнику (позднее бригадиру) Тилтмену, тогда главному криптоаналитику Блетчли Парк, после двухнедельной кропотливой работы осенью 1941 г. получить два открытых текста из разности двух записанных криптотекстов «глубины два».

Теперь известно, что этот случай произошел 30 августа 1941 г.: с одним и тем же индикатором HQIBPEXEZMUG были записаны два текста примерно из 4000 символов, совпадающие в первых семи символах; символы с 51-го по 120-й процитированы ниже вместе с их разностями, вычисленными Тилтменом:

	51 52 53 54 55	56 57 58 59 60	61 62 63 64 65	66 67 68 69 70	71 72 73 74 75	76 77 78 79 80	81 82 83 84 85
c''	UB23R	5WEVG	QI245	GRJML	CY50H	KAS1I	S5XUN
c'	YUHVH	3HEE0	TG2HH	1QJXV	K1BJM	K2OMZ	YVIN3
d	lvtsv	bu01g	um0mp	sx0en	er3j4	0uxaq	tm3jq

	86 87 88 89 90	91 92 93 94 95	96 97 98 99 00	01 02 03 04 05	06 07 08 09 10	11 12 13 14 15	16 17 18 19 20
c''	SRZZB	DBB1C	LSQHH	UH5XD	0FN3J	3VOCA	DJCDN
c'	HMC3D	UQ34Z	R2MRM	OH*JQ	PWUEY	CDRG1	LDATI
d	zp1rt	cc5q1	oejv4	10*pv	pvjgv	yqlhm	35fbr

Если бы Тилтмен испытывал в качестве вероятного слова очень частое /geheim2/, он бы уже дважды нашел понятное дополнение,

	61 62 63 64 65	66 67 68 69 70	71 72 73 74 75	76 77 78 79 80	81 82 83 84 85	86 87 88 89 90	91 92 93 94 95
p''	****g	eheim	2****	*****	**geh	eim2*	*****
d	um0mp	sx0en	erej4	0uxaq	tm3jq	zp1rt	cc5q1
p'	****n	2deut	s****	*****	**era	ttac*	*****

/n2deuts/, которое легко дополняется до /an2deutsch/ и /erratac/, приводящее к /2militaerattache2/; брешь заполняется с помощью /an2deutschen/. Таким образом, получается фрагмент открытого текста p' из 29 символов,

	61 62 63 64 65	66 67 68 69 70	71 72 73 74 75	76 77 78 79 80	81 82 83 84 85	86 87 88 89 90	91 92 93 94 95
p''	****g	eheim	2****	*****	**geh	eim2*	*****
d	um0mp	sx0en	erej4	0uxaq	tm3jq	zp1rt	cc5q1
p'	***an	2deut	schen	2mili	taera	ttach	e2***

который дает 29 символов для открытого текста p'':

	6162636465	6667686970	7172737475	7677787980	8182838485	8687888990	9192939495
p''	***lg	eheim	22kr2	233zz	01geh	eim22	kr***
d	um0mp	sx0en	erej4	0uxaq	tm3jq	zp1rt	cc5q1
p'	***an	2deut	schen	2mili	taera	ttach	e2***

Таким образом, немцы совершили традиционную ошибку, продублировали /geheim/; дублирование полной группы /1geheim2233zz0/ приводит к расширению, которое с учетом двух орфографических ошибок становится осмысленным. Теперь возникает подозрение, что остаток сообщения p'' просто сдвинут относительно p', и притом на 39 позиций, так как /an2deutschen2militaerattache2/ в 103-й позиции снова дает осмысленную последовательность, а именно /lage11nr33mwoou211g/ (т. е. *lage nr. 2997-g*):

	8687888990	9192939495	9697989900	0102030405	0607080910	1112131415	1617181920
p''	eim22	kr233	3z112	**an2	deuts	chen2	milit
d	zp1rt	cc5q1	oejv4	100pv	pvjgv	y4lhm	35fbr
p'	ttach	e2iw2	atgen	*lage	11nr3	3mwoo	u211g

Таким образом, дальнейшее дешифрование можно осуществлять автоматически, просматривая вперед 39 символов — совершенно аналогично ситуации с автоключом, описанной Шенноном (разд. 8.7.2). Сообщалось, что Тилтмен закончил дешифрование за 10 дней, что могло быть вызвано ошибками в прослушивании, более серьезными, чем показал ретроспективный анализ.

Тот факт, что сложение производилось по модулю 2 (и, следовательно, совпадало с вычитанием), для метода был безразличен, и лишь упрощал рутинную работу. Правда, шифрование было взаимообратным, (фактор, благоприятствующий дешифрованию), но, все-таки не собственным взаимообратным.

Эти два сообщения, скорее всего, были небольшими. Важнее всего было то, что выявленный фрагмент ключа примерно из 4000 символов к тому времени неизвестной машины (которой англичане дали кодовое наименование tunny (тунец)) получался вычислением $k = c' + p' \pmod 2$; он начинался с отрезка

	61 62 63 64 65	66 67 68 69 70	71 72 73 74 75	76 77 78 79 80	81 82 83 84 85	86 87 88 89 90	
k	***QO	3VRGC	RZFRT	JOVCQ	SXUIO	2NFYX	
	LO	LOOOO	OLLOO	LOOOL	LLLOO	OOLLL	16
	LO	LLLLL	LOOLO	LOLLL	OOLLO	OOOOO	8
	LO	OLOOL	OOLOO	OOLLL	LLLLO	LLLLL	4
	OL	LLLLL	LOLLO	LLLLO	OLOOL	OLLOL	2
	LL	LLOLO	OLOOL	OLLOL	OLOOL	OOOLL	1

	91 92 93 94 95	96 97 98 99 00	01 02 03 04 05	06 07 08 09 10	11 12 13 14 15	16 17 18 19 20	
k	IWX2Y	DH4JT	MIEZP	DNJIC	YRB5U	YFMKM	
	OLLOL	LOOLO	OOLLO	LOLLO	LOLOL	LLOLO	16
	LLOOO	OOOLO	OLOOL	OOLLL	OLOLL	OOOLO	8
	LOLLL	OLOOO	LLOOL	OLOOL	LOOOL	LLLLL	4
	OOLOO	LOLLO	LOOOO	LLLLL	OLLOO	OLLLL	2
	OLLOL	LLOOL	LOOLL	OOOOO	LOLOO	LOLOL	1

С помощью этого восстановленного ключевого фрагмента гипотетической машины *tunny* можно было проанализировать ключевой генератор этой немецкой шифромашины. Прежде всего, следовало найти периоды отдельных ключевых дисков (наличие которых можно было угадать по аналогии с сименсовской телетайпной шифровальной машиной T 52). С помощью индикатора можно было смутно догадаться, что там было 12 ключевых дисков. Поскольку ни один из каналов от k_1 до k_5 ключа k не имел длину периода меньше 100, можно было предположить, что каждый канал был зашифрован композицией (по крайней мере) двух ключевых дисков. Англичане назвали их χ_i и ψ_i, где первыми применялись *Хи*-диски, а вторыми — *Пси*-диски. Исследованиями периодичности сначала были установлены периоды ключевых *Хи*-дисков, в частности, 41 для χ_1, как показано на рис. 142. Затем были определены периоды ключевых *Пси*-дисков (43 для ψ_1) и выяснен способ работы двух оставшихся движущихся дисков (разд. 9.1.4).

Это было сделано в основном юным математиком Таттом из Тринити Колледжа в Кембридже, который позднее стал хорошо известным специалистом по теории графов. К февралю 1942 г. была полностью выяснена структура машины, использованной для передачи сообщений HQIBPEXEZMUG (впоследствии эту машину назвали ZMUG); после окончания войны правильность реконструкции была подтверждена трофейными машинами. Криптоаналитик

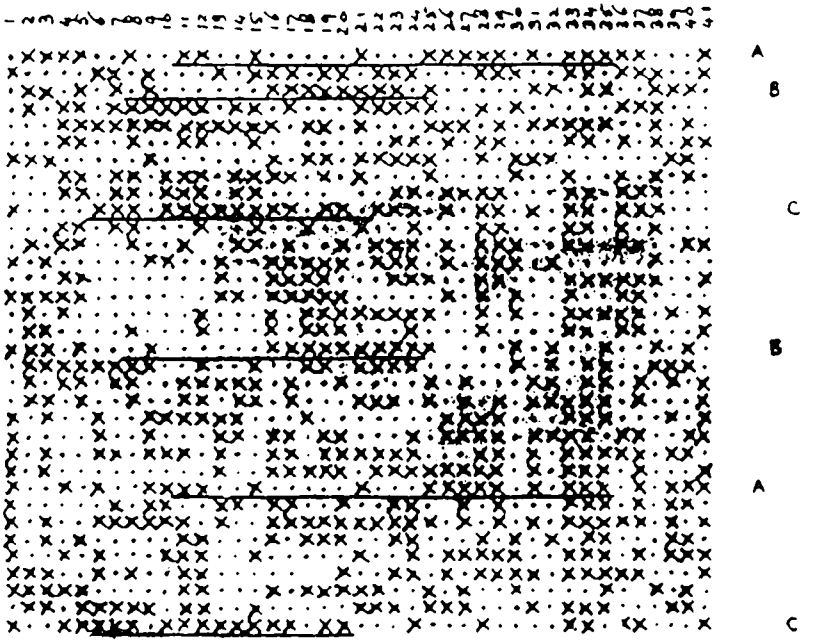


Рис. 142. Исследование периода ключевого диска χ_1. Повторения Касиски A, B и C. В Блетчли Парк точка представляла O, а крестик — L

Гуд мог с гордостью заявить: «Мы не располагали немецкими tunny до последних дней войны в Европе». Действительно, это был чистый криптоанализ. Первая задача состояла в нахождении дневной ключевой последовательности — периодических **O–L** последовательностей, порождаемых выступами на ключевых дисках («взлом колеса»). Остающаяся практическая работа состояла в нахождении начальной установки ключевых дисков («установка дисков для каждого сообщения»). Вероятно, эта задача решалась при помощи протягивания вдоль криптотекста вероятных слов, — и когда это приводило к успеху, макет *tunny*, построенный Бродхестом, печатал расшифрованный открытый текст.

Эта работа направлялась Ньюменом, который был убежденным сторонником чистого криптоанализа. Был ключевой текст и наблюдаемый криптотекст, и эти тексты должны были быть согласованы (для такого приведения предположительно применялся тест *Kanna*). Эта проблема могла быть решена функционально с помощью принципа «пилки дров» (разд. 17.3.2). В мае 1943 г. начала действовать первая модель быстрой машины HEATH ROBINSON (которая обсуждалась в разд. 17.3.3). Ее две петли, одна для криптотекста, другая для ключевого текста были перфорированы на лентах. Синхронизация этих лент требовала значительных усилий. Поэтому Флауэрс, поддержанный Бродхестом, Чандлером и Кумбсом, разработал усовершенствованную версию, где ключевая лента была заменена внутренним устройством, а вторая

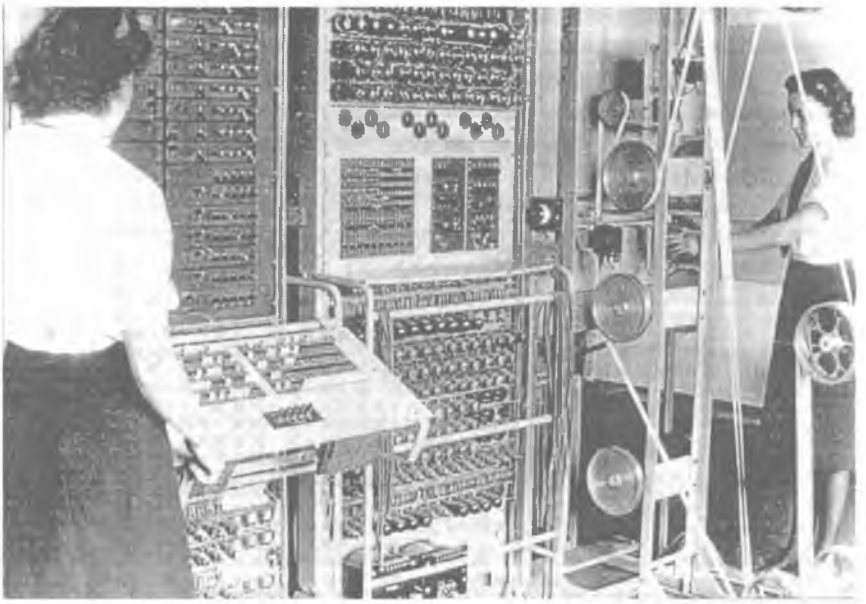


Рис. 143. Частичный вид машины COLOSSUS (предположительно II). Ясно видны замкнутая ленточная петля для криптотекста, штепсельный коммутатор и массив ламп, вероятно, типа Mullard EF 36

лента прочитывалась фотоэлектрически со скоростью 5000 символов в секунду. Важнее всего было то, что в новой машине применялись электронные переключательные схемы и, соответственно, она была быстрой. В декабре 1943 г. была готова модель, названная COLOSSUS-I. В феврале 1944 г. она была принята на вооружение и использовалась против *tunny*. Это был первый функционирующий электронный компьютер в мире. А 1 июня 1944 г., как раз в день высадки союзников в Нормандии, начал использоваться усовершенствованный COLOSSUS-II (рис. 143).

В нем внутренние электронные схемы примерно из 1500 ламп работали с кольцевыми счетчиками в n-значном коде[3]. COLOSSUS благодаря штепсельному коммутатору позволял использовать гибкое программирование элементарных булевых операций и бинарную арифметику с 5-кратной параллелизацией. Усовершенствованный COLOSSUS-II с примерно 2000 лампами был способен также выполнять условные ветвления, он имел «панель логического переключения» для заранее заданных, а также изменяемых вручную, булевых операций. Всего было построено 10 машин COLOSSUS.

Не следует считать очернением англичан утверждение, что машины COLOSSUS были в основном ориентированы на использование примитивных операций сравнения, на вычисление лишь в очень специальном смысле и что их управление было на том же уровне, что и у машин Цузе. Применялись ли они для других целей, кроме определения «установки дисков» — остается неизвестным.

Согласно Бурке, в США успешной криптоаналитической электронной разработки, сравнимой с английским COLOSSUS, не существовало до конца Второй мировой войны из-за неудач в попытках Буша построить электронный COMPARATOR. Но «ко времени японской капитуляции американцы построили электронные машины, использовав вдвое больше ламп, чем в английском COLOSSUS» (Бурке).

Англичане после 1943 г. добавили к своим успехам взлом немецкой радиотелетайпной линии связи, использовавшей в основном машины SZ 40 и SZ 42, и зашифрованный машиной ENIGMA поток сообщений, который был, главным образом, предназначен для немецкого Люфтваффе. Несмотря на то, что их дешифрование ввиду высшего криптологического уровня иногда требовало большого времени (обычно 4 дня), получаемая стратегическая информация стоила этих усилий. Эти взломы линий связи включали с ноября 1942 г. линию между Берлином и армейской группой E в Салониках, а с января 1943 г. линию между армейской группой C в Риме и танковой армией фельдмаршала Эрвина Роммеля в Тунисе. С мая 1943 г. дала течь линия между Берлинским центром в Штраусберге и армейской группой C (фельдмаршал Альберт Кесселринг), а с марта 1943 г. — также линия между Берлином и армейской группой Юг в Виннице. Как одно из последствий этого, немецкая атака про-

[3] Джонсон был дезинформирован в 1978 г., предполагая, что COLOSSUS был спроектирован для работы против шифромашины Сименса T 52 (tunny), и поэтому упомянул лишь 10 кольцевых счетчиков вместо 12, соответствующих 12 ключевым дискам шифромашины SZ 40.

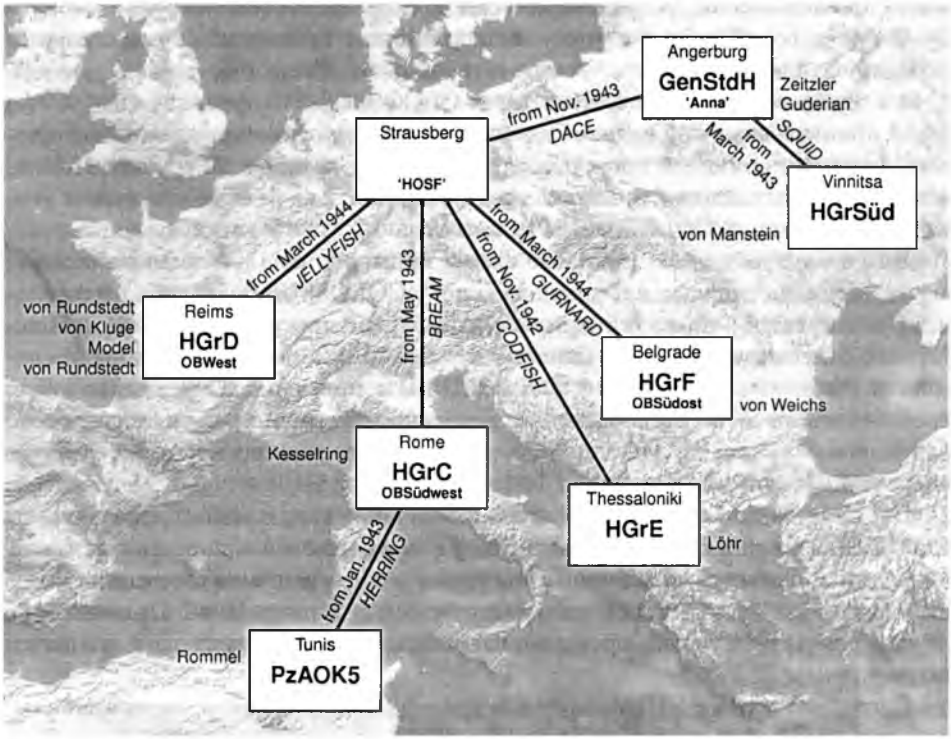


Рис. 144. Некоторые беспроволочные линии телетайпной связи, оснащенные SZ 40 и SZ 42, взломанные машиной COLOSSUS Блетчли Парк (Хинсли)

тив Курска в июле 1943 г. превратилась в поражение[4]. Эти взломы были в 1944 г. дополнены новыми, одной из побед был взлом линии между Берлином и командующим армейской группой Запад фельдмаршалом Гердом фон Рунштедтом (рис. 144). А немцы ни о чем не подозревали. Они каждые 3 месяца изменяли *Хи*-диски, каждые 6 месяцев — *Пси*-диски (позднее — каждый месяц), а движущиеся диски меняли ежедневно.

Но англичанам тоже не повезло. 10 июня 1944 г., через четыре дня после дня D[5] и спустя 10 дней после пуска COLOSSUS MARK II они потеряли свой вход в линию из Берлина к фон Рундштедту, а в июле также в линию из Берлина к Кессельрингу; и только в сентябре 1944 г. ситуация исправилась. Причиной этого стало радикальное дополнение к машине Лоренца, подобное *функции открытого текста* на машине Сименса.

[4] Советское командование имело собственные источники информации о планах немцев под Курском. — *Прим. ред.*

[5] Высадка союзников на атлантическом побережье. — *Прим. ред.*

Теперь все чаще и чаще машины COLOSSUS приходили на помощь. Их растущий успех против SZ 42 оказался очень кстати, чтобы компенсировать трудности с дешифровкой ENIGMA. Английские успехи достигли кульминации в марте 1945 г.; с этого момента и до своего крушения Вермахт больше не предоставлял достаточного объема работы даже для Блетчли Парк.

19.2.7. Осётр. Английские усилия в меньшей степени были напрвлены против машины Сименса *Schlüsselfernschreibmaschine* (ключевая машина дальнего письма, *нем.*) T 52 (*Geheimschreiber* (тайнописец, *нем.*)), хотя две такие машины были взяты в качестве трофеев в Северной Африке частями английской Восьмой армии[6]. Причиной было то, что машины T 52a и T 52c использовались исключительно на кабельных линиях, и частично потому, что синхронизация раннего T 52 была недостаточно устойчивой на шумных радио-каналах. Поэтому там имелось гораздо меньше сообщений, пригодных для дешифрования. Кроме того, так как немецкие военно-воздушные силы, которые использовали главным образом T 52c и T 52e, имели низкую криптоаналитическую стойкость на своих линиях, обслуживаемых ENIGMA, то не было настоятельной необходимости слушать линии, на которых применялись T 52. С другой стороны, взлом канала связи ENIGMA немецкой Армии был значительно труднее (разд. 19.7) из-за более высокой дисциплины обслуживающего персонала. Однако, благодаря небрежности немецких операторов, некоторые *Geheimschreiber* шифры также были взломаны (первый в 1942 г. на линии Сицилия-Ливия и на линии от Эгейского моря до Сицилии под кодовым названием «Макрель»), несмотря на более трудную криптоаналитическую ситуацию. Машины T 52a и T 52d использовали кроме 32 подстановок также 32 перестановки пяти телетайпных бит, 30 из которых были различными; проводка, приводящая к этим перестановкам, была известна из немецких и американских патентов. Методы атаки, которые оказались полезными против SZ 42, удалось экстраполировать, хотя восстановление ключа было более трудоемким. Строение ключевых дисков было также в основном известно. Изменения в T 52a, которые привели к T 52c, были минимальными; «нерегулярное» движение T 52d и T 52e оказалось совсем не таким нерегулярным, каким могло бы быть. Шифры с *Klartextfunktion* (разд. 8.7.3), включающие автоключи, были крайне неподатливыми для незаконного дешифровальщика, — правда, и для законного получателя тоже, если радиоканал был зашумлен. Они практически не использовались.

В мае и июне 1940 г. шведский математик Бьюрлинг — гений, подобный Тьюрингу, работавший для *Försvarets Radioanstalt* (FRA), взломал, работая над немецким T 52a без использования вычислительной техники, телетайпную линию связи в Осло, проходившую над шведской территорией. Он использовал немецкое стереотипное применение «сдвинутого на одну букву» символа перед символом «пробела» и их (пробелов) изобилие. Он заметил, что если два открытых текста 1 и 2 имеют один общий бит и четыре различных, то это

[6] Тот факт, что COLOSSUS использовался главным образом против SZ 42, а не против T 52, как ранее предполагалось, был обнаружен Маликом (1980).

Рис. 145. Копия немецкой T 52, созданная шведским криптоаналитическим бюро FRA

же должно выполняться в случае, когда они оба зашифрованы в одной и той же ключевой позиции («дифференциальный криптоанализ»). Таким способом он сумел восстановить пробелы между словами. Некоторые детали того, что он делал дальше, он унес с собой в могилу. Как бы то ни было, из сообщения, перехваченного 25 мая 1940 г. и подтвержденного сообщением двумя днями позже, он оказался в состоянии полностью восстановить, наконец, машину T 52a/b и построил ее точные копии, названные «арреп» (рис. 145). 17 июня 1942 г. немцы были предупреждены об этом финским источником, но не отреагировали должным образом. В июле 1942 г. шведы смогли взломать даже сообщения T 52c. Взлом прекратился в мае 1943 г., когда немцы изменили индикаторные процедуры.

19.3. Наложение согласованного перешифрованного кода

До сих пор приводились примеры непосредственных наложений — поскольку криптотексты были уже согласованы — такими они были при функционировании швейцарской армейской машины ENIGMA, которая использовала одну и ту же начальную установку для всех сообщений одного дня. Но если два криптотекста зашифрованы при разных начальных установках (механически порожденных) ключей, они должны быть взаимно согласованы (подогнаны), чтобы добиться их наложения в правильном положении. Достичь этого можно, как показано в разд. 17.1, с помощью *Kanna* исследования. Криптотексты предположительно находятся в фазе, как только их взаимная *Kanna* становится максимальной и близкой к $\varkappa_S$ — при условии, что ключи вообще пересекаются.

19.3.1. Использование индикатора. Иногда тексты допускают более простое согласование. Если для перешифрования кода должны использо-

ваться часто меняющиеся ключи, то можно рекомендовать начинать каждое сообщение с новой ключевой позиции одной и той же ключевой последовательности. Избегая заранее спланированной процедуры, принято указывать начальную позицию в начале сообщения. Этот так называемый индикатор *Spruchschlüssel* — не путать с дискриминантом, *Kenngruppe*, который обозначает используемую систему — может означать что-либо из страницы или строки в книге, используемой в качестве ключевого текста для начальной установки дисков шифрующей машины. Это, конечно, прячет ключ, но не препятствует согласованию. Что касается наложения, то использование индикаторов — всего лишь иллюзия сложности, в том случае. когда из индикаторов двух сообщений можно подсчитать или как-нибудь определить их фазовую разность.

Поэтому в профессиональном шифре индикатор сам должен быть зашифрован (как, например, в ENIGMA). Если индикатор просто указывает страницу или строку книги, то надо знать лишь число строк на странице, чтобы подсчитать фазовый сдвиг. Это число обычно приблизительно известно.

Такая ситуация рассматривается в следующем примере Кана 4-значного кода с 4-значными числами в качестве составляющих. Индикатор перед сообщением включает два знака для страницы и два знака для строки. Если пять сообщений зашифровываются при помощи одной и той же страницы,

(i)	6218	6260	7532	8291	2661	6863	2281	7135	5406	7046	9128	...
(ii)	6216	3964	3043	1169	5729	3392	1952	7572	2754	7891	6290	...
(iii)	6218	4061	6509	4513	1881	0398	3402	8671	4326	8267	6810	...
(iv)	6218	5480	9325	3811	4083	5373	4882	8664	8891	6337	5914	...
(v)	6217	7260	8931	8100	5787	6807	2471	0480	9892	1199	8426	...

то они сразу могут быть согласованы:

			1	2	3	4	5	6	7	8	9	10	
(i)			6260	7532	8291	2661	6863	2281	7135	5406	7046	9128	...
(ii)	3964	3043	1169	5729	3392	1952	7572	2754	7891	6290	6719	7529	...
(iii)			4061	6509	4513	1881	0398	3402	8671	4326	8267	6810	...
(iv)			5480	9325	3811	4083	5373	4882	8664	8891	6337	5914	...
(v)		7260	8931	8100	5787	6807	2471	0480	9892	1199	8426	1710	...

Для частотного анализа столбцов шифрокодовых групп обычно не имеется достаточного материала. В случае линейной подстановки, в частности, для ад-

1

0000	5101	2209	1880	8339
5909	0000	7108	6789	3238
8801	3902	0000	9681	6130
9220	4321	1429	0000	7559
2771	7872	4970	3551	0000

5

0000	9391	6575	1590	4492
1719	0000	7284	2209	5101
4535	3826	0000	5025	8927
9510	**8801**	5085	0000	3902
6618	5909	2183	7108	0000

Рис. 146. Два примера таблиц разностей

разность в $\mathbb{Z}_{10}^4$	столбец	запись
⋮	⋮	⋮
$8209 = 0480 - 2281$	6	(v)-(i)
$\rightarrow 8801 = 4061 - 6260$	1	(iii)-(i)
$\rightarrow 8801 = 5373 - 7572$	5	(iv)-(ii)
$9077 = 6509 - 7532$	2	(iii)-(i)
$9106 = 5914 - 6810$	10	(iv)-(iii)
$\rightarrow 9220 = 5480 - 6260$	1	(iv)-(i)
$\rightarrow 9220 = 1881 - 2661$	4	(iii)-(i)
$9308 = 3811 - 4513$	3	(iv)-(iii)
$\rightarrow 9391 = 1952 - 2661$	4	(ii)-(i)
$\rightarrow 9391 = 6337 - 7046$	9	(iv)-(i)
$\rightarrow 9391 = 6810 - 7529$	10	(iii)-(ii)
$9510 = 5373 - 6863$	5	(iv)-(i)
⋮	⋮	⋮

Рис. 147. Встречающиеся разности (упорядоченные)

дитивного перешифрования в $\mathbb{Z}_{10}^4$, как можно предположить в данном случае, может помочь *симметрия позиции*, введенная в разд. 18.6.2. Таким образом, метод разностей образует для каждого столбца таблицу разностей, два примера которой (для первого и пятого столбцов) представлены на рис. 146.

Ввиду большого размера восьми таблиц разностей, каждая из которых имеет 20 существенных элементов, имеет смысл упорядочить все элементы для нахождения кратности их встречаемости. Соответствующий фрагмент такой таблицы приведен на рис. 147.

Если теперь разности встречаются неоднократно, то соответствующие вычитаемые нужно вычесть в соответствующих столбцах. Как показано на рис. 148, для разности 8801 число *6260* вычитается в столбце 4, число *7572* в столбце 5; аналогично, для разности 9391 число *2661* вычитается в столбце 4, *7046* в столбце 9, *7529* в столбце 10. Здесь эти два вычитания покрывают уже вычитания, произведенные вначале для разности 9220, а это подтверждает, что фазы были согласованы правильно.

	1′	2	3	4′	5′	6	7	8	9′	10′	
(i)	**0000**	7532	8291	**0000**	**9391**	2281	7135	5406	**0000**	2609	...
(ii)	**5909**	5729	3392	**9391**	**0000**	2754	7891	6290	9773	**0000**	...
(iii)	**8801**	6509	4513	**9220**	3826	3402	8671	4326	1221	**9391**	...
(iv)	**9220**	9325	3811	2422	**8801**	4882	8664	8891	**9391**	8495	...
(v)	2771	8100	5787	4246	**5909**	0480	9892	1199	1480	4291	...
	6260			*2661*	*7572*				*7046*	*7529*	

Рис. 148. Частично приведенные сообщения

В этих сведенных столбцах на рис. 148 простокодовые группы **0000**, **9391**, **5909**, **8801**, **9220**, **9391** встречаются повторно. Группы внизу отмеченных штрихами столбцов дают относительный ключ. Дальнейшие приведения приводят все пять сообщений в одноалфавитно зашифрованный промежуточный текст, в относительный простокод, который может быть рассмотрен так, как это сделано в разд. 18.3.2.

Метод разностей не всегда работает так хорошо, как можно видеть в этом примере. Часто сначала находятся лишь островки взаимосвязанных групп, и нужен дальнейший материал, чтобы объединить их в архипелаги. Если же глубина сообщений недостаточна, то не исключено, что могут быть достигнуты лишь частичные решения. Более того, могут встретиться неправильные совпадения разностей. В нашем примере разность 1480 получается не только из 9-го столбца: $1480 = 8426 - 7046$, но также из 6-го столбца: $1480 = 4882 - 3402$. Это означает приведение 6-го столбца с помощью 3402. Однако, как мы обнаружим позднее, это привело бы к плохим простокодовым группам.

19.3.2. Кунце. Эксперты криптографической службы немецкого Министерства иностранных дел Пашке, Кунце, Шауффлер и Ланглотц имели весьма большой стаж в своей профессии. Все они в 1918 или 1919 гг. вошли в состав команды, возглавляемой тогда Сельховом. Пашке был номинальным главой лингвистической секции. Доктор Кунце был математиком (в то время редкость для криптоаналитической службы). Начав в 1921 г. атаку на французский код перешифрования, он окончательно взломал его в 1923 г. Кунце возобновил свою работу в 1927 г. Таким образом, он имел многолетний опыт по отслаиванию кода перешифрования. Это учреждение сначала называлось «Z-секцией подразделения 1 Отдела кадров и бюджета». После реорганизации 1936 г. оно стало называться Pers Z.

Трудоемкая работа по отслаиванию была полуавтоматизирована. Круг построил таких «роботов», как их называет Кан, частично из перфокарточного оборудования, частично из стандартных компонентов телекоммуникаций. С их помощью Мюллер, Фридрихс и другие преуспели в дешифровании дипломатического кода США, действовавшего с августа 1941 г. по лето 1943 г. Аллен Даллес, бывший тогда начальником американской секретной службы в Европе, ничего не подозревал, пока не получил предупреждения от Гизевиуса из немецкого Сопротивления.

Аналогичные идеи развивались в *Xu*, отделе шифров OKW, инженерами Ротшейдтом и Йенсеном, что уже отмечалось в разд. 18.6.3. А B-служба Военно-морского флота преуспела в раскрытии английских морских шифров.

Союзники и некоторые бюро нейтральных стран использовали одну и ту же технику. «Единственная наиболее общая криптоаналитическая процедура войны состояла в отслаивании многочисленных добавок от зашифрованного кода» (Дэвид Кан).

В 1936 г. Кунце сделал прекрасную работу, состоящую во взломе японской роторной машины ORANGE (разд. 8.5.7), а позднее машины RED. В 1997 г. Лейберих сообщил, что во время войны в Министерстве иностранных дел

работало 12 лингвистов (среди них Раве, который поддерживал связь с Хюттенхайном, сотрудничая с OKW *Хи*), работающих день за днем над сигналами PURPLE японского посла Хироси Осима. Министр иностранных дел Рейха Риббентроп рассматривал Per Z как особое оружие в борьбе с его врагами Герингом и Гиммлером.

19.4. Криптотекст—криптотекст компромиссы

На практике встречается трудная ситуация, когда сообщение должно быть повторено лишь с незначительными изменениями, например, при исправлении типографской ошибки. Если исправленное сообщение посылается снова с *тем же* ключом, то компромисс открытый текст—открытый текст начинается с позиции ошибки со всеми отрицательными эффектами, которые мы уже знаем. Если же исправленное сообщение посылается снова, но с *другим* ключом, то имеет место компромисс криптотекст—криптотекст до позиции ошибки.

19.4.1. Криптотекст—криптотекст компромисс ключей. Этот компромисс случается довольно часто, если одно и то же сообщение или, по крайней мере, большая его часть зашифровывается дважды или более (каждый раз с новым ключом). Существует большая опасность взлома шифра, если это делается в одной и той же системе; если получающиеся в результате криптотексты «изоморфны» (разд. 2.6.3), то они имеют равную длину, что обычно очень заметно. Классический пример такой атаки был проведен в декабре 1938 г. и в январе 1939 г. над парой радиосообщений от румынского военного атташе в Париже в его министерство иностранных дел, которые отличались по длине лишь на две 5-буквенные группы. Согласно Хюттенхайну, *Хи* достиг успеха в дешифровании этих сообщений; оказалось, что лишь один фрагмент открытого текста (тетрадь 17) первого сообщения был заменен фрагментом открытого текста (тетрадь 15 вместо 17) во втором сообщении.

Криптотекст—криптотекст компромисс ключей присущ системам связи, если циркулярное сообщение должно быть послано сотням получателей, причем, каждому с отдельным ключом. Похоже, что немецкие криптологи недооценивали эту опасность так, что их офицеры связи не были достаточно предупреждены о ней, и эта небрежность продолжалась до конца войны. Контрадмирал Штуммель, отвечающий за криптоаналитическую надежность радиосообщений немецких военно-морских сил, ввел в 1943 г. большое число ключевых сетей, а в 1944 г. дал каждой подводной лодке собственный ключ. Предполагалось, что это даст противнику так много индивидуальной работы, что в результате добавит немцам криптоаналитической надежности. Но это была саморазрушающая сложность: «...она была активно полезна, потому что одно и то же сообщение часто появлялось зашифрованным различными ключами, иногда в разные дни.» (Носквис). Даже Штуммель не мог справиться с формулировкой общих приказов отдельно для каждой ключевой сети и тем более для каждой лодки.

Когда 1 февраля 1942 г. 4-роторная ENIGMA была введена только для субмарин, компромиссы стали часто случаться при передаче общих прика-

зов, зашифрованных 3-роторной машиной ENIGMA, для других судов. Здесь уместно привести предупреждение главы 40-й комнаты в 1914 г. сэра Эвинга: «Никогда не следует смешивать ваши шифры. Подобно смешиванию ваших вин, это может привести к измене самим себе». Но как бы то ни было, штаб *Гроссадмирала* Деница делал это.

Строго говоря, мы не можем говорить о криптотекст-криптотекст компромиссе ключей, если они публичные. Но заметим, что публичны лишь шифрующие ключи, а не расшифровывающие, а именно последние мы имеем в виду — в симметричном методе шифрования никакой разницы нет. На самом деле, риск криптотекст-криптотекст компромисса присущ и публичным ключевым системам.

В жаргоне Блетчли Парк криптотекст—криптотекст компромисс называется «поцелуй» (*англ.*, kiss). Нельзя лучше выразить радость от такой неожиданной удачи. К счастью для англичан, менее крупные суда немецкого флота не имели машин ENIGMA, а пользовались простым биграммным шифром. На больших кораблях не имели этого шифра или не любили им пользоваться. Если какие-нибудь сообщения, например, предупреждения о плавающих минах, должны были быть переданы быстро, то никто не беспокоился о том, чтобы переформулировать открытый текст. Англичане иногда сами провоцировали такие ситуации с целью создания криптотекст—криптотекст компромисса трудных 4-роторных машин ENIGMA с легко взламываемыми биграммными подстановками. Они с британским юмором называли это «садоводством». Эта техника означала переход к классической ситуации компромисса открытый текст—криптотекст, где дешифрованное сообщение снабжает «шпаргалку» не только вероятными, но определенно содержащимися словами.

7 мая 1941 г. немецкое метеорологическое судно *Мюнхен* было захвачено английским военным флотом и машина Wetterkurzschlüssel (краткий погодный шифратор, нем.) попала в руки англичан. Из метеорологических сообщений для подводных лодок возник целый поток «поцелуев», наполнивший «шпаргалки» Блетчли Парк для машины ENIGMA, и это продолжалось до 1944 г. Игроки в бридж называли это «cross-ruffs» (побить козырем, *англ.*)[7]. Захват подводной лодки U-559 30 октября 1942 г. имел результатом компромисс новой 4-роторной машины ENIGMA с помощью метеорологических сообщений.

Урок состоит в том, что параллельное применение кода, перешифрованного одноразовым ключом, порожденным машиной, вместе с повторно используемыми добавками, в случае, когда последний шифр уже взломан, компрометирует этот «индивидуальный» ключ и приводит к реконструкции машины, которая его произвела. Такое несчастье приключилось с немецким шифром Министерства иностранных дел, который, предположительно из-за нехватки ключевого материала, на линии Берлин—Дублин использовал систему перешифрования FLORADORA, которая уже (см. разд. 9.2.1) была взломана

[7]Операция «cross-ruffs» была успешно проведена в июле 1918 г. Чайлдсом из G.2A.6, AEF с использованием телеграммы Макензена об отводе немецких войск в Румынии.

англичанами. Ключ, таким образом, можно было исследовать; оказалось он порождался модифицированной машиной Lorenz SZ40, уже реконструированной в Блетчли Парк. Таким образом, вся связь немецкого Министерства иностранных дел, считавшаяся невзламываемой, была раскрыта.

19.4.2. Сведение к компромиссу открытый текст—открытый текст. Для случая шифра ВИЖЕНЕР, в частности, для шифрования с помощью добавок, компромисс шифротекст—шифротекст может быть сведен просто к компромиссу открытый текст—открытый текст: открытый текст рассматривается как ключевой текст, а ключевой текст — как открытый текст. Это просто другой случай обмена ролей, с которым мы встретились в разд. 19.2.3. Это значит, что применимы методы наложения и симметрии, и не требуется, чтобы многоалфавитный шифр был периодическим. Предварительное условие для замены ролей снова состоит в том, что все ключи являются осмысленными текстами и указаны частотные характеристики и/или шаблоны.

Например, пусть даны пять сообщений равной длины

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22 23 24 25	26 27 28 29 30
(i)	T C C V L	E S K P T	X M P V W	H Y M V G	X B O R V	C W A R F
(ii)	V L L B V	C K W F P	E H E C F	C G N Z E	K K K V I	H D D I D
(iii)	M Y Y R D	M J W M C	U I G L O	K M X L R	E W H X M	R J H A S
(iv)	B K Q T Z	B Z W K W	Z X G Z O	V T B A T	K W M G M	R J K L P
(v)	M Y Y V H	B W J D X	C P C Z O	H V T S I	V M E B S	O H R A U

	31 32 33 34 35	36 37 38 39 40	41 42 43 44
(i)	R R D Y C	T K L B L	M G L W
(ii)	S V F K Q	A J V C R	F K L K
(iii)	H B R N U	T R V G J	X J P W
(iv)	K O W H U	C B D U F	T V E F
(v)	S D A N I	T Y H F K	Z Z W G

1

0	24	**7**	18	**7**
2	0	**9**	20	**9**
19	17	0	11	0
8	6	15	0	15
19	17	0	11	0

13

0	11	9	9	13
15	0	24	24	**2**
17	**2**	0	0	**4**
17	**2**	0	0	**4**
13	24	22	22	0

18

0	25	15	11	19
1	0	16	12	20
11	10	0	22	4
15	14	4	0	8
7	6	22	18	0

22

0	17	5	5	15
9	0	14	14	24
21	12	0	0	10
21	12	0	0	10
11	**2**	16	16	0

27

0	19	13	13	15
7	0	20	20	22
13	6	0	0	**2**
13	6	0	0	**2**
11	4	24	24	0

36

0	19	0	17	0
7	0	**7**	24	**7**
0	19	0	17	0
9	**2**	**9**	0	**9**
0	19	0	17	0

Рис. 149. Шесть таблиц разностей

и имеется причина ожидать, что использована линейная подстановка. Для каждого столбца определены разности над $\mathbb{Z}_{26}$. Шесть из них показывают отдельные совпадения в разностях 4, 7 и 11, как это отмечено жирным шрифтом на рис. 149.

В колонке 18 отмечено жирным шрифтом одно число **11** как сумма **4** и **7**, в колонке 1 разность 11 исключается, так как 11 и 19, дополнение разности **7**, стоят в одной строке.

Среди других разностей, встречающихся в колонке 36, **2** и **9** имеются также в колонке 1, **2** в колонке 13 и **9** в колонке 22. В таблице разностей колонки 13 две разности **2** второго столбца должны различаться от разности **2** в пятой колонке; одна из этих разностей встречается в колонке 22. Потом окажется, что разности **9** и **11** в первой строке колонки 13 случайны.

Беря теперь первую колонку в качестве эталона и выравнивая остальные пять колонок на основе этих разностей, создадим следующий каркас:

	1	13′	18′	22′	27′	36′
(i)	T	G	M	M	M	M
(ii)	V	V	N	V	T	T
(iii)	M	X	X	H	Z	M
(iv)	B	X	B	H	Z	V
(v)	M	T	T	X	X	M
	0	*9*	*0*	*15*	*10*	*7*

Этот каркас можно протестировать столбцами 5, 6, 7, 10, 12, 19, 20, 24, 30, 31, 33, 37, 38. В результате получим:

	5′	6′	7′	10′	12′	19′	20′	24′	30′	31′	33′	37′	38′
(i)	X	P	T	X	B	H	Z	X	X	W	V	M	B
(ii)	H	N	L	T	W	L	X	B	V	X	A	L	L
(iii)	P	X	K	G	X	X	K	D	K	M	M	T	L
(iv)	L	M	A	A	M	M	M	M	H	P	R	D	T
(v)	T	M	X	B	E	E	B	H	M	X	V	A	X
	14	*15*	*25*	*22*	*11*	*14*	*7*	*20*	*8*	*21*	*5*	*24*	*10*

Повторная встречаемость четырех символов M, T, V, X свидетельствует, что мы на верном пути. Другие частые символы B, L, A, G, K могут быть использованы для продолжения порождения разностей. Это позволяет выравнять 40 из 44 столбцов:

	1′	2′	3′	4′	5′	6′	7′	8′	9′	10′	11′	12′	13′	14′	15′	16′	17′	18′	19′	20′	21′	22′
(i)	T	E	B	V	X	P	T	L	U	X	Z	B	G	G	B	G		M	H	Z	X	M
(ii)	V	N	K	B	H	N	L	X	K	T	G	W	V	N	K	B		N	L	X	K	V
(iii)	M	A	X	R	P	X	K	X	B	G	W	X	X	W	T	J		X	X	K	E	H
(iv)	B	M	P	T	L	M	A	X	P	A	B	M	X	K	T	U		B	M	M	K	H
(v)	M	A	X	V	T	M	X	K	I	B	E	E	T	K	T	G		T	E	B	V	X
	0	*24*	*1*	*0*	*14*	*15*	*25*	*25*	*21*	*22*	*24*	*11*	*9*	*15*	*21*	*1*		*0*	*14*	*7*	*0*	*15*

	23′	24′	25′	26′	27′	28′	29′	30′	31′	32′	33′	34′	35′	36′	37′	38′	39′	40′	41′	42′	43′	44′
(i)	O	X	K		M	B	K	X	W	H	Y	L	B	M	M	B	G	Z			A	X
(ii)	K	B	X		T	E	B	V	X	L	A	X	P	T	L	L	H	F			A	L
(iii)	H	D	B		Z	I	T	K	M	R	M	A	T	M	T	L	L	X			E	X
(iv)	M	M	B		Z	L	E	H	P	E	R	U	T	V	D	T	Z	T			T	G
(v)	E	H	H		X	W	T	M	X	T	V	A	H	M	A	X	K	Y			L	H
	0	*20*	*11*		*10*	*25*	*7*	*8*	*21*	*10*	*5*	*13*	*1*	*7*	*24*	*10*	*21*	*12*			*11*	*25*

Как и раньше, добавки, записанные в линии сигнатуры, должны прибавляться к буквам промежуточного текста для получения символов оригинального криптотекста. Таким образом, они сами являются шифрами ЦЕЗАРЯ для псевдоключа (начального открытого текста), который читается так:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
A	*Y*	*B*	*A*	*O*	*P*	*Z*	*Z*	*V*	*W*	*Y*	*L*	*J*	*P*	*V*	*B*	*	*A*	*O*	*H*	*A*	*P*	*A*	*U*	*L*	*	*K*	*Z*	*H*	*I*

31	32	33	34	35	36	37	38	39	40	41	42	43	44
V	*K*	*F*	*N*	*B*	*H*	*Y*	*K*	*V*	*M*	*	*	*L*	*Z*

Теперь требуется перебор: среди 26 возможных выравниваний прибавление числа 19 приводит к следующему фрагментарному английскому открытому тексту:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
t	r	u	t	h	i	s	s	o	p	r	e	c	i	o	u	*	t	h	a	t	i	t	n	e	*	d	s	a	b

31	32	33	34	35	36	37	38	39	40	41	42	43	44
o	d	y	g	u	a	r	d	o	f	*	*	e	s

В открытом тексте соответствующая полная цитата из автобиографии Черчилля 1949 г. читается так: «*In wartime, truth is so precious that she should always be attended by a bodyguard of lies* (в военное время правда так драгоценна, что она всегда должна сопровождаться ложью, *англ.*)» (Черчилль Рузвельту и Сталину, ноябрь 1943 г.).

Эти пять псевдо-открытых текстов (пять начальных ключей) можно легко восстановить. Они взяты из «детской книги», написанной не для детей:

«*Alice was beginning to get very tired of sitting by he*[*r sister on the bank...*] (Алиса начала уставать от сидения рядом с сестрой на берегу..., *англ.*)»,

«*Curiouser and curiouser*! — *cried Alice* (*she was so much s*[*urprised...*] (Все любопытней и любопытней! — кричала Алиса (она была так сильно удивлена))»,

«*They were indeed a queer-looking party that assemble*[*d on the bank*] (Они действительно выглядели очень странной группой, собравшейся на берегу)»,

«*It was the White Rabbit, trotting slowly back again, an*[*d looking...*] (Это был Белый Кролик, бежавший медленной рысью снова назад и смотревший...)»,

«*The Caterpillar and Alice looked at each other for so*[*me time in silence*]» (Гусеница и Алиса в течение некоторого времени молча глядели друг на друга).

19.5. Метод Синкова

19.5.1. Прямое произведение ключей. В самом общем случае многоалфавитных шифров с неродственными алфавитами может быть апробирован метод, который способен работать даже всего с двумя шифрованиями одного и того же открытого текста, если два ключа являются периодическими с известными периодами разной длины.

Следующий пример такого компромисса криптотекст—криптотекст, впервые опубликованный Синковым в 1968 г., объясняет эту процедуру, примененную в засекреченной работе 1938 г. Фридмана (рассекреченной в 1984 г.).

Пусть имеются два сообщения одного дня, каждое состоит из 149 символов:

(i)	WCOAK	TJYVT	VXBQC	ZIVBL	AUJNY	BBTMT
	JGOEV	GUGAT	KDPKV	GDXHE	WGSFD	XLTMI
	NKNLF	XMGOG	SZRUA	LAQNV	IXDXW	EJTKI
	YAOSH	NTLCI	VQMJQ	FYYPB	CZOPZ	VOGWZ
	KQZAY	DNTSF	WGOVI	IKGXE	GTRXL	YOIP
(ii)	TXHHV	JXVNO	MXHSC	EEYFG	EEYAQ	DYHRK
	EHHIN	OPKRO	ZDVFV	TQSIC	SIMJK	ZIHRL
	CQIBK	EZKFL	OZDPA	OJHMF	LVHRL	UKHNL
	OVHTE	HBNHG	MQBXQ	ZIAGS	UXEYR	XQJYC
	AIYHL	ZVMQV	QGUKI	QDMAC	QQBRB	SQNI

Поскольку оба криптотекста имеют равные длины, возникает подозрение, что соответствующие открытые тексты идентичны. Сначала проводим исследование периода. Оказывается, что ключ для первого криптотекста предположительно имеет период 6, а ключ для второго — предположительно 5. В таком случае 30 является общим периодом обоих (неизвестных) ключей. Тогда каждое совпадение символов между двумя текстами должно повторяться на расстоянии 30 позиций. Действительно, написанное выше показывает, что совпадение XX в столбце 12 повторяется в столбце 42 как совпадение DD, в столбце 72 как совпадение ZZ и т. д. Это наблюдение подтверждает предположение, что оба криптотекста принадлежат к одному и тому же открытому тексту. В действительности

$$12+30i=\begin{cases}6\ (\mathrm{mod}\ 6),\\ 2\ (\mathrm{mod}\ 5),\end{cases}\qquad 15+30i=\begin{cases}3\ (\mathrm{mod}\ 6),\\ 5\ (\mathrm{mod}\ 5).\end{cases}$$

Поэтому 6-й алфавит первого сообщения должен совпадать со 2-м алфавитом второго сообщения, а 3-й алфавит первого сообщения должен совпадать с 5-м алфавитом второго сообщения. Подсчет соответствующего Xu дает высокие значения, подкрепляющие эту гипотезу.

Теперь метод Синкова разлагает эти два сообщения в соответствии с двумя используемыми ключами. Шесть алфавитов, применяемых для шифрования первого сообщения, назовем α, β, γ, δ, ε, ζ, а пять алфавитов, используемых для шифрования второго сообщения, назовем ι, κ, λ, μ, ν.

Таким образом, начало этого разложения читается так:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
α	W						J						B						B						Y					
β		C						Y						Q						L							B			
γ			O						V						C						A						B			
δ				A						T						Z						U						T		
ϵ					K						V						I						J						M	
ζ						T						X						V						N						T
ι	T						J				M					E					E					D				
κ		X						X					X				E					E					Y			
λ			H						V					H					Y					Y				H		
μ				H						N					S					F					A				R	
ν					V						O					C					G					Q				K

Используя эту диаграмму можно найти новые входы. Действительно, поскольку оба криптотекста соответствуют одному и тому же открытому тексту, то столбец 2, столбец 7 и столбец 12 (все показывающие значение X для ключа κ), могут быть совмещены. Кроме того, столбец 16 и столбец 21 (оба показывающие значение E для ключа ι); столбец 17 и столбец 22 (оба показывающие значение E для ключа κ); столбец 18 и столбец 23 (оба показывающие значение Y для ключа λ) тоже могут быть совмещены. Это приводит к следующему построению:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
α	W	J	B				J					J	B						B						Y			B		
β		C					C	Y				C		Q						L							B			
γ			O						V				O		C	A					A						B	O		
δ			T	A						T			T			Z	U				Z	U						T		
ϵ					K						V						I	J				I	J						M	
ζ		X				T	X					X						V					V	N						T
ι	T						J				M					E					E					D				
κ		X						X					X				E					E					Y			
λ			H						V					H					Y					Y				H		
μ				H						N					S					F					A				R	
ν					V						O					C					G					Q				K

Тем же способом совмещение может быть сделано также в строках ι, κ, λ, μ, ν: столбец 13 и столбец 19, оба показывающие B для ключа α, могут быть совмещены и т. д. В итоге мы получаем следующую таблицу:

	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20	21 22 23 24 25	26 27 28 29 30
α	W J B M	D J P B	J B W M	J N B U	J N X Y	L B J D
β	Q C K E X	C Y X K	C K Q E	C A K L	C A T B	K C
γ	L A O C V	K A W V O	N A O L C	A Z O G	A Z Q	S B O A K
δ	G Z T A F	Z F T	Z T G A	Z U X T I	Z U X L	T Z
ϵ	Y M D R K	F M K D	V M D Y R	M I J D W	M I J S	D M F
ζ	I X Q Z D	T X D Q	X Q I Z	X E V Q P	X E V N G	Y Q X T
ι	T E U Z	J E S Z	M E T U	E C Q	E C H O	D E J
κ	I X Q Z D	T X D Q	X Q I Z	X E V Q P	X E V N G	Y Q X T
λ	J S H B	S V H	I S H J B	S Y H N	S Y M	A H S
μ	S R F H N	R G N F	R F S H	R M F Y	R M A B	Q F R
ν	L A O C V	K A W V O	N A O L C	A Z O G	A Z Q	S B O A K
$\rightarrow$	A B C H D	E B F D C	G B C A H	B I J C K	B I J L Q	M N C B E

	31 32 33 34 35	36 37 38 39 40	41 42 43 44 45	46 47 48 49 50	51 52 53 54 55	56 57 58 59 60
α	J B	Y U J B	P B	W B J H M	P W Y D	W B J W
β	C G K	T L G C K	X X Y K X	Q K C E	Y Q T	X Q K C Q
γ	A O N	Q G A O	V V W O V	L O A C	W L Q K	V L O A L
δ	Z T E	L I Z T	F F T F	G T Z E A	G L	F G T Z G
ϵ	M D P V	S W M D	K K D K	Y D M P R	Y S F	K Y D M Y
ζ	X H Q	G P H X Q	D D Q D	I Q X Z	I G F T	D I Q X I
ι	E M	O Q E	Z Z S Z	T E U	S T O J	Z T E T
κ	X H Q	G P H X Q	D D Q D	I Q X Z	I G F T	D I Q X I
λ	S K H	M N K S H	V H	J H S B	V J M	J H S J
μ	R F I	B Y R F	N N G F N	S F R I H	G S B J	N S F R S
ν	A O N	Q G A O	V V W O V	L O A C	W L Q K	V L O A L
$\rightarrow$	B O C P G	Q K O B C	D D F C D	A C B P H	F A Q R E	D A C B A

	61 62 63 64 65	66 67 68 69 70	71 72 73 74 75	76 77 78 79 80	81 82 83 84 85	86 87 88 89 90
α	N B Y D	J M B W	Y M R J	Y B N	N B J W	M B W
β	A K T	C E G K Q	T E U C	T K A V	A K C Q	E K X Q
γ	Z O N Q K	A C O L	Q C A	Q O Z F	I Z O A L	C J O V L
δ	X T L	Z A T G	L A Z	L T X	X T Z G	A T F G
ϵ	J D V S F	M R D Y	S R M	D A D J	J D M Y	R D K Y
ζ	V Q G T	X Z H Q I	G Z X	G J Q V	V Q X I	Z K Q D I
ι	C M O J	E U T	O U E	O C X	L C E T	U Z T
κ	V Q G T	X Z H Q I	G Z X	G J Q V	V Q X I	Z K Q D I
λ	Y H I M	S B K H J	M B D S	M H Y	Y H S J	B H J
μ	M F B	R H F S	B H P R	B F M K	M F R S	H F N S
ν	Z O N Q K	A C O L	Q C A	Q O Z F	I Z O A L	C J O V L
$\rightarrow$	J C G Q E	B H O C A	Q H S T B	Q U C J V	W J C B A	H X C D A

	91 92 93 94 95	96 97 98 99 00	01 02 03 04 05	06 07 08 09 10	11 12 13 14 15	16 17 18 19 20
α	YNB	XTUMU	BM Y	WLP	MJ UZ	BWUM
β	TAK	LEL	KEJT	XQ YB	EC L	VKQLE
γ	QZO E	GCG	NOC Q	VLBWS	CA GR	FOLGC
δ	LXTS	IAI	TA L	FG	AZ I	TGIA
ϵ	SJD H	WRW	VDR S	KY	RMOW	DYWR
ζ	GVQ	NBPZP	QZ G	DIY	ZX P	QIPZ
ι	OC	H QUQ	M U O	ZT SD	UE Q	X TQU
κ	GVQ	NBPZP	QZ G	DIY	ZX P	QIPZ
λ	MYH	NBN	IHB M	JAV	BSEN	HJNB
μ	BMFT	A YHY	FHXB	NS GQ	HR Y	KFSYH
ν	QZO E	GCG	NOC Q	VLBWS	CA GR	FOLGC
$\rightarrow$	QJCYZ	L1KHK	GCH2Q	DANFM	HB3K4	VCAKH

1..	21 22 23 24 25	26 27 28 29 30	31 32 33 34 35	36 37 38 39 40	41 42 43 44 45	46 47 48 49
α	KWNMW	NY	UYO	U YXM	UBMJL	PBU
β	QAEQ	XATBX	LT VG	LXT E	LKEC	YKL
γ	LZCL	VZQSV	GQ FI	GVQ C	GOCAB	WOG
δ	GXAG	FXL F	IL	IFL A	ITAZ	TIE
ϵ	YJRY	KJS K	WS	WKS R	WDRM	DWP
ζ	IVZI	DVG D	PG H	PDGNZ	PQZXY	QP
ι	ATCUT	ZCODZ	QO X	QZOHU	Q UE	S Q
κ	IVZI	DVG D	PG H	PDGNZ	PQZXY	QP
λ	JYBJ	YM	NMU K	N M B	NHBSA	VHN
μ	SMHS	NMBQN	YB K	YNBAH	YFHR	GFYI
ν	LZCL	VZQSV	GQ FI	GVQ C	GOCAB	WOG
$\rightarrow$	5AJHA	DJQMD	KQ6VO	KDQLH	KCHBN	FCKP

Как и следовало ожидать, алфавиты, обозначенные γ и ν, идентичны[8)], идентичны и алфавиты, обозначенные ζ и κ. Теперь все места в таблице заполнены, и некоторые столбцы, такие, как 8-й и 17-й, возможно, совпадают. Действительно, имеется в целом 32 различных столбца, которые *ad hoc* (специально для этого, *лат.*) обозначены A, ..., Z и 1, 2, ..., 6 и выписаны, более или менее, в порядке их образования в процессе совмещения. Тридцать два символа — это больше, чем 26 букв общего алфавита. На самом деле, никто не запрещает противнику использовать алфавит открытого текста с более чем 26 символами. Но наиболее вероятно, что некоторые столбцы соответствуют одному и тому же символу открытого текста. Таким образом, мы предположительно достигли одноалфавитно зашифрованного промежуточного текста, но это шифрование содержит омофоны — довольно удивительный результат.

К счастью, окажется, что жертвами омофонии являются не наиболее частые символы — как это часто делается с целью выравнивания частот — а

[8)] Для дальнейшей ручной работы мы будем отождествлять соответствующие строки, но в программном исполнении проще их повторить.

лишь редкие символы; они слишком редки, чтобы обеспечить достаточно материала для заполнения соответствующих мест.

19.5.2. Промежуточный шифр. Он задается линией сигнатуры таблицы

```
A B C H D   E B F D C   G B C A H   B I J C K   B I J L Q   M N C B E
B O C P G   Q K O B C   D D F C D   A C B P H   F A Q R E   D A C B A
J C G Q E   B H O C A   Q H S T B   Q U C J V   W J C B A   H X C D A
Q J C Y Z   L 1 K H K   G C H 2 Q   D A N F M   H B 3 K 4   V C A K H
5 A J H A   D J Q M D   K Q 6 V O   K D Q L H   K C H B N   F C K P
```

и ясно показывает частотное распределение английского языка с С-пиком и примерно равными частотами для B, A, H, D, O, N. Из многих путей, которые могут дать вход, мы используем предположение, что вероятным словом является /treasurysecretary/ (министр финансов, *англ.*), соответствующее начальному шаблону *1234567538231427*. С восемью буквами достигается многое:

```
t r e a s   u r y s e   c r e t a   r I J e K   r I J L Q   M N e r u
r O e P c   Q K O r e   s s y e s   t e r P a   y t Q R u   s t e r t
J e c Q u   r a O e t   Q a S T r   Q U e J V   W J e r t   a X e s t
Q J e Y Z   L 1 K a K   c e a 2 Q   s t N y M   a r 3 K 4   V e t K a
5 t J a t   s J Q M s   K Q 6 V O   K s Q L a   K e a r N   y e K P
```

Во второй строке глаз ловит /yesterday/, но это помогает немного. Чуть более раннее /congress/ дает больше. Теперь определены 12 букв, и лишь две из наиболее частых *etaonirsh* еще неизвестны /i/ и /h/:

```
t r e a s   u r y s e   c r e t a   r I J e n   r I J L o   M N e r u
r g e d c   o n g r e   s s y e s   t e r d a   y t o R u   s t e r t
J e c o u   r a g e t   o a S T r   o U e J V   W J e r t   a X e s t
o J e Y Z   L 1 n a n   c e a 2 o   s t N y M   a r 3 n 4   V e t n a
5 t J a t   s J o M s   n o 6 V g   n s o L a   n e a r N   y e n d
```

Теперь мы избавляемся от некоторых омофонов: /henry/ в первой строке означает, что I омофонно с F и означает /y/. Для /i/ дело обстоит сложнее, но в последней строке мы можем прочесть /no sings/, если 6 омофонно с D и означает /s/. Таким образом:

```
t r e a s   u r y s e   c r e t a   r y h e n   r y h L o   M N e r u
r g e d c   o n g r e   s s y e s   t e r d a   y t o R u   s t e r t
h e c o u   r a g e t   o a S T r   o U e h i   W h e r t   a X e s t
o h e Y Z   L 1 n a n   c e a 2 o   s t N y M   a r 3 n 4   i e t n a
5 t h a t   s h o M s   n o s i g   n s o L a   n e a r N   y e n d
```

что дает во второй строке /to muster/, в третьей строке /approve higher taxes/ (S и T омофонны), в четвертой строке /help finance a costly war in vietnam/, откуда, наконец, в первой строке возникает имя /henry h fowler/, и открытый текст читается так:

```
t r e a s   u r y s e   c r e t a   r y h e n   r y h f o   w l e r u
r g e d c   o n g r e   s s y e s   t e r d a   y t o R u   s t e r t
h e c o u   r a g e t   o a p p r   o v e h i   g h e r t   a x e s t
o h e l p   f i n a n   c e a c o   s t l y w   a r i n v   i e t n a
m t h a t   s h o w s   n o s i g   n s o f a   n e a r l   y e n d
```

Теперь 32 столбца свелись к 21 символу; 5 букв отсутствуют в открытом тексте. Фрагментарная шифровальная таблица имеет вид:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
α	M				B	X		N	T		L	K	U	Y	R		J	O	W	D	Z			P	
β	E		J		K		G	A	V				L	T	U		C	X	Q			B		Y	
$\gamma=\nu$	C		N		O		I	Z	F		B		G	Q	E		A	V	L	K	R	S	J	W	
δ	A			E	T			X			S		I	L			Z	F	G					U	
ϵ	R		V	P	D			J	O				W	S	H		M	K	Y	F	A			I	
$\zeta=\kappa$	Z				Q	N	H	V	B		Y	F	P	G			X	D	I	T	J		K	E	
ι	U		M			H	L	C	X			A	Q	O			E	Z	T	J		D		S	
λ	B		I		H		K	Y	E		A		N	M	D		S	U	J					V	
μ	H		X	I	F	A		M	K		T	J	Y	B	P		R	N	S			Q		G	
	H		G	P	C	L	O	J	V		N	R	K	Q	S		B	D	A	E	U	M	X	F	
$\rightarrow$			2				W		1		Y	5			T			6			4			I	
									3						Z											

19.5.3. Восстановление. Полная шифровальная таблица может быть восстановлена, если алфавиты были получены сдвигом первичного алфавита. Оказывается, имеет место именно этот случай, и метод, развитый Фридманом (упомянутый в разд. 18.8.2), позволяет восстановить первичный алфавит. Это помогает выполнить принцип Рорбаха, хотя для дешифрования и не обязательно.

Фридман начал с наблюдения, что в требуемой шифровальной таблице все *столбцы* получаются из какого-то одного циклическим сдвигом. Возьмем, например, строки λ и μ, и выберем в этих строках под произвольной буквой (скажем, /t/) символы J и S на некотором (неизвестном) расстоянии k, потом под буквой /r/ символы S и R, под /s/ — U и N, потом N и Y, Y и M, M и B, B и H, H и F, наконец, V и G — все с одним и тем же расстоянием k. Таким образом, мы имеем уже три цепочки

$$\text{J-S-R}, \quad \text{U-N-Y-M-B-H-F}, \quad \text{V-G}$$

с расстоянием k. Но символы J и S находятся также под буквой /r/ в строках α и λ (на том же расстоянии k), тогда на том же расстоянии находятся также символы W и J, R и D, O и U, U и N, P и V, L и A этих строк. Эти цепочки могут быть расширены новыми звеньями; теперь мы имеем

$$\text{W-J-S-R-D}, \quad \text{O-U-N-Y-V-B-H-F}, \quad \text{P-V-G}, \quad \text{L-A}.$$

Расстояние между символами J и D равно $3k$; в строках ι и α находим под буквой /u/ символы J и D, таким образом, не только U и M, O и Y, но также C и N, X и T, A и K, Q и U, E и J, Z и O, T и W, S и P имеют расстояние $3k$. Теперь мы имеем цепочки

T-E-*-W-J-S-R-D-P-V-G, Q-C-O-U-N-Y-M-B-H-F-*-X, L-A-*-*-K.

Мы можем теперь замкнуть эти цепные фрагменты, заметив, что W и G, которые находятся в строках α и δ, имеют расстояние $7k$; таким образом, это также является расстоянием между J и Z, B и T, M и A, U и I, H и E, Y и L, которые замыкают цикл:

T-E-K-W-J-S-R-D-P-V-G-Z-Q-C-O-U-N-Y-M-B-H-F-I-X-L-A-.

Но это не обязательно исходный порядок. Применяя метод разд. 18.8.1, замечаем, что пятая степень

T-S-G-U-H-A-J-V-O-B-L-W-P-C-M-X-K-D-Q-Y-I-E-R-Z-N-F-

приносит успех: последовательность

H A J V O B L W P C M X K D Q Y I E R Z N F T S G U

получается чтением по столбцам из пароля HOPKINS методом, описанным в разд. 3.2.5:

H O P K I N S
A B C D E F G
J L M Q R T U
V W X Y Z

С помощью этой последовательности строится *tabula recta* (табл. 26).

Для определения заглавной строки можно сделать следующее. Взять какую-либо «богатую» строку (т. е. имеющую много определенных символов открытого текста), вроде названной $\gamma = \nu$, переупорядоченную в соответствии с вышесказанным:

r x s e l t y a u o g p v h c i w n
$\gamma = \nu$ H A J V O B L W P C M X K D Q Y I E R Z N F T S G U

Тогда другие строки, рассмотренные к этому моменту, попадут на место и определят дальнейшие символы открытого текста. В итоге фрагментарно получаем заглавную строку:

* r x s e l t y * a f m u * o * g p v h c i * w n d

где все еще отсутствуют лишь редкие символы открытого текста: /b/, /j/, /k/, /q/, /z/.

Теперь эта заглавная строка тоже обнаруживает свою тайну: она построена методом, описанным в разд. 3.2.5 с помощью пароля /johns/.

j	o	h	n	s
a	b	c	d	e
f	g	i	k	l
m	p	q	r	t
u	v	w	x	y
z				

Это позволяет установить пять пропущенных символов.

Таблица шифрования (табл. 26) попадает под определение «тройного» ключа (разд. 8.2.3). Пароль /johns/ и HOPKINS очевидно намекают на Джонса Гопкинса (1795–1873 гг.), американского финансиста и филантропа[9]. Соответствующими ключами являются CIPHER и GROUP, как можно видеть из входов $\alpha\beta\gamma\delta\varepsilon\zeta$ и $\iota\kappa\lambda\mu\nu$ в табл. 26.

19.6. Криптотекст—криптотекст компромисс: дублирование

> Двойное шифрование установок каждого текста было грубой ошибкой.
>
> *Гордон Уэлчман*, 1982 *г.*

Поляки выделялись высоким уровнем своих криптоаналитических способностей: они выиграли войну против России в 1920 г. с помощью криптоанализа.

Как сообщил в 1967 г. Козачук[10], один типичный случай компромисса криптотекст—криптотекст позволил с 1932 г. Польскому бюро шифров под руководством майора Лангера и его криптоаналитической службе BS 4 (Цежкий и юные математики Режевски, Ружицкий, Зыгальский) проникнуть в шифр немецкого Вермахта, радиопередачи которого в восточных прусских провинциях, зашифрованные машиной ENIGMA, давали им богатый запас криптотекстов.

Это была типичная проблема машинного шифрования с ключевой последовательностью, порожденной самой машиной (разд. 8.5). Если каждое сообщение начинается с его собственных начальных установок, то непосредственное использование наложения подавляется. Но существовало широко распространенное убеждение, что слишком трудно и чревато ошибками заранее подготавливать новые начальные установки для каждого сообщения. Это имело место не только для машины ENIGMA, но было общей проблемой рас-

[9] Университет Джонса Гопкинса в Мэриленде, США, был местом, где во время Второй мировой войны разрабатывались дистанционные взрыватели.

[10] В книге «Bitwa o taemnice: *Służby wywiadowcze Polski i Rzeszy Niemieckiej* 1922–1939 гг.», широко освещенной на Западе (ее обзор был опубликован в Геттингенском журнале в 1967 г.). Однако в 1968 г. Уатт в предисловии к книге Ирвинга «*Breach of Security*» сообщил сенсационный факт, что в 1939 г. Великобритания «получила от Польской военной разведки ключи и машины для декодирования немецких официальных военных и дипломатических шифров». Это казалось невероятным, пока в 1973 г. французский генерал Бертран, который был вовлечен в это дело, не подтвердил это.

		j a f m u z o b g p v h c i q w n d k r x s e l t y
δ	H	H A J V O B L W P C M X K D Q Y I E R Z N F T S G U
	A	A J V O B L W P C M X K D Q Y I E R Z N F T S G U H
	J	J V O B L W P C M X K D Q Y I E R Z N F T S G U H A
	V	V O B L W P C M X K D Q Y I E R Z N F T S G U H A J
λ	O	O B L W P C M X K D Q Y I E R Z N F T S G U H A J V
	B	B L W P C M X K D Q Y I E R Z N F T S G U H A J V O
	L	L W P C M X K D Q Y I E R Z N F T S G U H A J V O B
	W	W P C M X K D Q Y I E R Z N F T S G U H A J V O B L
$\gamma=\nu$	P	P C M X K D Q Y I E R Z N F T S G U H A J V O B L W
α	C	C M X K D Q Y I E R Z N F T S G U H A J V O B L W P
	M	M X K D Q Y I E R Z N F T S G U H A J V O B L W P C
	X	X K D Q Y I E R Z N F T S G U H A J V O B L W P C M
	K	K D Q Y I E R Z N F T S G U H A J V O B L W P C M X
	D	D Q Y I E R Z N F T S G U H A J V O B L W P C M X K
	Q	Q Y I E R Z N F T S G U H A J V O B L W P C M X K D
	Y	Y I E R Z N F T S G U H A J V O B L W P C M X K D Q
β	I	I E R Z N F T S G U H A J V O B L W P C M X K D Q Y
ϵ	E	E R Z N F T S G U H A J V O B L W P C M X K D Q Y I
$\zeta=\kappa$	R	R Z N F T S G U H A J V O B L W P C M X K D Q Y I E
	Z	Z N F T S G U H A J V O B L W P C M X K D Q Y I E R
	N	N F T S G U H A J V O B L W P C M X K D Q Y I E R Z
	F	F T S G U H A J V O B L W P C M X K D Q Y I E R Z N
	T	T S G U H A J V O B L W P C M X K D Q Y I E R Z N F
	S	S G U H A J V O B L W P C M X K D Q Y I E R Z N F T
ι	G	G U H A J V O B L W P C M X K D Q Y I E R Z N F T S
μ	U	U H A J V O B L W P C M X K D Q Y I E R Z N F T S G

Таблица 26. Шифровальная таблица, полученная сдвигами заглавной линии

пределения и управления ключей. Поэтому, как уже отмечалось в разд. 19.3, применялись индикаторы для «текстовых установок», или «установок сообщения». Конечно, эти установки не должны были передаваться открытым текстом, а должны были быть как-то зашифрованы.

Так появилась идея использовать для этой цели саму шифровальную машину. Специфической слабостью машины ENIGMA было то, что шифрование индикатора основывалось *исключительно* на самой машине ENIGMA. Лишь много позднее, не ранее мая 1941 г., был создан биграммный шифр для военно-морских машин ENIGMA.

Другой слабостью было то, что открытый индикатор дублировался перед шифрованием (по изложенным ниже причинам), и тем самым вводился крошечный криптотекст—криптотекст компромисс в постоянном месте, а именно в начале сообщения. Все это являлось недостатками режима работы, а не самой машины.

Легкомысленная идея о дублировании индикатора выглядит возвратом к рекомендациям для коммерческой машины ENIGMA 1924 г.

С другой стороны, надежность машины ENIGMA казалась высокой. Дневной ключ, используемый всеми машинами какой-либо «ключевой сети» в течение дня, определял, помимо штепсельного соединения, порядок (а позднее также и выбор) трех роторов, внутренние установки каждого ротора и основные установки роторов. Соответствующая используемой машине произвольно выбранная трехбуквенная группа, т. е. открытый индикатор, называемый в Блетчли Парк текстовыми установками, дублировалась и шифровалась, и результирующая 6-буквенная группа — зашифрованный продублированный индикатор — передавалась сразу же после (открытой) преамбулы, включающей сигнал вызова, время отправления и количество букв в криптотексте. Затем следовало сообщение, зашифрованное с использованием переданных установок (индикатора). Законный получатель сначала расшифровывал своим дневным ключом зашифрованный индикатор, проверял, правильно ли он был дублирован, и в результате находил установки для дешифрования собственно криптотекста. Подобная процедура шифрования применялась до 15 сентября 1938 г.

Ослабление надежности собственно криптотекста, вызванное использованием индикатора, принималось немцами без колебаний, поскольку индикатор был хорошо защищен машиной ENIGMA, которая рассматривалась ими как недешифруемая. Никто не заметил, что это был логически порочный круг. Как бы то ни было, ENIGMA представлялась как машина с достаточно высокой комбинаторной сложностью, и когда после 15 июля 1928 г. в эфире стало появляться все больше и больше немецких радиосигналов, зашифрованных ENIGMA, Польское бюро шифров сначала оказалось не в состоянии взломать ее, несмотря на то, что было хорошо осведомлено о коммерческой машине ENIGMA (которая с 1926 г. была представлена на открытом рынке). (Разумеется, внутренняя электропроводка роторов в военной машине ENIGMA 1930 г. отличалась от коммерческой версии.)

Причиной дублирования открытого индикатора, «удвоенного шифрования текстовых установок» было частое нарушение радиосигналов посторонним шумом. А если бы зашифрованный индикатор был испорчен в течение передачи, это послужило бы причиной бессмыслицы при законной расшифровке криптотекста, со всеми рисками повторной передачи. Даже речи не может быть о повторной передаче криптосообщения ради обнаружении ошибки. Таким образом, возможность обнаружения ошибки была ограничена открытым индикатором. Но это приводит к гораздо более опасному компромиссу, которого можно было избежать. И вот наконец 1 мая 1940 г. от дублирования отказались, а сообщения, зашифрованные ENIGMA, продолжались.

Как вообще Польское бюро шифров узнало об индикаторе? Было обнаружено, что два сообщения, начинающиеся одной и той же 6-буквенной группой, имели высочайший индекс совпадения: *Kappa* была близка к κ_d. Таким образом, дешифровальщики использовали наложение. Но тогда 6-буквенная группа определила начальную позицию роторов — другими словами, это был

индикатор. Но поскольку было ясно, что он не был открытым, то он мог быть только каким-то образом зашифрован.

19.6.1. Франция I. Неизвестно, имели ли поляки причины поверить, что немцы настолько глупы, что используют саму машину ENIGMA для зашифрования индикатора. Как бы то ни было, они получили помощь от своих французских друзей в 1931 г.

Французский шпион Шмидт с кодовым именем ASCHE (Ашé), работавший до 1938 г. в шифровальном отделе немецкого Министерства обороны, который позднее был раскрыт и казнен в июле 1943 г., получил в октябре 1931 г. руководства по использованию ENIGMA, по процедуре шифрования и даже дневной ключ на сентябрь и октябрь 1932 г. (включая установочное кольцо и установки коммутатора на эти два месяца). Он передал эти материалы французскому резиденту, тогда майору (позднее генералу) Бертрану («Болеку»), который, в свою очередь, направил их Польскому бюро шифров.

19.6.2. Польша I. Юный помощник Цежкого, талантливый Мариан Режевски[11] (1905–1980 гг.) теперь должен был выяснить, как это можно было использовать. Согласно сообщению Козачука в 1984 г., это происходило следующим образом.

19.6.2.1. Он уже угадал, что каждое сообщение начиналось с 6 букв зашифрованного удвоенного индикатора. Пусть P_1, P_2, ..., P_6 обозначают перестановки, действующие на 1-ю, 2-ю, ..., 6-ю буквы открытого текста, при условии, что шифрование начинается с одной и той же основной кольцевой установки.

Из $aP_i = X$ и $aP_{i+3} = Y$ $(i = 1, 2, 3)$ вытекает, что $XP_i^{-1}P_{i+3} = Y$. Собственно взаимообратный характер машины ENIGMA был известен. Знание символов X и Y, стоящих на 1-й или 4-й или 2-й и 5-й или 3-й и 6-й позициях криптотекста, таким образом, накладывают условия на три произведения P_iP_{i+3} неизвестных перестановок P_1, P_2, P_3, P_4, P_5, P_6.

На рис. 150 представлено 65 зашифрованных удвоенных индикаторов. Они показывают, что для P_1P_4 символ a переходит в себя (1.), аналогично, символ s переходит в себя (35.), тогда как символ b переходит в c и обратно (2., 4.), символ r переходит в w и обратно (30., 53.). Для остальных же символов оказывается, что они принадлежат циклам

(d v p f k x g z y o) (5., 49., 27., 7., 17., 57., 8., 63., 61., 26)

и

(c i j m u n q l h t) (6., 12., 15., 21., 48., 23., 28., 19., 9., 45).

[11] Интуитивные способности Режевски иллюстрируются следующим эпизодом. В коммерческой машине ENIGMA контакты на входном кольце принадлежат буквам в том же порядке, что и на коммутаторе. В этом, кажется, и состоит ее отличие от военной ENIGMA. Режевски сказал себе: «Немцы полагаются на порядок», — и испытал алфавитный порядок (см. разд. 7.3.2) — и это был именно он. Когда в июле 1939 г. об этом решении рассказали Кноксу, который долго мучился над этой проблемой, тот, по сообщению Режевски, пришел в бешенство, когда узнал, как это было просто.

1. AUQAMN	14. IND JHU	27. PVJ FEG	40. SJM SPO	53. WTMRAO
2. BNHCHL	15. JWF MIC	28. QGALYB	41. SJM SPO	54. WTMRAO
3. BCTCGJ	16. JWF MIC	29. QGALYB	42. SJM SPO	55. WTMRAO
4. CIK BZT	17. KHB XJV	30. RJL WPX	43. SUG SMF	56. WKI RKK
5. DDBVDV	18. KHB XJV	31. RJL WPX	44. SUG SMF	57. XRS GNM
6. EJP IPS	19. LDR HDE	32. RJL WPX	45. TMNEBY	58. XRS GNM
7. FBRKLE	20. LDR HDE	33. RJL WPX	46. TMNEBY	59. XOI GUK
8. GPBZSV	21. MAWUXP	34. RFC WQQ	47. TAA EXB	60. XYWGCP
9. HNOTHD	22. MAWUXP	35. SYX SCW	48. USE NWH	61. YPC OSQ
10. HNOTHD	23. NXD QTU	36. SYX SCW	49. VII PZK	62. YPC OSQ
11. HXVTTI	24. NXD QTU	37. SYX SCW	50. VII PZK	63. ZZY YRA
12. IKG JKF	25. NLU QFZ	38. SYX SCW	51. VQZ PVR	64. ZEF YOC
13. IKG JKF	26. OBU DLZ	39. SYX SCW	52. VQZ PVR	65. ZSJ YWG

Рис. 150. Шестьдесят пять наблюдаемых зашифрованных удвоенных индикаторов одного и того же дневного ключа

Короче говоря, P_1P_4 имеет два 1-цикла, два 2-цикла и два цикла из 10 символов, и поскольку это охватывает все 26 символов, то P_1P_4 полностью определено. Цикловое определение полно, если каждый символ встречается по крайней мере один раз на первой, второй и третьей позиции; как правило, это требует от 50 до 100 сообщений — такое количество было, конечно, результатом напряженного дня учебных маневров.

Для P_2P_5 и P_3P_6 работа аналогична. В итоге получается успешный результат:

$$P_1P_4 = (\mathrm{a})(\mathrm{s})(\mathrm{bc})(\mathrm{rw})(\mathrm{dvpfkxgzyo})(\mathrm{eijmunqlht}),$$
$$P_2P_5 = (\mathrm{axt})(\mathrm{blfqveoum})(\mathrm{cgy})(\mathrm{d})(\mathrm{hjpswizrn})(\mathrm{k}),$$
$$P_3P_6 = (\mathrm{abviktjgfcqny})(\mathrm{duzrehlxwpsmo})$$

1-циклы («female» — женщина, самка, *англ.*)[12] играют особую роль: поскольку каждая из перестановок P_1, P_2, ..., P_6, ввиду их взаимообратного характера, состоит всего из двух циклов, то равенство $P_iP_{i+3}x = x$ означает, что имеется символ y такой, что $P_ix = y$, а $P_{i+3}y = x$, т. е. как P_i, так и P_{i+3} содержат один и тот же цикл (xy). В данном примере обе перестановки P_1 и P_4 содержат 2-цикл (a s).

Одна теорема теории групп о произведениях собственных взаимообратных перестановок утверждает, что циклы произведения P_iP_{i+3} встречаются парами равной длины, если

P_i содержит 2-циклы (x_1y_1), (x_2y_2), ..., $(x_\mu y_\mu)$ и
P_{i+3} содержит 2-циклы (y_1x_2), (y_2x_3), ..., $(y_\mu x_1)$, то

[12] В жаргоне Блетчли Парк термин «female» произошел от польского каламбура te same (то же самое, *пол.*) — samiczka (самочка, *пол.*). Однако, большинство людей в Блетчли Парк не знали этого польского происхождения термина «female», и находили свои собственные объяснения, вроде «female skrew» (гайка, *англ.*) для отверстия с резьбой.

P_iP_{i+3} содержит μ-циклы $(x_1x_2 \ldots x_\mu)$ $(y_\mu y_{\mu-1} \ldots y_1)$.

Таким образом, если один из циклов произведения P_iP_{i+3} записать в обратном порядке ($\leftarrow$) под другим циклом той же длины, то 2-циклы перестановок P_i можно читать по вертикали, если циклы правильно согласованы. Остается задача — согласовать циклы. Она может быть решена перебором — для вышеприведенных P_1P_4 — из 2×10 испытаний, а для вышеприведенных P_2P_5 — из 3×9 испытаний.

19.6.2.2. Но Мариан Режевски нашел более короткий путь. Он заметил, что действительно применяемые зашифрованные индикаторы показывают отклонения от равномерного распределения, которые вероятно означают, что немецкие шифровальщики, подобно большинству людей, играющих в лотерею, были неспособны выбрать начальные установки действительно случайными. Таким образом, Режевски направил свое внимание на выделяющиеся шаблоны, и оказался прав. В действительности, немецкие инструкции безопасности в этом пункте были не слишком ясными, и немецкий офицер, который отдавал приказ брать в качестве установки конечную позицию роторов в предыдущем сообщении, мог доказать, что он был уверен, что установки менялись после каждого сообщения. Таким образом, существовала общая практика использования даже стереотипных 3-буквенных групп вроде /aaa/, /bbb/, /sss/. Когда весной 1933 г. даже простое повторение букв было недвусмысленно запрещено, было уже слишком поздно. Поляки уже имели свои входы в ENIGMA. Позднее возникла плохая привычка использования соседних букв на клавиатуре: /qwe/, /asd/ (горизонтально), /qay/, /cde/ (вертикально) и т. д.

Частотный аргумент Режевски состоял в том, что наиболее часто встречающийся зашифрованный индикатор SYX SCW, встречающийся пять раз (35.–39.), должен соответствовать наиболее выделяющемуся шаблону. Имелось еще некоторое число аналогичных индикаторов, которые требовали проверки. Допустим, проверяется открытый индикатор aaa. Это соответствует в P_1 наличию 2-цикла (a s), в P_2 это дает 2-цикл (a y), в P_3 — 2-цикл (a x); это соответствует в P_4 2-циклу (a s), в P_5 это дает 2-цикл (a c), в P_6 дает 2-цикл (a w). Таким образом, для P_3 и P_6 фаза двух циклов

$$\begin{array}{c} \downarrow \\ \rightarrow (\mathrm{a\ b\ v\ i\ k\ t\ j\ g\ f\ c\ q\ n\ y}) \\ \leftarrow (\mathrm{x\ l\ h\ e\ r\ z\ u\ d\ o\ m\ s\ h\ w}) \end{array}$$

уже определена; с помощью зигзага можно подсчитать 2-циклы из P_3 и P_6, начинающиеся с (a x):

$$P_3 = (\mathrm{ax})(\mathrm{bl})(\mathrm{vh})(\mathrm{il})(\mathrm{kr})(\mathrm{tz})(\mathrm{ju})(\mathrm{gd})(\mathrm{fo})(\mathrm{cm})(\mathrm{qs})(\mathrm{np})(\mathrm{yw}),$$
$$P_6 = (\mathrm{xb})(\mathrm{lv})(\mathrm{hi})(\mathrm{ek})(\mathrm{rt})(\mathrm{zj})(\mathrm{ug})(\mathrm{df})(\mathrm{oc})(\mathrm{mq})(\mathrm{sn})(\mathrm{py})(\mathrm{wa}).$$

P_3 содержит среди других 2-цикл (q s). Таким образом, открытый индикатор для AUQ AMN (1.) имеет шаблон **s; так как P_1 содержит среди других

2-цикл (a s), шаблон имеет вид s*s. Если теперь угадано, что открытый индикатор для AUQ AMN читается sss, то в P_2 кроме (a y) определяется также (s u). Таким образом, фаза для циклов P_2 и P_5 также полностью определена:

$$\begin{array}{c} \downarrow \qquad\qquad \downarrow \\ \rightarrow (\mathrm{a\ x\ t})\ (\mathrm{b\ l\ f\ q\ v\ e\ o\ u\ m})\ (\mathrm{d}) \\ \leftarrow (\mathrm{y\ g\ c})\ (\mathrm{j\ n\ h\ r\ z\ i\ w\ s\ p})\ (\mathrm{k}). \end{array}$$

В зигзаге можно подсчитать 2-циклы из P_3 и P_6, начинающиеся с (a y):

$$P_2 = (\mathrm{ay})(\mathrm{xg})(\mathrm{tc})(\mathrm{bj})(\mathrm{ln})(\mathrm{fh})(\mathrm{qr})(\mathrm{vz})(\mathrm{ei})(\mathrm{ow})(\mathrm{us})(\mathrm{mp})(\mathrm{dk}),$$
$$P_5 = (\mathrm{yx})(\mathrm{gt})(\mathrm{ca})(\mathrm{jl})(\mathrm{nf})(\mathrm{hq})(\mathrm{rv})(\mathrm{ze})(\mathrm{io})(\mathrm{wn})(\mathrm{sm})(\mathrm{pb})(\mathrm{kd}).$$

Другим часто встречающимся зашифрованным индикатором был RJL WPX, встречающийся 4 раза (30.–33.). Соответствующий открытый индикатор имеет шаблон *bb. P_1 может иметь лишь 2-циклы (r b) и (r c). В первом случае с более вероятным открытым индикатором bbb: P_1 содержит 2-цикл (b r), P_4 — 2-цикл (r c). Для согласования 10-циклов можно использовать другой индикатор, скажем, LDR HDE (19.–20.). Так как P_3 и P_6 содержат циклы (r k) и (k e), P_2 и P_5 содержат (d k) и (k d), то открытый индикатор имеет шаблон *kk. Это снова наводит на мысль о стереотипе kkk с тем результатом, что P_1 и P_4 содержат 2-циклы (l k) и (k h). Таким образом, фаза для циклов из P_1 и P_4 полностью определена:

$$\begin{array}{c} \downarrow \qquad\qquad \downarrow \\ \rightarrow (\mathrm{a})\ (\mathrm{b\ c})\ (\mathrm{d\ v\ p\ f\ k\ x\ g\ z\ y\ o}) \\ \leftarrow (\mathrm{s})\ (\mathrm{r\ w})\ (\mathrm{i\ e\ t\ h\ l\ q\ n\ u\ m\ j}) \end{array}$$

и таким образом

$$P_1 = (\mathrm{as})(\mathrm{br})(\mathrm{cw})(\mathrm{di})(\mathrm{ve})(\mathrm{pt})(\mathrm{fh})(\mathrm{kl})(\mathrm{xq})(\mathrm{gn})(\mathrm{zu})(\mathrm{ym})(\mathrm{oj}),$$
$$P_4 = (\mathrm{sa})(\mathrm{rc})(\mathrm{wb})(\mathrm{iv})(\mathrm{ep})(\mathrm{tf})(\mathrm{hk})(\mathrm{lx})(\mathrm{qg})(\mathrm{nz})(\mathrm{uy})(\mathrm{mo})(\mathrm{jd}).$$

Вместе первые три перестановки имеют следующий вид:

$$P_1 = (\mathrm{as})(\mathrm{br})(\mathrm{cw})(\mathrm{di})(\mathrm{ev})(\mathrm{fh})(\mathrm{gn})(\mathrm{jo})(\mathrm{kl})(\mathrm{my})(\mathrm{pt})(\mathrm{qx})(\mathrm{uz})$$
$$P_2 = (\mathrm{ay})(\mathrm{bj})(\mathrm{ct})(\mathrm{dk})(\mathrm{ci})(\mathrm{fh})(\mathrm{gx})(\mathrm{ln})(\mathrm{mp})(\mathrm{ow})(\mathrm{qr})(\mathrm{su})(\mathrm{vz})$$
$$P_3 = (\mathrm{ax})(\mathrm{bl})(\mathrm{cm})(\mathrm{dg})(\mathrm{ei})(\mathrm{fo})(\mathrm{hv})(\mathrm{ju})(\mathrm{kr})(\mathrm{np})(\mathrm{qs})(\mathrm{tz})(\mathrm{wy})$$

19.6.2.3. Теперь возможно восстановление всех открытых индикаторов, использованных в тот день (рис. 151). Плохие привычки шифровальщиков машины ENIGMA очевидны. Во-первых, применение стереотипов, что приводит к частому использованию одних и тех же индикаторов, чего не должно было случиться ни в коем случае. Во-вторых, взгляд на клавиатуру ENIGMA (рис. 152) показывает, что всего два индикатора из сорока, а именно abc и uvw не являются клавиатурными стереотипами (вместо этого они являются алфавитными стереотипами). Ни шифровальщики, ни офицеры связи даже вообразить не могли, что мирные тренировочные передачи с наивно придуманным боевым сценарием могут выдать так много секретов машины ENIGMA.

AUQ	AMN	: sss	IKG	JKF	: ddd	QGA	LYB	: xxx	VQZ	PVR	: ert
BNH	CHL	: rfv	IND	JHU	: dfg	RJL	WPX	: bbb	WTM	RAO	: ccc
BCT	CGJ	: rtz	JWF	MIC	: ooo	RFC	WQQ	: bnm	WKI	RKK	: cde
CIK	BZT	: wer	KHB	XJV	: lll	SYX	SCW	: aaa	XRS	GNM	: qqq
DDB	VDV	: ikl	LDR	HDE	: kkk	SJM	SPO	: abc	XOI	GUK	: qwe
EJP	IPS	: vbn	MAW	UXP	: yyy	SUG	SMF	: asd	XYW	GCP	: qay
FBR	KLE	: hjk	NXD	QTU	: ggg	TMN	EBY	: ppp	YPC	OSQ	: mmm
GPB	ZSV	: nml	NLU	QFZ	: ghj	TAA	EXB	: pyx	ZZY	YRA	: uvw
HNO	THD	: fff	OBU	DLZ	: jjj	USE	NWH	: zui	ZEF	YOC	: uio
HXV	TTI	: fgh	PVJ	FEG	: tzu	VII	PZK	: eee	ZSJ	YWG	: uuu

Рис. 151. Сорок различных дешифрованных индикаторов

19.6.2.4. Польское бюро шифров, конечно, узнало кое-что из содержания дешифрованных сигналов. Но гораздо более важным было то, что этот компромисс поставил под угрозу коммутацию роторов машины ENIGMA.

Вначале анализ индикаторов включал только первые 6 букв, это был главным образом только быстрый ротор R_N, который двигался, а два другие ротора в 20 из 26 случаев простаивали. Этого вместе с материалом от ASCHE, оказалось достаточно для Режевски, чтобы восстановить коммутации быстрого ротора, а поскольку порядок роторов в то время менялся каждый квартал (с 1936 г. — каждый месяц, позднее — ежедневно), то каждый ротор имел удовольствие быть однажды изученным Польским бюро шифров. А поскольку все открытые индикаторы одного дня были дешифрованы, то все сообщения этого дня могли быть дешифрованы с помощью точных копий ENIGMA, построенных поляками. Но что можно сказать о следующем дне?

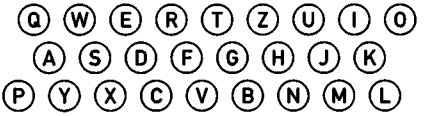


Рис. 152. Клавиатура машины ENIGMA

19.6.2.5. Восстановление основной кольцевой установки было выполнено с помощью одной теоремы из теории групп, ярко охарактеризованной Девуром, который назвал ее «теоремой, которая выиграла Вторую мировую войну». Она гласит:

Подстановки S и TST^{-1} имеют одно и то же цикловое разложение.

Поэтому Мариан Режевски и его сотрудники использовали тот факт, что длины циклов в трех заслуживающих внимания подстановках P_iP_{i+3}, $i=1, 2, 3$, не зависят от выбора переходных контактов (и во всяком случае от кольцевой установки). Число существенно различных цикловых разложений равно числу разбиений числа $26/2=13$, которое равно 101; три такие разбиения — в приведенном выше примере — $10+2+1$, $9+3+1$ и 13 — вообще однозначно характеризуют $6\times 26\times 26\times 26\approx 10^5$ основных кольцевых установок. Режевски, вместе с Ружицким и Зыгальским, были теперь способны с

помощью копии ENIGMA создать каталог для каждого роторного порядка, содержащий разбиения на циклы для всех основных кольцевых установок. С этой целью на фабрике AVA в Варшаве было построено электромеханическое устройство, названное «циклометром». Бюро шифров окончило этот каталог в 1937 г.; теперь, чтобы найти дневной ключ, требовалось не более 10–20 минут. К несчастью для поляков, 1 ноября 1937 г. немцы изменили отражающий ротор.

19.6.2.6. Оставалась проблема нахождения кольцевых установок. Переборные испытания можно было упростить, воспользовавшись замечанием Режевски, сделанным в 1932 г. благодаря материалу ASCHE: большинство открытых текстов начинается на /апх/, где /х/ заменяет пробел между словами. Согласно предостережению Керкхоффса, следует предполагать, что машина побывала в «плохих руках», и потому глупо использовать стереотипные начала. Мы вернемся к этому тривиальному случаю компромисса открытый текст — криптотекст в разд. 19.7.

19.6.2.7. Некоторое время поляки применяли также метод, называемый metoda rusztu (метод решеток, *пол.*), который, по выражению Козачука, был «ручным и утомительным». Он был применим, пока число переходных контактов было малым (шесть до 1 октября 1936 г.), и служил для определения кольцевой установки быстрого ротора. В открытой литературе других подробностей нет.

19.6.3. Польша II. Все эти успехи оказались возможными только по причине взаимообратного[13)] характера шифрующего ротора машины ENIGMA; оказалось, что отражающий ротор Шербиуса и Корна был грандиозной иллюзорной сложностью. Введение таблиц биграмм в 1937 г. прекратило чтение сообщений, зашифрованных военно-морской ENIGMA. В 1938 г. ситуация ухудшилась еще больше. Германия изменила процедуру шифрования 15 сентября, а с 15 декабря ввела четвертый и пятый роторы, которые теперь давали $60 = 5 \times 4 \times 3$ вместо прежних $6 = 3 \times 2 \times 1$ возможных роторных порядков.

19.6.3.1. Поляки должны были найти коммутацию новых роторов быстро, и им повезло. Среди сообщений, которые они регулярно дешифровали, были сигналы от Службы безопасности (СД) — службы разведки нацистской партии. СД не изменила процедуру шифрования, только ввела дополнительные два ротора в декабре 1938 г. Эти роторы время от времени переходили в положение быстрого ротора, и их коммутация могла быть восстановлена тем же способом, что и раньше с первоначальными тремя роторами.

Применение двух методов, один из которых, возможно, был скомпрометирован, явилось серьезной ошибкой.

Приведем историю о том, как Польское бюро шифров сумело прочитать сообщения СД. Офицеры СД подозревали всех, и потому кодировали свои сообщения вручную перед тем как давать их оператору ENIGMA для перешифрования. Поляки, дешифруя все сообщения, получаемые с помощью

[13)] Шифровальные машины Бориса Хагелина, которые тоже были взаимообратными, не страдали от *этого* дефекта, хотя все же М-209 была взломана немцами, начиная с 1942 г. на североафриканском театре боевых действий.

ENIGMA, получали бессмысленные тексты и сначала думали, что криптотекст был зашифрован другой системой. Но однажды в 1937 г. было прочитано трехбуквенное слово /ein/ (один, *нем.*). Это могло означать только, что некоторая группа открытого текста была ошибочно перепутана с кодом; вероятно, номер 1 не был транскрибирован, и оператор ENIGMA не нашел ничего лучше, чем послать /ein/. Тогда поляки легко сумели взломать простой ручной шифр.

19.6.3.2. Новая процедура шифрования, которая действовала до конца апреля 1940 г., не использовала одни и те же основные установки для всех сообщений текущего дня (как это делалось раньше), а для каждого сообщения выбирались произвольные основные установки, которые должны были предшествовать передаваемому сообщению. С этими выбранными основными установками (в Блетчли Парк их называли «индикаторными установками»), как и раньше, случайно выбранный открытый индикатор дублировался и шифровался, и затем использовался в качестве установок для шифрования основного текста.

Приведем пример. Для сообщения, начинающегося (после открытой преамбулы) с RTJWA HWIK..., rtj — есть основная установка, WAH WIK — зашифрованный удвоенный индикатор, причем, зашифрованный с помощью rtj. В дальнейшем в такой ситуации мы будем писать rtj | WAH WIK.

Законный получатель использует основную установку rtj, чтобы из WAH WIK найти открытую индикаторную пару (которая имеет шаблон *123123*); и затем расшифровывает криптотекст, используя ее первые три буквы (настоящий индикатор) как установку для расшифровки основного текста.

Пока кольцевые установки и порядок роторов не попали к противнику, он ничего не может поделать с выставленными напоказ основными установками. Пространство поиска содержит 1054560 возможностей (26^3 кольцевых установок, 6 порядков роторов, а с декабря 1938 г. еще и 10 выборов трех роторов из пяти).

19.6.3.3. Методы Режевски и его друзей, использованные до сих пор, больше не работали, так как они основывались на кратном использовании одной и той же основной установки на весь день. Но немцы, хоть это и невероятно, сохраняли дублирование открытого индикатора[14)], и тем самым позволяли атаковать поиском шаблона (т. е. шаблона *123123* в известной позиции). Этот метод мог бы работать также и раньше, но теперь не было другого выбора, кроме как преодолевать трудности резко возросшего объема работы. Поэтому поляки подумали о механизации. В октябре 1938 г. Режевски заказал шесть машин, каждая из которых моделировала один из 6 роторных порядков, на фабрике AVA, и опробовал на каждой из них 17576 кольцевых установок. На это уходило не более 110 минут.

«Правильные» кольцевые установки были найдены с помощью 1-циклов следующим способом. Была построена машина из трех пар роторных ком-

[14)] Двойное шифрование текстовых установок продолжалось, согласно Твинну, для 4-роторной машины ENIGMA Абвера до конца войны.

плектов ENIGMA. В каждой паре позиции всех роторов сдвигались на три; позиция роторных комплектов первой пары сдвигалась на единицу относительно позиции роторных комплектов второй пары, которая, в свою очередь, была сдвинута на единицу относительно позиции роторных комплектов третьей пары.

Как только накопилось достаточное количество материала, чтобы обеспечить наличие таких трех зашифрованных удвоенных индикаторов, что один и тот же символ появлялся один раз на первой и четвертой, один раз на второй и пятой и один раз на третьей и шестой позициях, — как буква W в таблице

```
rtj | WAH WIK
dqx | DWJ MWR
hpl | RAW KTW
```

и таким же способом, как в разд. 19.6.2.1, обнаруживался 1-цикл, («неподвижная точка»), дающий надежду на возможную атаку (см. рис. 153).

Машина стартовала с трех начальных установок rtj, dqx, hpl, и пробный символ W загружался несколько раз, до тех пор пока в каждой из трех пар

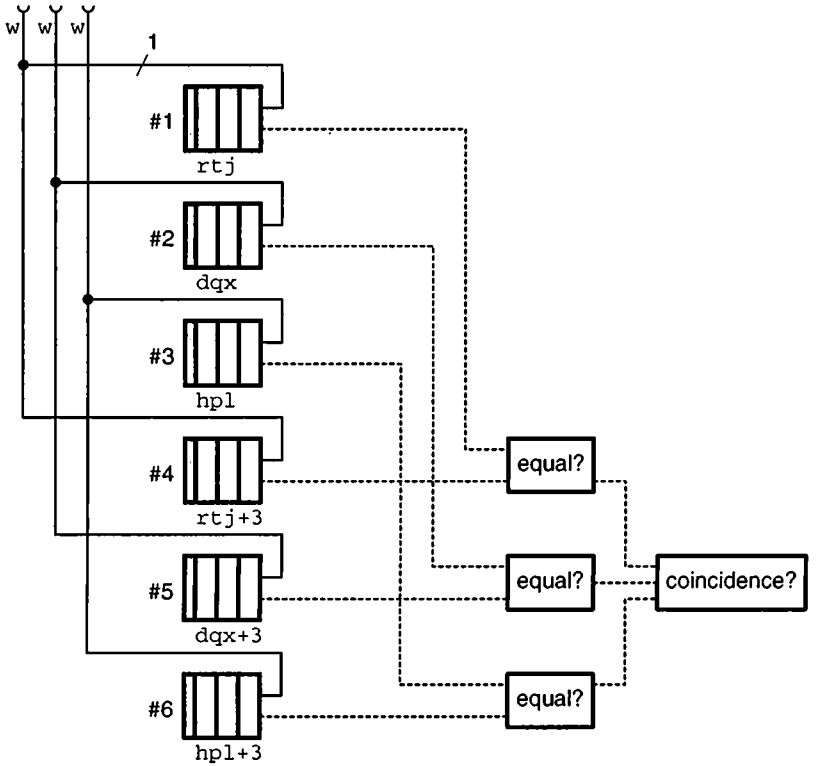


Рис. 153. Абстрактная схема польской бомбы (версия 1938 г.)

не встретился один и тот же символ, т. е. пока не был найден шаблон *123123*. Такое совпадение, как спусковой крючок, приводило в действие простую схему, останавливающую машину; это и привело поляков к названию *bomba*[15]. Конечно, случались и промахи.

Если таким способом удавалось обнаружить кольцевые установки, то сравнение зашифрованных двойных индикаторов с зашифрованиями, производимыми копией машины ENIGMA (с найденным кольцевыми установками), позволяло восстановить переходные контакты. Таким способом, все сообщения данного дня (позднее — восьмичасового периода) из одной и той же ключевой сети (за время войны существовало до 120 ключевых сетей) могли быть дешифрованы.

Бомба была очень чувствительна к переходным коммутациям, и этот метод работал только когда опробуемый символ (в примере буква W) не был «подключен». Вероятность, что это случится, была примерно 50%, пока использовались пять из восьми штепселей. Если оказывались допустимыми три разные неподвижные точки (как сначала думали англичане), то эта вероятность падала до 12.5%.

19.6.3.4. Другой тип механизации примерно в то же время был развит Хенриком Зыгальским (1907–1978 гг.), а именно «перфокарточный каталог» для определения дневного ключа с 10–12 неподвижными точками. Для всех шести роторных упорядочений данных роторов R_L, R_M и R_N было подсчитано, будет ли для перестановок P_1P_4, P_2P_5 и P_3P_6, полученных при некоторых основных установках $\langle R_L\rangle\ \langle R_M\rangle\ \langle R_N\rangle$ роторов возможна хотя бы одна неподвижная точка. Для каждой из 26 букв $\langle R_L\rangle$ это было записано в $\langle R_M\rangle \times \langle R_N\rangle$ матрицу с помощью перфорированных отверстий («female») (рис. 154); грубо говоря, 40% квадратов на каждом листе содержали отверстия. Для того, чтобы сделать возможным полное покрытие, были использованы листы из 51×51 полей с продублированными горизонталями и вертикалями. При наложении листы выстраиваются в соответствии с их основными установками $\langle R_M\rangle\ \langle R_N\rangle$, и кольцевые установки определяются, как правило, однозначно, как только в распоряжении оказывается от 10 до 12 неподвижных точек. Самое важное, что этот метод был нечувствителен к применению переходных контактов и сохранил свою пользу даже после введения 10 штепселей после 19 августа 1939 г. — до тех пор, пока сохранялось двойное шифрование.

Немцы всегда старались с помощью соответствующей надежности шифрования не оказаться жертвами наложения Керкхоффса — и оказались жертвами тривиальной слабости, дублирования индикатора.

19.6.4. Великобритания. На встрече 9 января 1939 г. в Париже польский подполковник Лангер пополнил свои французские связи контактами с их британскими коллегами. С возрастанием опасности войны требовалось более тесное сотрудничество. Результатом этих контактов стала встреча 25 июля 1939 г. в Варшаве Кнокса, ведущего английского криптоаналитика из Ми-

[15] Согласно Лисицкому, она была так названа Ружицким по аналогии с названием десерта «ice-cream bombe», за которым ему и его друзьям пришла идея такой машины.

нистерства иностранных дел, и его босса Деннистона, руководителя государственной Школы кодов и шифров — с британской стороны, майора Бертрана и капитана Бракенье — с французской. Польскую же сторону представляли Цежкий, Лангер и главный шеф полковник Майер. Режевски, Ружицкий и Зыгальский гордо представили все свои результаты в местечке Пыры, южнее Варшавы. По этому случаю французы, так же, как и англичане, получили польские копии ENIGMA со всеми пятью роторами.

Со времени кризиса, который привел к Мюнхенской конференции в сентябре 1938 г., англичане настойчиво думали, куда бы эвакуировать свою криптологическую службу, известную как «комната 47» Министерства иностранных дел, размещенную по адресу Вестминстер, Бродвей 56 (Уайтхолл). Наконец, такое место было найдено в Блетчли Парк, находящемся в семи милях к северу от Лондона. И перед тем, как в 1939 г. разразилась война, там уже была размещена и усилена Школа кодов и шифров (GC&CS). Одной из ее задач было дешифрование ENIGMA. Этим занималась группа, возглавляемая Кноксом и Тьюрингом, которая, подобно полякам, использовала перфорированные листы (которые здесь назывались «простынями Джеффри» в честь наблюдавшего за их приготовлением Джеффри (умер в 1940 г.) — Кнокс узнал о них на встрече в Пыры. То же самое имело место и в отношении польской *bomby*, которой англичане тоже занимались.

Стречи[16] в разное время пытался наладить контакт с юным Тьюрингом, который уже имел репутацию логика и с детских лет интересовался криптологией и GC&CS. Глава криптоаналитической службы Кнокс был классическим ученым, который в 1915 г. предпочел работе в Кембридже комнату 40 Адмиралтейства, и уже проводил опыты с коммерческой машиной ENIGMA, которая применялась в Италии. 4 сентября 1939 г., на второй день Второй мировой войны Тьюринг прибыл в Блетчли Парк. Он работал над дальнейшим развитием польской *bomby*, и в этом сотрудничал с Уэлчманом (1906–1985 гг.), который тоже прибыл 4 сентября, Тьюринг экспериментировал с релейными схемами (разд. 5.7.3), и следовательно, интересовался криптологией не только теоретически. Его контакты с GC&CS могли начаться в 1936 г.

В январе 1940 г. в Блетчли Парк в первый раз взломали ключ машины ENIGMA немецких Военно-воздушных сил, а именно ключ RED для 6 января, и продолжали это делать впоследствии.

19.6.4.1. Когда в середине января 1940 г. Тьюринг встречал Режевски (который после оккупации Польши немцами бежал во Францию) в Гретц-Арманвильерсе, его, по словам Режевски, очень заинтересовали идеи поляков о возможных способах восстановления коммутаций роторов ENIGMA. К тому времени он уже подошел к своим собственным идеям (см. разд. 19.7), но конечно еще не имел понятия, как далеко они его заведут. Было вполне естественно, что Тьюринг пытался улучшить польскую *bomby*, чтобы сделать ее нечувствительной к перемене коммутаций, подобно листам Зыгальского. Ан-

[16] Оливер Стречи, муж феминистки Рей Стречи, отец известного специалиста в области информатики Кристофера Стречи, брат писателя Литтона Стречи, заменил в 1941 г. в канадской службе бывшего майора армии США Ярдли.

гличане, как и поляки, боялись, что с дальнейшим уменьшением неподвижных точек их методы станут бесполезными.

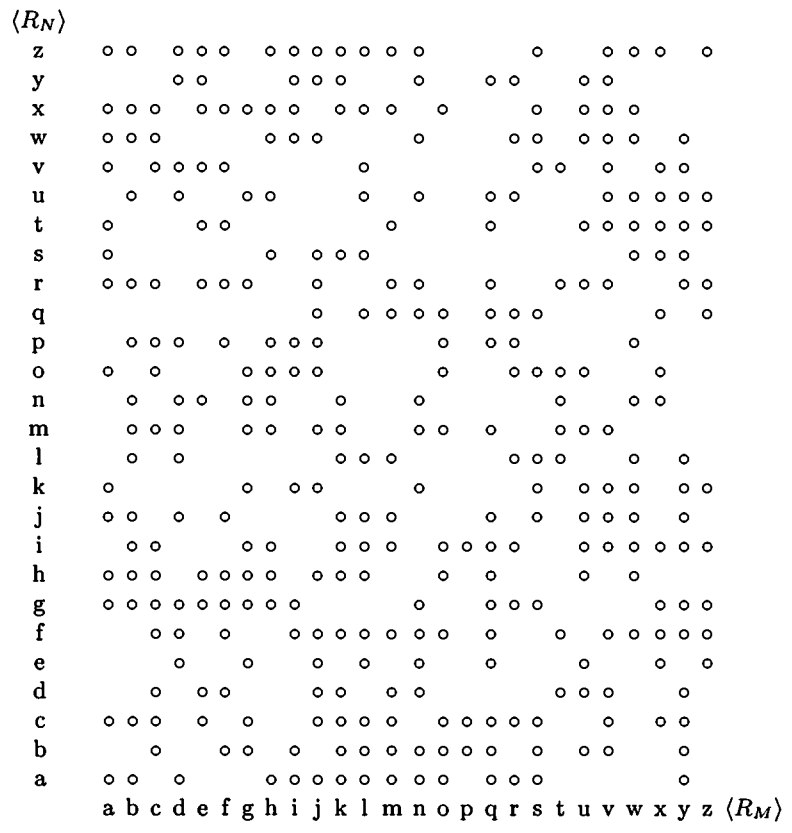


Рис. 154. Страница K_{14}^{413} Зыгальского для роторного порядка IV–I–III, $\langle R_L \rangle = k$; отверстия показывают возможность 1-циклов для P_1P_4 при основных установках $\langle R_L \rangle$ $\langle R_M \rangle$ $\langle R_N \rangle$

При конструировании *bomby* поляки, по существу, следовали электрической схеме ENIGMA, в которой существовал единственный путь — от клавиши, которая нажимается для ввода испытуемой буквы открытого текста, через отражатель до электролампы, которая отмечает соответствующую букву криптотекста на выходе. Для шифрования это было достаточно практично. Но Тьюринг хотел в версии *bomby* 1939 г. испытывать параллельно все 26 букв, чтобы видеть, какие выходы они могут иметь; это позволило бы, как вспоминала Джоан Мюррей, урожденная Кларк, произвести «одновременный беглый просмотр» всех 26 возможностей испытуемой буквы. Таким образом, Тьюринг думал о замене роторов роторами «Тьюринга», каждый из которых имеет на входной и выходной стороне два концентрических кольца контактов и, имитирует ротор машины ENIGMA. Два кольца контактов имеет и отражатель.

Вся кострукция связана 26-жильным кабелем с двухсторонним скрамблером (коммутатором). В результате такая модификация моделирует классическую подстановку, производимую машиной ENIGMA:

$$P_i = S_i U S_i^{-1}, \qquad i = 1, 2, \ldots, 26^3.$$

Ввиду взаимообратного характера, скрамблер должен быть симметричным относительно входа-выхода.

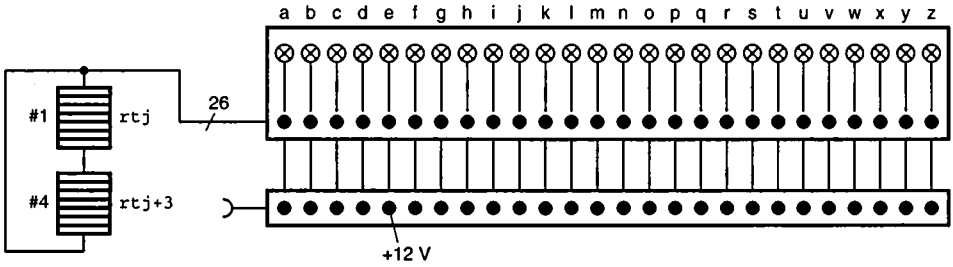


Рис. 155. Гипотетическая версия Тьюринга польской *bomby* (с «одновременным просмотром»). Диаграмма одного из трех циклов

В такой версии Тьюринга польская *bomba* равносильна трем замкнутым циклам, каждый из которых строится из пары удвоенных скрамблеров, один такой цикл показан на рис. 155. Тьюринг таким образом умудрялся отслаивать перешифрование с помощью коммутации, механически. И он выяснил, что 1-циклы (называемые «female») в листах Зыгальского, будучи неподвижными точками некоторого отображения, могут быть определены с помощью итерационного процесса с обратной связью. Если этот процесс расходится, то это указывает, что кольцевая установка в рассматриваемом случае не имеет неподвижной точки; если же он сходится, то это дает неподвижную точку. Логик Тьюринг, хорошо знакомый с методом *reductio ad absurdum*, таким образом, обращается к общему принципу обратной связи.

Технически различие сходимости и расходимости производилось «тестовым регистром», присоединенным 26-жильным кабелем к шине обратной связи (на рис. 155 от #4 до #1). На провод, соответствующий испытываемой букве (в нашем примере W), подается напряжение. В случае расходимости загораются все лампочки испытываемого регистра. В случае же сходимости цикл обратной связи неподвижной точки оказывается электрически изолированным от остальных проводов; поэтому либо загорается лишь одна лампа (лампа, принадлежащая W), либо загораются все лампы кроме этой одной, — в зависимости от того, правильно или нет выбрана коммутация.

Батарея скрамблеров должна была действовать одновременно. Таким образом, Тьюринг мог смоделировать польскую *bomba*. Но дальнейшее развитие получил другой, гораздо более общий путь (разд. 19.7).

В последнем квартале 1939 г. устройство Тьюринга было усовершенствовано настолько, что Блетчли Парк было разрешено просить Британскую ком-

панию табуляционных машин в Летчворте построить машину, которая также была названа BOMBE (бомба, *англ.*). Ответственным инженером проекта, над которым работало 12 человек, был Кин. Машина была построена в марте 1940 г. Позже Ким с таким же успехом построил 4-х роторную BOMBE MAMMOTH.

Кстати, Уэлчмен, когда он был еще новичком, в 1939 г. независимо пришел к тем же заключениям, хотя вначале он не был включен в группу дешифрования ENIGMA. Он также переоткрыл листы Зыгальского, не зная, что Джеффрис работает с ними в соседнем здании. Об идеях Тьюринга он ничего не знал.

19.6.4.2. Перед встречей в Пыры Тьюринг уже мог думать о применении вероятных слов для взлома ENIGMA, однако после этой встречи, услышав об аппарате *bomba*, он стал думать об автоматизации этого метода. Главным преимуществом проектируемого устройства было, по мнению Тьюринга, то, что он не только находил порядок роторов, подобно листам Зыгальского, но также находил, по крайней мере, один штекер.

Предположительно идеи Тьюринга стимулировались опасением Кнокса, что немцы снова изменят свою процедуру шифрования и откажутся от дублирования индикатора. В этом случае, обладая значительной информацией об обычаях и стилях немцев, полученной из дешифрованных сообщений, англичане надеялись создать обратные связи, эффективно работающие с вероятными словами, широко используемыми немцами: /wettervorhersage biskaya/ (прогноз погоды Бискайского залива, *нем.*), /wettervorhersage deutsche bucht/ (...немецкой бухты) и т. д., или /obersturmbahnführer/, /obergruppenführer/ и т. д., или /keine besonderen ereignisse/ (никаких необычных происшествий). Таким образом, Блетчли Парк был подготовлен, когда 1 мая 1940 г. незадолго перед кампанией во Франции, произошло следующее изменение: армия и военно-воздушный флот прекратили дублирование индикатора и вывели польскую *bomba* и листы Зыгальского—Джеффри из действия. (О военно-морском флоте, перешифровающем индикаторы сообщений биграммными подстановками, см. в разд. 4.1.2.)

Прототип BOMBE Тьюринга, названный «Victory», был поставлен на вооружение 18 марта. Начиная с августа 1940 г. за ним последовали другие, названные «Agnes», «Jumbo», «Funf»; «Ming» был изготовлен 26 мая 1941 г. Еще об этом см. в разд. 19.7.

Ходж сказал о Тьюринге: «Именно он первым сформулировал принцип автоматического поиска логической согласованности, основанный на *вероятном слове*». Кроме того, Тьюринг создал аппарат BOMBE, являющийся в известном смысле универсальным устройством.

19.6.4.3. Однако, если метод вероятного слова не работает, то существует еще одна фундаментальная возможность наложения согласованных текстов (разд. 19.3). При этом возникает необходимость в согласовании и, как следствие, в подсчете числа повторений или совпадений. Механизм для выполнения такой работы был назван «бэнберизмом», потому что длинные листы бумаги, содержащие сообщение в коде из 26 символов были произведены в

Бэнбери, маленьком городке близ Оксфорда. Тьюринг развил для этой цели особую шкалу оценок параллелей, которая означает придание весов повторениям, где, например, два биграммных повторения оцениваются выше, чем четыре односимвольных повторения. Тьюринг использовал логарифмическую единицу [ban (бан)] (гл. 12), десятичный аналог шенноновой бинарной единицы информации [бит].; 1 [бан]= $\log_2$ 10 [бит], 1 [децибан]≈ 0.332 [бит], практическая единица, используемая в Блетчли Парк, соответствует 1 [dB] (1 [децибан]).

Тьюринг и Шеннон, очевидно, развили свои идеи независимо друг от друга, и встретились лишь примерно в конце 1942 г. «Тьюринг и я никогда не говорили о криптографии» (Клод Шеннон). Вполне возможно, что до июля 1941 г. никто в Блетчли Парк не знал ни *Kanna*, теста Фридмана (разд. 17.1), ни *Xu*, теста Кульбака (разд. 17.4). Несмотря на попытки Гуда, остается неясным, как Тьюринг нашел то, что он назвал «частотой повторений» в своем «Докладе об ENIGMA», написанном в конце лета или в начале осени 1940 г.

Банберизм был, фактически, развитием одного метода поляков («часового метода» Ежи Ружицкого). Согласование криптотекстов заключалась в определении разности фаз открытых индикаторов. По крайней мере, таким образом был обнаружен быстрый ротор R_N (начальная установка которого обозначалась третьей буквой текстовой установки). Девур и Кру построили следующий пример: из нескольких пар согласованных сообщений третья позиция зашифрованного индикатора показывает следующие наблюдаемые значения:

первое сообщение	R	E	D	T	N	M	K	B	M	N
второе сообщение	F	K	Y	Y	Q	Q	O	Y	C	K
разность фаз по mod 26	07	12	04	02	14	21	06	11	06	03

Эти данные образуют две цепи:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
R				Q			F		M						C	N			K						O
D		T		Y															B						

При скольжении открытого текста вдоль первой цепи находится взаимообратная пара (а именно (k n)) для следующего сдвига:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
	F		M						C	N			K					O	R				Q		

Используя все возможности для взаимообратных пар, получаем следующий каркас:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
	F	J	M		B				C	N		D	K	T		Y	U		O	R			Q		

который совместим со второй цепью:

												B				D	T	Y							

Таким образом, 14 начальных позиций быстрого ротора обнаружены. С помощью этого метода для роторов I, II, ..., V может быть определена позиция одного из зубцов (разд. 8.5.2), и тогда может быть определено, какой из роторов использовался в качестве быстрого. Это помогало в польской bomba и английской BOMBE сократить число роторных упорядочений, которые необходимо тестировать, с 60 до $12 = 4 \times 3$.

19.6.5. Франция II. После поражения Польши Режевски, Зыгальский и Ружицкий бежали во Францию через Румынию. В конце сентября 1939 г. они начали работать во французской группе радиоперехвата под руководством Бертрана (почтовое управление Брюно) в Шато де Виньоль около Грец-Арменвилье, 45 км южнее Парижа. До нападения Германии на Францию в мае 1940 г. группа Z работала с листами Зыгальского, сделанными в Блетчли Парк. Они расшифровали примерно такое же число сообщений как в Блетчли Парк. В основном это были сообщения немецкого военного командования (ключевая сеть GREEN: сообщения от 28 октября 1939 г., расшифрованы 17 января 1940 г.), и Люфтваффе (ключевая сеть RED: сообщения от 6 января 1940 г., расшифрованы 25 января 1940 г.), а также, после немецкого вторжения в Норвегию, сообщения оккупационной администрации Тронхейма (ключевая сеть YELLOW: начиная с 10 апреля 1940 г.).

После поражения Франции П. С. Брюно прибыл в Алжир — в Оран, 24 июня 1940 г., а затем в октябре того же года перебрался в Шато де Фузэ вблизи Юзэ, находящийся в неоккупированной части Франции. Так как с 1 мая 1940 г. листы Зыгальского стали бесполезны, полякам и англичанам пришлось некоторое время довольствоваться только правилами «Cillis» и Херивела (см. разд. 19.7). Польское подразделение под руководством подполковника Гвидо Лангера, которое англичане называли «Expositur 3000», было эвакуировано из Франции 9 ноября 1942 г., после того как союзники высадились в Северной Африке, а немцы заняли остававшуюся свободной часть Франции. Режевски и Зыгальский провели некоторое время в испанской тюрьме и прибыли в Лондон 3 августа 1943 г. Там они продолжили свою криптоаналитическую работу, однако их держали вдали от Блетчли Парк, где применялись бомбы Тьюринга—Уэлчмена и машины COLOSSUS.

19.7. Компромисс открытый текст—криптотекст: циклы обратной связи

Система ENIGMA была практически раскрыта англичанами в мае 1940 г. Они достигли успеха, используя атаку вероятного слова и компромисс открытый текст—криптотекст вместе с обратной связью Тьюринга (и Уэлчмена), которой они занимались с 1939 г. Тьюринг создал английскую BOMBE таким образом, что она позволяла использовать обратную связь. Этим был продолжен метод, который использовал Режевски в 1932 г. (см. разд. 19.6.2.6). Так что содержательно этот параграф скорее относится к концу 14-й главы.

Англичане были вынуждены ежедневно выполнять большой объем работы. Между тем, им сильно помогало постоянное нарушение немцами даже простейших правил криптографической надежности. Херивел заметил в мае 1940 г., что установки, используемые для передачи первого сообщения часто бывали очень близки (если фактически не совпадали) с позицией колес в дневных установках («совет Херивела»), что было вопиющей небрежностью. Кроме того, продолжалось использование стереотипных индикаторов, которые англичане называли «Cillis»[17]. Также существовало нарушение, когда основные установки использовались в качестве индикатора; Деннис Бэббидж называл это JABJAB. Когда немецкое начальство, наконец, прореагировало на это, ущерб был уже непоправимым. Самой низкой была дисциплина криптобезопасности в военно-воздушных силах напыщенного выскочки Геринга. Еще в апреле 1940 г., до того, как заработала BOMBE, математики и лингвисты в Блетчли Парк ухитрялись регулярно читать ENIGMA-сообщения Люфтваффе (ключевая сеть RED — красный, *англ.*). Аналогичного успеха с сообщениями военно-морского флота (ключевая сеть DOLPHIN — дельфин, *англ.*, «Heimische Gewässer» — родные воды, *нем.*, позднее «Hydra») им пришлось ждать до июля 1941 г. В декабре 1940 г. англичане добились успеха во взломе радиосообщений СС (ключевая система ORANGE — оранжевый, *англ.*). С сентября 1941 г. линия ENIGMA-сообщений с Берлином фельдмаршала Роммеля перестала быть надежной, а с середины 1942 г. англичане добились устойчивых взломов линий связи Люфтваффе (ключевые сети WASP — оса, *англ.*, GADFLY — овод, HORNET — шершень, SCORPION). Наиболее трудно поддающейся дешифрованию, по мнению англичан, была радиосвязь сухопутных войск, что было следствием основательной тренировки операторов. До весны 1942 г. ни одна линия коммуникаций ENIGMA сухопутных войск (кроме одной в России) не была взломана.

19.7.1. BOMBE Тьюринга. В общей атаке вероятного слова Тьюринг (и параллельно Уэлчмен) использовали вместо трех изолированных удвоенных циклов польской bomby целую систему циклов с обратной связью, образованную группой первых десяти, а позднее двенадцати скрамблеров. Такие системы циклов с обратной связью получаются из сопоставления вероятных слов и фрагментов криптотекста. К счастью, для достаточно длинных вероятных слов метод несовпадающего перебора (разд. 14.1) позволяет исключить многие позиции, к тому же заметные вероятные слова часто встречаются в начале или в конце сообщения (если не используется русское соединение). Поэтому вполне реально установить новую систему циклов с обратной связью для каждого сопоставления; их существует не слишком много.

Следующий пример[18] возвращает к Девуру и Кру. Рассмотрим криптотекст

---

[17] Иногда интерпретируемое как «sillies» (глупости, *англ.*). Уэлчмен: «Я не знаю, откуда взялся этот термин».

[18] Роторный порядок IV, I, II, отражатель B. Кольцевая установка 000, коммутации VO WN CR TY PJ QI. Текстовая установка tgv.

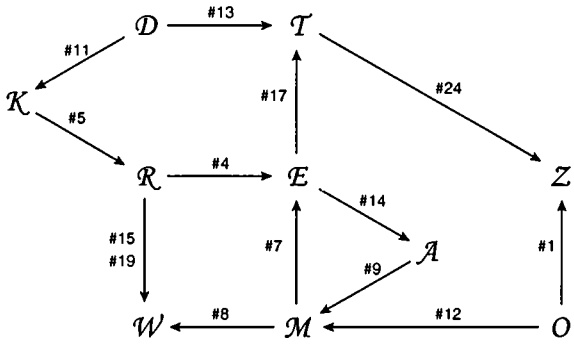


Рис. 156. Система циклов открытого текста/шифротекста

o b e r k o m m a n d o d e r w e h r m a c h t
Z M G E R F E W M L K M T A W X T S W V U I N Z

OVRLJ BZMGE RFEWM LKMTA WXTSW VUINZ GYOLY
FMKMS GOFTU EIU...

и предположим, что /oberkommandoderwehrmacht/ — вероятное слово. Третья слева позиция, не исключенная несовпадающим перебором, дает «шпаргалку»

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
o b e r k o m m a n d o d e r w e h r m a c h t
O V R L J B Z M G E R F E W M L K M T A W X T S W V U I N Z

с 24 парными соответствиями букв, пронумерованными от #1 до #24.

Пары, состоящие из одной буквы открытого текста и одной буквы криптотекста, можно представить при помощи ориентированного графа, показанного на рис. 156. Взаимообратный характер этих соответствий позволяет перейти к ориентированному графу. Из этого графа можно выделить подграф, называемый «меню», — в нашем примере это граф с восемью вершинами, показанный на рис. 157 в верхнем правом углу. Каждый цикл в этом подграфе устанавливает некоторую обратную связь в устройстве Тьюринга BOMBE. Меню с 6 буквами и 4 циклами, конечно, более выгодно, чем с 12 буквами и одним циклом: оно уменьшает опасность промахов.

Десять скрамблеров, соответствующих такому подграфу с десятью переходами, теперь соединены (26-жильными кабелями), и тестовый регистр связан, скажем, с $\mathcal{E}$ (рис. 157). К одному из входов, скажем, к e, подано напряжение. Позиции 14, 9, 7 образуют цикл («замыкание»): обозначая внутренние контакты через a, b, c, ... , y, z, коммутацию через T и подстановку, выполняемую скрамблером $\#i$ через P_i, получаем соотношения:

$$eT = mTP_7, \quad mT = aTP_9, \quad aT = eTP_{14} \quad \text{или} \quad eT = eTP_{14}P_9P_7.$$

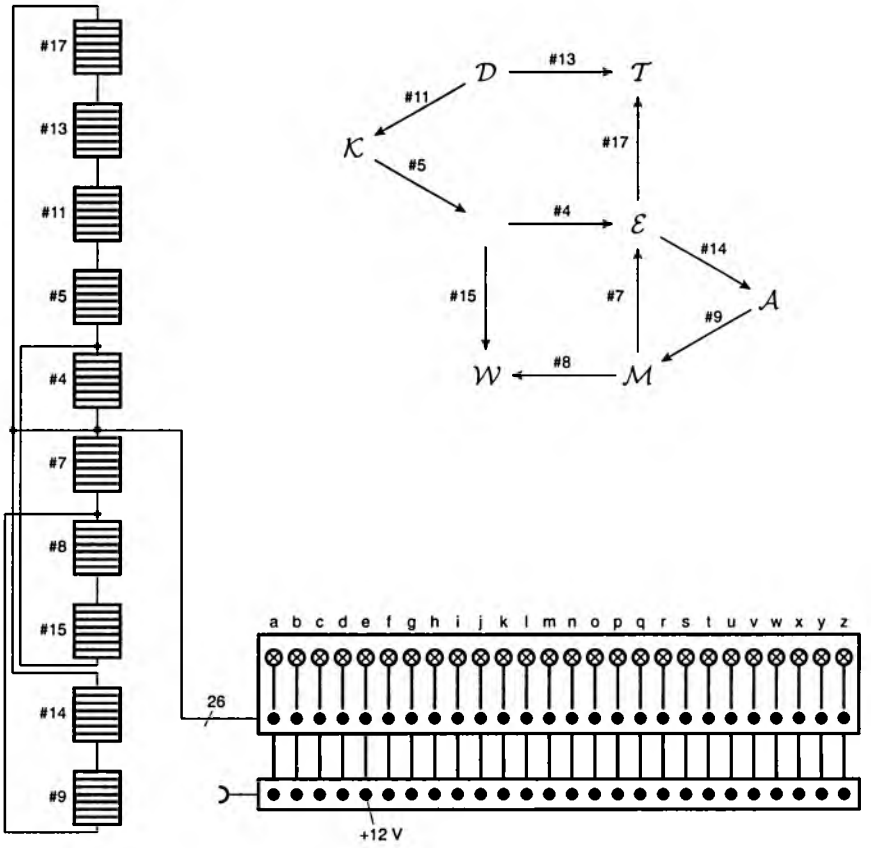


Рис. 157. Устройство Тьюринга BOMBE для системы циклов с обратной связью (рис. 156)

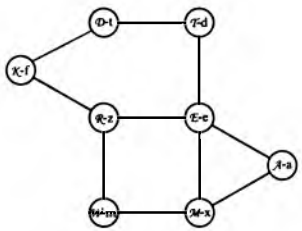


Рис. 158. Противоречивая «останавливающая» конфигурация

Гордон Уэлчмен

Таким образом, eT является неподвижной точкой для $P_{14}P_9P_7$. Но позиции 4, 15, 8, 7 тоже образуют цикл: сначала мы имеем соотношения

$$eT = rTP_4, \quad wT = rTP_{15}, \quad wT = mTP_8, \quad eT = mTP_7,$$

а поскольку скрамблерные подстановки взаимообратны, мы получаем

$$eT = mTP_7, \quad mT = wTP_8, \quad wT = rTP_{15}, \quad rT = eTP_4, \quad \text{или} \quad eT = eTP_4P_{15}P_8P_7.$$

Таким образом, eT является неподвижной точкой подстановки $P_4P_{15}P_8P_7$.

Кроме того, позиции 4, 5, 11, 13, 17 образуют цикл: снова используя взаимообратность скрамблерных подстановок, мы получаем соотношения

$$eT = rTP_4, \quad rT = kTP_5, \quad kT = dTP_{11}, \quad dT = tTP_{13}, \quad tT = eTP_{17}.$$

Таким образом, подстановка $P_{17}P_{13}P_{11}P_5P_4$ тоже имеет неподвижную точку eT.

Допустим, что позиция скрамблеров неверна. Тогда обычно (т. е. если существует достаточное количество циклов) напряжение распространяется на всю систему, и все лампочки тестовых регистров загораются. В таком случае релейная схема обнаруживает, что имеет место случай расходимости, и передвигает скрамблеры на следующую позицию.

Теперь предположим, что позиция скрамблеров «верна», т. е. она использовалась для шифрования (так что скрамблер #4 отображает $rT = /\mathrm{r}/$ в $eT = \mathrm{E}$). Тогда возможны два подслучая. Если коммутация выбрана правильно, т. е. вход e, к которому подведено напряжение, соответствует символу /e/, то напряжение не распространится дальше, и кроме лампочки, соответствующей /e/, ни одна лампочка не загорится. Но если коммутация выбрана неверно, то напряжение обычно (т. е. если существует достаточное количество циклов) распространяется дальше на всю оставшуюся систему, и загорятся все лампочки кроме одной, которая и укажет коммутацию. В обоих подслучаях сходимости машинные установки и загоревшиеся лампочки можно записать. Позиция скрамблеров определяет кольцевые установки, которые могут быть и ошибочными. Их правильность можно быстро выяснить, использовав найденные установки для дешифрования текста.

Эту возможность атаки цикла с обратной связью Тьюринга начисто проглядел в своем докладе 23-летний Хазенегер, ответственный за надежность ENIGMA в секции Безопасности собственных шифров шифротдела OKW. Как было показано выше, эта атака существенно опиралась на взаимообратный характер шифра ENIGMA; правда, она в принципе должна была бы работать и для невзаимообратных скрамблеров, хотя такие циклы встречаются гораздо реже. Например, единственным верным циклом на рис. 156 является цикл

7	9	14
m	a	e
E	M	A

и чтобы атака была успешной, потребовались бы очень длинные вероятные слова или еще более длинное меню. Например, в цикловой системе с обратной связью на рис. 156 система циклов должна была бы позволить соединить

вершину $\mathcal{U}$, соответствующую $\mathcal{A}$, или вершину $\mathcal{V}$, соответствующую $\mathcal{M}$, или чуть больше. В примере имеется, по сравнению с рис. 157, даже на один цикл больше: из $\mathcal{T}$ через $\mathcal{Z}$, $\mathcal{O}$, $\mathcal{M}$ и $\mathcal{A}$ к $\mathcal{T}$. Но это увеличило бы необходимое число скрамблеров в устройстве.

19.7.2. BOMBE Тьюринга—Уэлчмена. Гордон Уэлчмен усовершенствовал тьюрингову атаку цикла с обратной связью, приняв в расчет все соотношения, являющиеся следствием взаимообратного характера коммутации ENIGMA. Всякий раз, когда BOMBE Тьюринга останавливалась, ее узлам вроде $\mathcal{A}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{K}$ и т. д. сопоставлялись определенные внутренние контакты. На рис. 158 показана такая «останавливающая» конфигурация, включающая две интер-

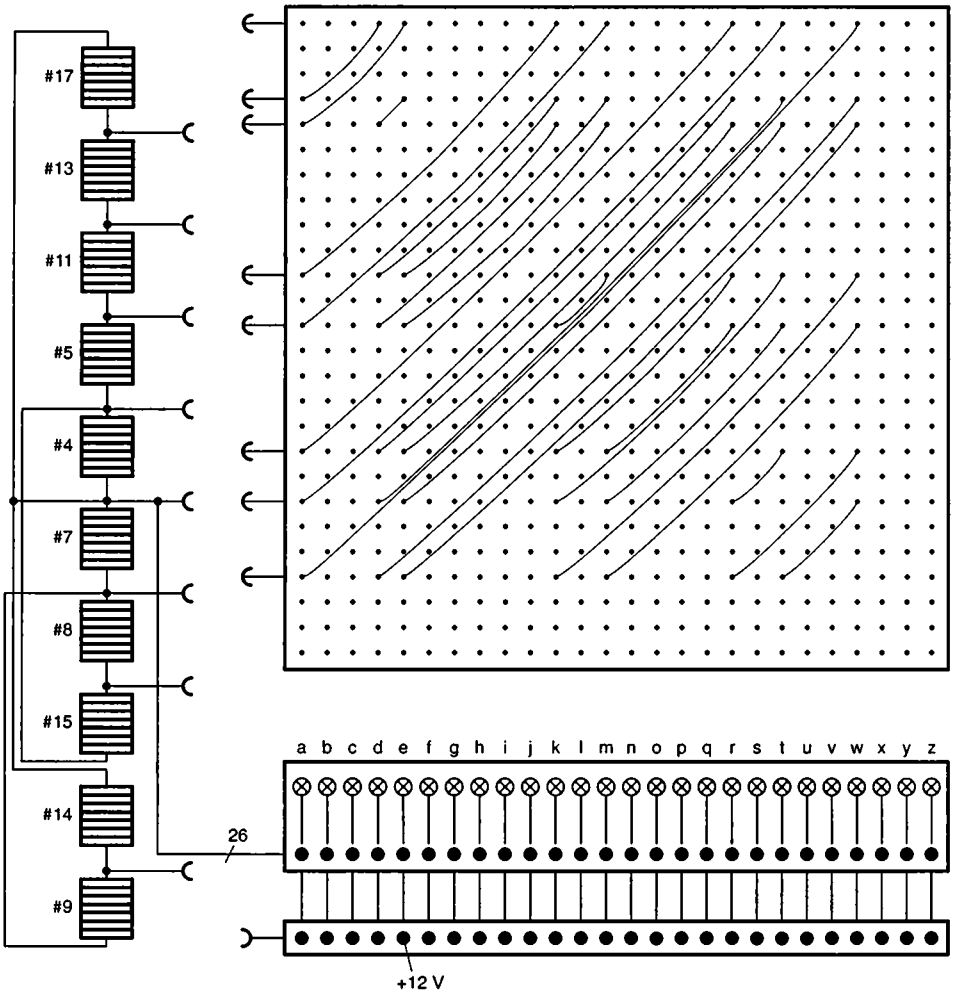


Рис. 159 а. Устройство BOMBE Уэлчмена для системы циклов с обратной связью (рис. 156)

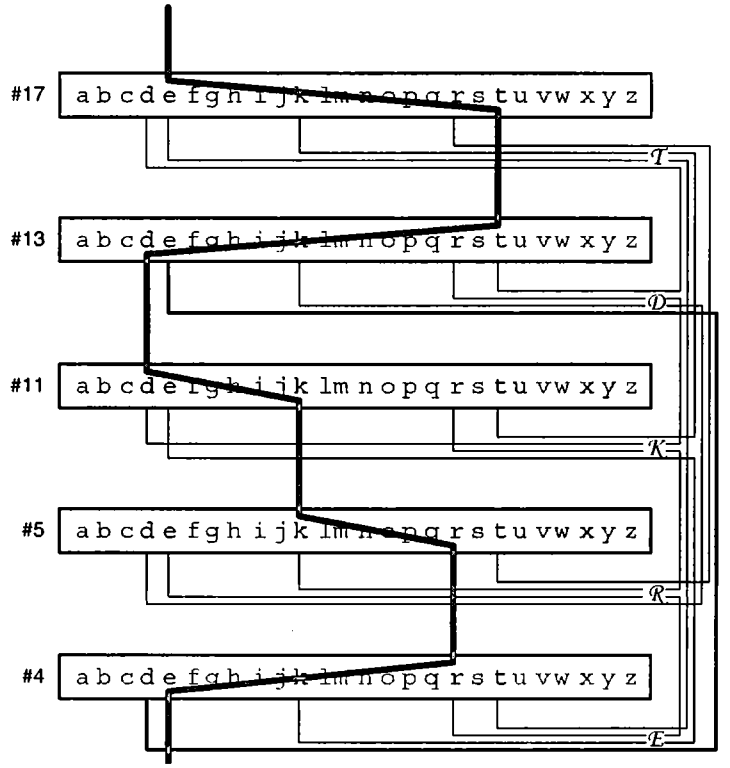


Рис. 159 b. Функционирование «диагонального коммутатора» Уэлчмена для примера из рис. 156

претации $\mathcal{A}$-a и $\mathcal{E}$-e, две интерпретации $\mathcal{D}$-t и $\mathcal{T}$-d, означающие коммутацию t-d, и две интерпретации $\mathcal{W}$-m и $\mathcal{M}$-x, которые, однако, противоречат инволютивному характеру коммутаций. Но BOMBE не должна «останавливаться» в такой конфигурации; полученное противоречие должно привести к расходимости внутри электропроводки ENIGMA.

Уэлчмен в ноябре 1939 г. создал простую электротехническую конструкцию — «диагональный коммутатор», изображенный на рис. 159 a. Его функционирование объясняется на рис. 159 b. Предположим, что $ET = dTP_{11}P_5P_4$. Жирная линия показывает, как соединение, устанавливаемое скрамблерами от e в шине $\mathcal{E}$ до d в шине $\mathcal{D}$ дополняется диагональным коммутатором с фиксированным соединением от d в шине $\mathcal{E}$ до e в шине $\mathcal{D}$.

С усовершенствованием Уэлчмена тьюрингова атака цикла с обратной связью достигла своей полной силы, и эффективность BOMBE сильно возросла. Теперь требовалось гораздо меньше циклов, чтобы наполнить текстовый регистр. А это не только помогало сэкономить скрамблеры, но позволяло также

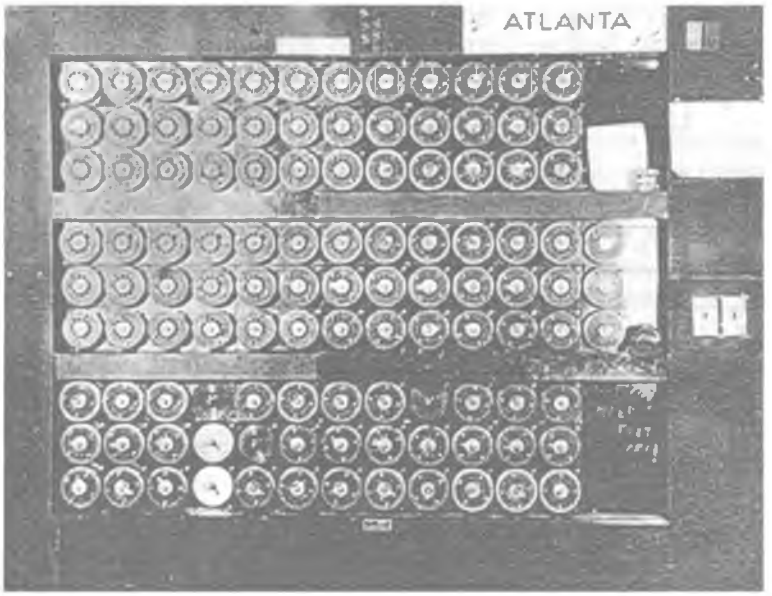


Рис. 160 а. Британская BOMBE «Atlanta» (стандартная модель)

уменьшить размеры шпаргалок и таким образом увеличивало шансы того, что средний ротор останется неподвижным. Так Уэлчмен стал настоящим героем истории с BOMBE. Девур и Кру в 1985 г. сформулировали это следующим образом (способным утешить Хазенегера):

«Сомнительно, чтобы до идеи Уэлчмена додумался кто-нибудь еще, поскольку даже выдающиеся лица, включая Тьюринга, отнеслись к ней скептически, когда Уэлчмен впервые высказал ее».

19.7.3. Еще BOMBEs. Люди в Блетчли Парк, которые назвали систему скрамблеров вместе с тестовым регистром и диагональным коммутатором «бомбой», не знали о польском приоритете названия и идеи. «Agnes» — первый экземпляр BOMBE Тьюринга—Уэлчмена после прототипа «Victory» (который еще не имел диагонального коммутатора) был готов в середине августа 1940 г.; ему требовалось 15 минут для полного перебора порядка колес. Весной 1941 г. в работе было 8 машин BOMBE и 12 были готовы к концу этого года. Они были построены British Tabulating Machine Company. Их число возросло до 30 в августе 1942 г., до 60 в марте 1943 г. и до 200 к концу Второй мировой войны.

В США как армия, так и военно-морской флот разработали высокоскоростные версии машины BOMBE, которые находились в эксплуатации и несколько лет после окончания войны. С октября 1943 г. на вооружении американской армии находилась настоящая релейная машина X68-003, скон-

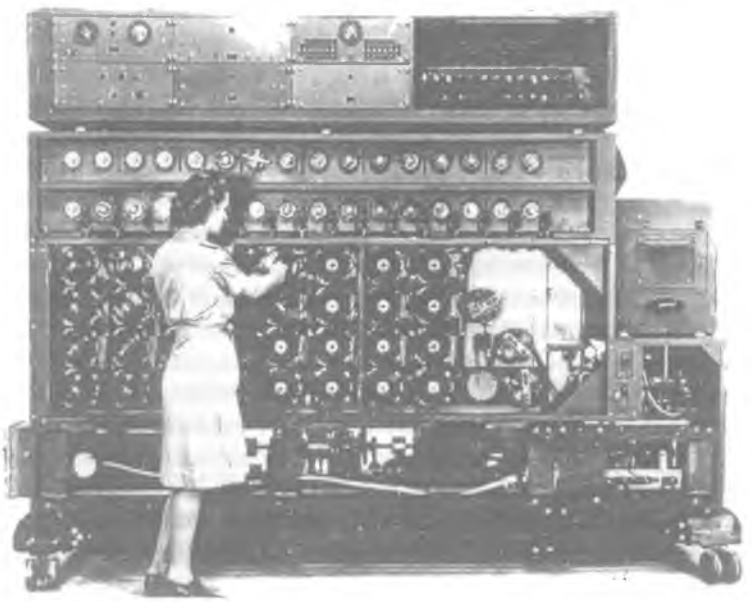


Рис. 160 b. BOMBE Военно-Морского флота США для 4-х роторной ENIGMA

струированная Уильямсом и оборудованная 144 скрамблерами. Она получила известность как MADAME X[19]; используя поэтапное переключение, машина позволяла быстро менять шпаргалку. Релейная имитация скрамблеров была медленной, зато при этом не использовались вращающиеся детали. Усовершенствованную с помощью лаборатории Белла машину применяли против 3-роторной машины ENIGMA и против 4-роторных машин немецкого военно-морского флота. Лишь одна MADAME X была действительно построена. Ее производительность соответствовала производительности шести или восьми английских BOMBEs. Ее конструирование и постройка стоили миллион долларов, что по отношению цена/производительность было несравнимо с BOMBEs военно-морского флота.

Дэш, экспериментировавший с быстрыми схемами счетчиков заряженных частиц и имевший заслуженную репутацию в электронике, принял в сентябре 1942 г. амбициозное поручение от американских военно-морских сил построить 350 машин BOMBE, каждую в несколько раз больше, чем BOMBE Тьюринга—Уэлчмена. Заказчики хотели получить эти машины к весне 1943 г. Дэш и его группа «полагали, что американская технология и методы массового производства могли творить чудеса» (Бурке). Дэш, однако, отклонил просьбу Венгера построить электронную версию: «Электронная BOMBE невозможна». Он был мудрым человеком и посчитал, что для «super-BOMBE»

[19]Неясно, было ли это имя намеком на Агнес Дрисколл, смелого борца за чистый криптоанализ (разд. 17.3.4), которую называли в Op-20-G Madame «X».

потребуется 20000 электронных ламп, тогда как англичанам даже для их COLOSSUS'а понадобилось всего около 2000.

Англичане не могли не помочь военному флоту своих союзников. Энгстром в июле 1942 г. послал своего представителя в Блетчли Парк. Тьюринг 7–13 ноября 1942 г. пересек на корабле «Королева Елизавета» Атлантический океан и отбыл снова из Нью-Йоркской гавани ночью 23 марта 1943 г. на «Шотландском экспрессе», вернувшись в Англию 29 марта. Он принес большую пользу за четыре месяца своего пребывания в США. Там он встретился с Шенноном, работавшим в лаборатории Белла над шифрованием речи. Служба Op-20-G была хорошо обеспечена деньгами и способными людьми, но меры безопасности там были жестче, чем в атомном проекте. Несмотря на все усилия, на постройку машин потребовалось больше времени, чем предполагалось, и к весне 1943 г. были готовы лишь два прототипа, ADAM и EVE. Франклин Делано Рузвельт оказывал проекту личную поддержку. Тем временем ситуация в Атлантике для Союзников улучшилась, в основном, благодаря дешифрованиям Блетчли Парк. У Дэша возникли проблемы с быстро вращающимися электромеханическими скрамблерами со щеточными контактами. К середине июня появилась надежда на скорое преодоление трудностей. Но когда 26 июня 1943 г. оказалось, что 13 моделей вообще не функционируют, возникла реальная опасность того, что весь проект может быть закрыт. Тем не менее Дэш не сдался. Механические проблемы шаг за шагом преодолевались и надежность увеличилась. В сентябре 1943 г. первые машины, построенные NCR, были отправлены из Дейтона в Вашингтон, где они начали работать. В середине ноября 50 машин BOMBE уже функционировали и 30 были на пути к этому. В 1944 г. успех уже был несомненным, хоть это потребовало больше времени, чем оптимистически предполагалось, и стоило почти втрое больше, чем планировалось, но даже после всего BOMBE Дэша (проект был так засекречен, что машина даже не имела названия) стоила всего $ 45 тысяч.

Четырехроторная BOMBE американских ВМС содержала 16 четырехроторных скрамблеров и диагональный коммутатор Уэлчмена. Она была в 200 раз быстрее, чем польская bomba и в 20 раз быстрее, чем BOMBE Тьюринга-Уэлчмена (специалисты утверждали, что в 26 раз быстрее). Она была еще на 30% быстрее, чем английская BOMBE 1943 г. версии «COBRA» (работавшая против четырехроторной ENIGMA немецких ВМС. Машина Op-20-G догнала ее: в декабре 1943 г. дешифрование сообщения ENIGMA TRITON занимало в среднем только 18 часов, хотя в июне 1943 г. для этого требовалось 600 часов. В противоположность своим английским коллегам они локализовали позиции скрамблеров и управляли работой всей цифровой электроники с 1500 тиратронами (электронными лампами, наполненными газом). Было построено по меньшей мере 100 машин BOMBE Дэша. Они настолько доказали свою надежность, что к концу 1943 г. всю работу по дешифрованию ключевой сети TRITON немецких ВМС возложили на американские военно-морские силы — большой шаг вперед по сравнению с соперничеством середины 1942 г.

Соглашение BRUSA о сотрудничестве в области кодов и шифров, заключенное в мае 1943 г. между США и Великобританией, и Холденское[20] соглашение, заключенное между их военно-морскими силами в октябре 1942 г., «начали движение двух наций к беспрецедентному уровню сотрудничества» (Бурке) в криптоанализе. Но некоторые трения и напряженность оставались. «До соглашения UKUSA 1946 г. не было случая, чтобы две нации забыли то уникальное отношение доверия, которое поддерживалось на протяжении холодной войны» (Бурке).

VIPER (гадюка, *англ.*) и PYTHON — американские машины, действовавшие против японских роторных шифромашин. Они были построены из релейно-переключательных элементов и постепенно оснащались электронными узлами.

Наконец, к концу войны был сделан неизбежный переход к настоящим электронным машинам: Op-20-G построил машину RATTLER (сокрушительный удар, *англ.*), тогда как SIS в начале 1945 г. построил преемника релейной машины AUTOSCRITCHER, названного соответственно SUPERSCRITCHER[21]. Op-20-G построил также DUENNA, а англичане построили GIANT (гигант) — название, которое до сих пор отсутствует в открытой литературе. Все эти криптоаналитические машины предназначались для определения коммутаций в шифраторах. К разработке DELILAH приступили в 1943 г., сконструировали его в июне 1944 г. и закончили 6 мая 1945 г.

19.7.4. Приход компьютеров. Идея универсального снабженного программой компьютера, которая возникла в середине 1940 г. у Эккерта и Мочли и которая была тщательно разработана фон Нейманом и Голдстайном, стала довольно скоро, хоть и не открыто, оказывать влияние на использование машин в криптоанализе. Пендеграсс выступил за использование универсальных компьютеров. Начало положили в 1948 г. машина ABNER, созданная SIS (разработка которой потребовала четырех лет), и ATLAS, начало работ над которой в Op-20-G восходит к августу 1947 г. (см. также разд. 17.3.5 и разд. 18.6.3). Агентство национальной безопасности, являющееся преемником как SIS, так и Op-20-G, и ставшее высшим авторитетом в криптографии, дало громадный импульс возникающему компьютерному миру. Кампейн, Снайдер и Томаш сообщили о влиянии американских криптологических организаций на цифровую компьютерную индустрию.

Несколько раньше, в начале 1946 г., военно-морские офицеры запаса Энгстром, Норрис и Мидер основали частную компанию Engineering Research Associated, Inc. (ERA). Им оказывали помощь Томпкинс и Ховард, содействовали Ичус и Пендеграс из Op-20-G. ERA разрабатывала компьютеры в тесном контакте с Военно-морским министерством. Вехой в их работе был компью-

[20] Холден, капитан военно-морских сил США, начальник связи военно-морских сил.

[21] Выражение «scritchmus» пришло из жаргона Блетчли Парк («Я не могу теперь вспомнить, какая техника была прозвана скрипучей (scritchmus)» — писал Таунт), и этот метод был развит Деннисом Бэббиджем. См. также разд. 14.5. Однако Эрскин полагает, что «scritching» пришло из выражения «scuitching out contradictions» (удалить противоречия, *англ.*).

тер Task-13, переименованный в 1948 г. в ATLAS и выпущенный в декабре 1950 г. За этим последовала торговля компьютером ERA 1101 (анонсированного в 1951 г.). Его преемником был Task 29 (ATLAS II), который был завершен в 1953 г. и выпущен на рынок под именем ERA 1103 и сразу же заслуживший огромный успех. В 1952 г. E.R.A. стала дочерней компанией Remington Rand. В 1954 г. Remington Rand занимала вторую позицию на рынке с усовершенствованным 1101 A (и UNIVAC II, разработанным группой Эккерта—Мочли).

Их конкурент IBM анонсировал в 1951 г. Defense Calculator, переименованное затем в IBM 701. IBM начал торговлю в апреле 1953 г. с поставок IBM 701, а в октябре 1953 г. Remington Rand продала ATLAS II правительству. Компьютер IBM STRETCH появился в результате усовершенствования компьютера E.R.A. 1962 г. HARVECT.

В 1970-х гг. электроинженер Крэй (1925–1996 гг.), который раньше работал в E.R.A. под началом Энгстрома, создал собственную компанию, и в 1976 г. спроектировал компьютер CRAY-I. АНБ частично использует в своей работе коммерческую продукцию, в том числе компьютеры CRAY, конструкция которых скрывает некоторые криптоаналитические алгоритмы, используемые АНБ. Холодная война в ее нынешней миниатюрной форме ведется на уровне микросхем.

Глава 20

Линейный базисный анализ

Не было бы преувеличением сказать, что абстрактная криптография — то же самое, что абстрактная математика.

А. Адриан Алберт, 1941 *г.*

20.1. Приведение линейных многосимвольных подстановок

В благоприятных случаях линейная многосимвольная подстановка с шириной шифрования n может быть сведена к перебору ширины n, т. е. если угадано дешифрование некоторого множества из n часто встречающихся n-грамм криптотекста в некоторое множество из n n-грамм открытого текста. Иногда это оказывается легче, чем кажется на первый взгляд, например если можно угадать достаточно длинное вероятное слово, которое встречается в n разных фазах.

20.1.1. Пример. Чтобы расчеты можно было легко проверить, мы ограничимся примером с $n = 3$. Пусть дан криптотекст

FDYSW IJXNZ NSNRE NHUWA WMIEI EXWSX
ISIGQ JNTBD BWDPU.....

Предположим, что он получен с помощью линейной многосимвольной подстановки ширины 3 над стандартным алфавитом — возможно, это была вторая попытка в серии попыток с возрастающей шириной шифрования. Тогда криптотекст читается по триграммам над $\mathbb{Z}_{26}$:

5324 1822 8 92313 251318 **13174** 13720 **22 0 22** 12 8 4
8423 221823 418 8 **6 16 9** 13191 3 122 3 1520 .....

и мы предполагаем, что триграммы, выделенные жирным шрифтом, **13 17 4**, **22 0 22** и **6 16 9** появляются достаточно часто в дальнейшем криптотексте.

Ввиду очень частой встречаемости слова |ation| в английском, французском и немецком языках, мы можем предположить, что три триграммы открытого текста /ati/, /tio/ и /ion/ содержатся в открытом тексте; остается узнать, в каком порядке.

В $\mathbb{Z}_{26}$ эти триграммы открытого текста суть **0 19 8**, **19 8 14** и **8 14 13**. Поэтому матрица линейной подстановки X определяется следующим образом:

$$\begin{pmatrix} 0 & 19 & 8 \\ 19 & 8 & 14 \\ 8 & 14 & 13 \end{pmatrix} X = P \begin{pmatrix} 13 & 17 & 4 \\ 22 & 0 & 22 \\ 6 & 16 & 9 \end{pmatrix},$$

где P — некоторая перестановочная матрица, неизвестная в данный момент.

В $\mathbb{Z}_{26}$ имеется шесть решений для $6 = 3!$ перестановок. Представим, что после опробования двух или трех из них мы получаем

$$\begin{pmatrix} 0 & 19 & 8 \\ 19 & 8 & 14 \\ 8 & 14 & 13 \end{pmatrix} X = \begin{pmatrix} 22 & 0 & 22 \\ 6 & 16 & 9 \\ 13 & 17 & 4 \end{pmatrix},$$

что дает

$$X = \begin{pmatrix} 12 & 8 & 17 \\ 8 & 18 & 24 \\ 13 & 19 & 14 \end{pmatrix}.$$

Это представляется правильным, так как после замены чисел буквами матрица

$$X = \begin{pmatrix} M & I & R \\ I & S & Y \\ N & T & O \end{pmatrix}$$

приводит к «осмысленному» паролю

$$MINISTRIO[F],$$

что подтверждает решение в смысле Рорбаха.

20.1.2. Ловушка. Но имеется одна сложность. Если мы сейчас попытаемся дешифровать текст и найти обратную матрицу для X, мы с удивлением можем обнаружить, что ее не существует. Действительно, использование «осмысленного» пароля не гарантирует инъективности шифрования, и X не инъективна: вектор (0 13 0) подстановка X переводит в (0 0 0). Это означает, что даже законный дешифровальщик развеселится, взглянув на такое «правильное» решение:

5 3 24 принадлежат 8 0 5 $\hat{=}$ i a f и 8 13 5 $\hat{=}$ i n f;
18 22 8 принадлежат 14 4 12 $\hat{=}$ o e m и 14 17 12 $\hat{=}$ o r m, и т. д.

Полифоническое дешифрование дает:

```
a  e  i  c  b  f  g  a
i f o m d r e t i n o n a i o a
n  r  v  p  o  s  t  n
e  i  g  i  a  h  h
l a d o s a t o n b o t o r .....
r  v  t  v  n  u  u
```

И правильный открытый текст легко находится: «inform direction of national radio station about our...».

20.2. Восстановление ключа

Если некоторый квазинепериодический ключ многоалфавитной линейной многосимвольной подстановки ширины n образуется итерацией регулярной $n \times n$-матрицы A над $\mathbb{Z}_N$, то можно поменять ролями открытый и ключевой тексты. Вероятное слово длины k, где $k \geqslant n^2 + n$, передвигается вдоль криптотекста и вычитается из него для каждой позиции. То, что остается, в случае удачи, является ключевым фрагментом $(s_{M+1}, s_{M+2}, \dots, s_{M+n^2+n}, s_{M+k})$ длины $k \geqslant n^2 + n$. Для определения A в $\mathbb{Z}_N$ достаточно n уравнений

$$
\begin{aligned}
(s_{M+1}, s_{M+2}, \dots, s_{M+n})A &= (s_{M+n+1}, s_{M+n+2}, \dots, s_{M+2n}),\\
(s_{M+n+1}, s_{M+n+2}, \dots, s_{M+2n})A &= (s_{M+2n+1}, s_{M+2n+2}, \dots, s_{M+3n}),\\
(s_{M+2n+1}, s_{M+2n+2}, \dots, s_{M+3n})A &= (s_{M+3n+1}, s_{M+3n+2}, \dots, s_{M+4n}),\\
&\vdots\\
(s_{M+n^2-n+1}, s_{M+n^2-n+2}, \dots, s_{M+n^2})A &= (s_{M+n^2+1}, s_{M+n^2+2}, \dots, s_{M+n^2+n}).
\end{aligned}
$$

Так как $k > n^2 + n$, мы даже получаем переопределенную систему линейных уравнений.

В качестве примера, возьмем три пары чисел (1 0), (3 5), (23 22), получаемые из (1 0) двумя итерациями матрицы A, которая определяется условиями

$$(1\ 0)A = (3\ 5) \quad \text{и} \quad (3\ 5)A = (23\ 22),$$

так что результат в $\mathbb{Z}_{26}$ такой:

$$A = \begin{pmatrix} 3 & 5 \\ 8 & 17 \end{pmatrix}.$$

Если, в случае удачи, позиция вероятного слова подходит, то система может быть решена, и для переопределенного случая несколько таких систем разрешимы и дают общее решение, строго определяющее правильное решение. В случае же неудачи позиция вероятного слова не подходит, и система или одна из систем может не решаться. Если она случайно оказывается разрешимой, то ключевой текст может быть продлен, и вычитание его из криптотекста дает, как правило, бессмысленный текст — указывая на неудачу. Для достаточно длинного вероятного слова ключ обычно полностью обнаруживается, и промахи не должны встречаться.

20.3. Восстановление линейных регистров сдвига

Линейные регистры сдвига в широком смысле подпадают в качестве частичного случая под атаку, рассмотренную в разд. 20.2. В этом случае матрицей A является $n \times n$ сопровождающая матрица (разд. 8.6.1)

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & \alpha_k \\ 1 & 0 & 0 & \dots & 0 & \alpha_{k-1} \\ 0 & 1 & 0 & \dots & 0 & \alpha_{k-2} \\ & & \dots\dots\dots & & & \\ 0 & 0 & 0 & \dots & 0 & \alpha_2 \\ 0 & 0 & 0 & \dots & 1 & \alpha_1 \end{pmatrix}.$$

Если ключ порождается итерацией такой сопровождающей матрицы, то фрагмента ключевого текста длины $2n$ достаточно для восстановления сопровождающей матрицы, и таким образом можно восстановить весь ключ. Опять, чтобы было легко проверить расчеты, мы ограничимся примером с $n = 4$.

Пусть имеется следующий криптотекст:

C G V J F M C I H T X U F S D Y V L M R ... ...

Допустим также, что в данном случае имеет место многосимвольный шифр ВИЖЕНЕР с квазипериодической ключевой последовательностью, порожденной линейной многосимвольной подстановкой ширины 4 в $\mathbb{Z}_{26}$. В качестве вероятного слова мы возьмем /broadcast/.

Можно начать с гипотезы, что правильная позиция вероятного слова — начало сообщения. Это приводит к следующей ситуации

C G V J F	M C I H T	X U F S D	Y V L M R ...
2 6 21 9 5	12 2 8 7 19	23 20 5 18 3	24 21 11 12 17 ...
b r o a d	c a s t		
1 17 14 0 3	2 0 18 19		
1 15 7 9 2	10 2 16 14		

Отсюда получаем итерационное уравнение в $\mathbb{Z}_{26}$

$$\begin{pmatrix} 1 & 15 & 7 & 9 \\ 15 & 7 & 9 & 2 \\ 7 & 9 & 2 & 10 \\ 9 & 2 & 10 & 2 \\ 2 & 10 & 2 & 16 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & t_1 \\ 1 & 0 & 0 & t_2 \\ 0 & 1 & 0 & t_3 \\ 0 & 0 & 1 & t_4 \end{pmatrix} = \begin{pmatrix} 15 & 7 & 9 & 2 \\ 7 & 9 & 2 & 10 \\ 9 & 2 & 10 & 2 \\ 2 & 10 & 2 & 16 \\ 10 & 2 & 16 & 14 \end{pmatrix}$$

и переопределенную линейную систему

$$\begin{pmatrix} 1 & 15 & 7 & 9 \\ 15 & 7 & 9 & 2 \\ 7 & 9 & 2 & 10 \\ 9 & 2 & 10 & 2 \\ 2 & 10 & 2 & 16 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 10 \\ 2 \\ 16 \\ 14 \end{pmatrix},$$

которая не может быть решена. Первые четыре строки могут быть преобразованы по методу Гаусса в систему

$$\begin{pmatrix} 1 & 15 & 7 & 9 \\ 0 & 1 & 9 & 9 \\ 0 & 0 & 1 & 17 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 18 \\ 0 \\ 8 \end{pmatrix}$$

с решением

$$t_4 = 8, \quad t_3 = 20, \quad t_2 = 0, \quad t_1 = 24,$$

которое, очевидно, не удовлетворяет пятому уравнению.

Следующая гипотеза, которую надо проверить, состоит в том, что вероятное слово начинается со второй позиции открытого текста. Она приводит к следующей ситуации:

C G V J F	M C I H T	X U F S D	Y V L M R ...
2 6 21 9 5	12 2 8 7 19	23 20 5 18 3	24 21 11 12 17 ...
b r o a	d c a s t		
1 17 14 0	3 2 0 18 19		
5 4 21 5	9 0 8 15 0		

и итерационному уравнению в $\mathbb{Z}_{26}$

$$\begin{pmatrix} 5 & 4 & 21 & 5 \\ 4 & 21 & 5 & 9 \\ 21 & 5 & 9 & 0 \\ 5 & 9 & 0 & 8 \\ 9 & 0 & 8 & 15 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & t_1 \\ 1 & 0 & 0 & t_2 \\ 0 & 1 & 0 & t_3 \\ 0 & 0 & 1 & t_4 \end{pmatrix} = \begin{pmatrix} 4 & 21 & 5 & 9 \\ 21 & 5 & 9 & 0 \\ 5 & 9 & 0 & 8 \\ 9 & 0 & 8 & 15 \\ 0 & 8 & 15 & 0 \end{pmatrix}.$$

Это приводит к следующей переопределенной линейной системе

$$\begin{pmatrix} 5 & 4 & 21 & 5 \\ 4 & 21 & 5 & 9 \\ 21 & 5 & 9 & 0 \\ 5 & 9 & 0 & 8 \\ 9 & 0 & 8 & 15 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{pmatrix} = \begin{pmatrix} 9 \\ 0 \\ 8 \\ 15 \\ 0 \end{pmatrix},$$

которая может быть решена: первые 4 строки методом Гаусса можно преобразовать в

$$\begin{pmatrix} 1 & 6 & 25 & 1 \\ 0 & 1 & 23 & 7 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{pmatrix} = \begin{pmatrix} 7 \\ 18 \\ 23 \\ 3 \end{pmatrix},$$

с решением

$$t_4 = 3, \quad t_3 = 11, \quad t_2 = 4, \quad t_1 = 17,$$

которое, очевидно, удовлетворяет пятому уравнению.

Итерационная матрица A для продолжения ключевого текста, таким образом, имеет вид

$$A = \begin{pmatrix} 0 & 0 & 0 & 17 \\ 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 11 \\ 0 & 0 & 1 & 3 \end{pmatrix},$$

а ее обратная матрица:

$$A^{-1} = \begin{pmatrix} 12 & 1 & 0 & 0 \\ 7 & 0 & 1 & 0 \\ 9 & 0 & 0 & 1 \\ 23 & 0 & 0 & 0 \end{pmatrix}.$$

Поэтому ключ может быть дополнен до

2 5 4 21 5 9 0 8 15 0 15 7 25 4 24 23 20 9 19 3 ...,

что приводит к следующему дешифрованию:

C G V J F	M C I H T	X U F S D	Y V L M T ...
2 6 21 9 5	12 2 8 7 19	23 20 5 18 3	24 21 11 12 17 ...
2 5 4 21 5	9 0 8 15 0	15 7 25 4 24	23 20 9 19 3 ...
0 1 17 14 0	3 2 0 18 19	8 13 6 14 5	1 1 2 19 14 ...
a b r o a	d c a s t	i n g o f	b b c t o ...

«A broadcasting of BBC tonight announced the Allied invasion to be expected within forty-eight hours» (радио Би-Би-Си сегодня ночью объявило о высадке союзников, которая должна начаться в течении 48 часов (*англ.*)).

Итерационная матрица выведена из пароля

$$(FID)DLER \doteq (5\ 8\ 3)\ 3\ 11\ 4\ 17.$$

Для ключевой последовательности, порожденной бинарным линейным регистром сдвига, связанным с шифром ВИЖЕНЕР над $\mathbb{Z}_2$, т. е. ВЕРНАМ, также выполняется все, о чем говорилось выше. Поэтому чтобы избежать такой атаки, шифрование с помощью регистра сдвига должно быть обязательно нелинейным.

Глава 21

Анаграммирование

> Abadonner les méthodes de substitution pour celles de transposition a été changer son cheval borgne pour un aveugle.
>
> [Отказ от методов подстановок в пользу перестановки подобен замене одноглазой лошади на слепую, *фр.*]
>
> *Этьен Базерье*, 1901 *г.*

Перестановки были в течение некоторого времени излюбленным средством военных, в частности, в конце XVIII и начале XIX веков во Франции, Германии, Австрии и некоторых других странах. Они казались самыми подходящими в качестве полевых («окопных») кодов, лучшими среди простых кодов. Базерье полюбил их около 1900 г., при этом он приписывал перестановочным системам, которые на первый взгляд кажутся трудными, *иллюзорную сложность*. Криптоаналитики обычно любят противников, которые применяют простые перестановки (подобные ручным шифрам немецкого Абвера), потому что они обещают стать легкой добычей; подобно литературным опытам любителей, криптоанализ перестановок сравнительно бесхитростен.

21.1. Перестановка

Простая перестановка (разд. 6.2.1) с малой шириной шифрования n может быть охарактеризована для очень малого данного n систематическим изучением контактов в биграммах (возможно также в триграммах и тетраграммах). Во Второй мировой войне криптоаналитические службы немецкого мИДа (Pers Z) и Главного командования Вермахта (Хи) использовали специальные машины (специальные сравниватели и приборы оценки биграмм) (Рорбах, Йенсен) для полуавтоматического решения простых колонных перестановок и простых блочных перестановок. По существу, переборный метод «отрежь и приклей» из разд. 12.8.2 был механизирован. Кусок криптотекста сопоставлялся со всем криптотекстом во всех взаимных позициях, и для наблюдаемых биграмм были перемножены теоретические биграммные частоты,

потом были выделены позиции, где такое произведение было высоким. Этот метод полезен даже в том случае, когда не все столбцы, принимающие участие в перестановке, имеют равную длину.

Для армии США к концу войны SIS построил машину FREAK (каприз) — биграммный счетчик из электрических конденсаторов, заменившую построенный NCR в 1944 г. MIKE, который, согласно Бурке, был «огромной электромеханической штуковиной».

21.1.1. Пример. Рассмотрим пример такого «контактного метода» для следующего криптотекста

SSNKL HONIW MMEUN TAHUL INNAH NCINF CIERO
NACBA MZGHN KTHWC DESIN KCAIE ANIM

Распределение частот отдельных символов дает малое отклонение от распределения частот немецкого языка, что позволяет рассматривать эту криптограмму как перестановку. Общее число символов равно 64, что указывает на перестановку с квадратом 8×8 или с прямоугольником 4×16. Испытывая квадрат 8×8,

S I AHEMWC
SWHNRZCA
NMUCOGDI
KMLINHEE
LEINANSA
HUNFCKIN
ONNCBTNI
NTAIAHKM

мы возьмем столбец (а можно также строку), который содержит много частых символов, скажем, пятый: ERONACBA, и сопоставим его с другими столбцами. Результирующие биграммы и их частоты (в %) задаются следующей диаграммой (пустые клетки означают частоты ниже 0.5%)

ES 140	EI 193	EA 26	EH 57	EM 55	EW 23	EC 25
RS 54	RW 17	RH 19	RN 31	RZ 14	RC 9	RA 80
ON 64	OM 17	OU 3	OC 15	OG 5	OD 7	OI 1
NK 25	NM 23	NL 10	NI 65	NH 17	NE 122	NE 122
AL 59	AE 64	AI 5	AN 102	AN 102	AS 53	AA 8
CH 242	CU	CN	CF	CK 14	CI 1	CN
BO 8	BN 1	BN 168	BC	BT 4	BN 1	BI 12
AN 102	AT 46	AA 8	AI 5	AH 20	AK 7	AM 28

Сопоставление его со столбцом SSNKLHON ясно показывает частоты более высокие, чем с остальными столбцами. Перемножая эти частоты, получаем $1.41 \times 10^{14} \times 10^{-32} = 1.41 \times 10^{-18}$, тогда как остальные столбцы дают значения ниже $3.74 \times 10^{9} \times 10^{-32} = 3.74 \times 10^{-23}$.

Ввиду столь хорошего результата однозначно определяется следующий столбец, который мы будем тестировать с остальными — это первый столбец исходной таблицы SSNKLHON. Теперь сравнение с другими столбцами дает

SI 65	SA 36	SH 9	SM 12	SW 10	SC 89
SW 10	SH 9	SN 7	SZ 7	SC 89	SA 36
NM 23	NU 33	NC 5	NG 94	ND 187	NI 65
KM 1	KL 10	KI 7	KH 1	KE 26	KE 26
LE 65	LI 61	LN 4	LN 4	LS 22	LA 45
HU 11	HN 19	HF 2	HK 3	HI 23	HN 19
ON 64	ON 64	OC 15	OT 9	ON 64	OI 1
NT 59	NA 68	NI 65	NH 17	NK 25	NM 23

На этот раз выделяется сопоставление со столбцом WCDESINK, правда, не так отчетливо, как раньше, но произведение $3.50 \times 10^{12} \times 10^{-32} = 3.50 \times 10^{-20}$ все еще бесспорно максимальное, так как все остальные произведения меньше $5.39 \times 10^{11} \times 10^{-32} = 5.39 \times 10^{-21}$. Решившись продолжить тестирование со столбцом WCDESINK, получаем, что следующее сопоставление выделяет столбец AHULINNA. Если теперь выделенные до сих пор столбцы выписать рядом, результат будет такой:

E	S	W	A
R	S	C	H
O	N	D	U
N	K	E	L
A	L	S	I
C	H	I	N
B	O	N	N
A	N	K	A

Поразительно, но это уже открытый текст; а поскольку использованные до сих пор столбцы имеют номера 5, 1, 7 и 3, то можно предположить, что используется не квадрат 8×8, а прямоугольник 4×16, и для получения перестановки оставшихся столбцов надо взять ту же (считая каждый четный столбец продолжением предыдущего нечетного) и удвоить длину столбца:

M	I	C	H
Z	W	A	N
G	M	I	C
H	M	E	I
N	E	A	N
K	U	N	F
T	N	I	C
H	T	M	I

Полный открытый текст читается так:

«es war schon dunkel als ich in bonn ankam ich zwang mich meine ankunft nicht mi[t der automatik...]» (Генрих Бёлль «Глазами клоуна», 1963 г.).

21.1.2. Сдвинутые столбцы. В рассмотренном примере первый столбец открытого текста имел более частые буквы, чем остальные. Это бывает не всегда, поэтому кроме правых контактов надо исследовать и левые контакты. Как только достигнут самый первый (или самый последний) столбец, продолжать имеет смысл лишь со столбцами, сдвинутыми на одну позицию. Использование частот триграмм увеличивает число шагов перебора, но зато может дать более стабильные перестановки.

21.1.3. Предостережение. Мы видели, что простая перестановка с фиксированными схемами шифрования определенной ширины не обеспечивает надежности, если длина текста лишь в небольшое число раз больше этой ширины. Перестановка с шириной, равной длине текста, как правило, допускает более одного «осмысленного» решения даже для очень длинных текстов. Остроумный адвокат поэтому мог бы спасти Брата Тома Джонатана Свифта (разд. 6.3), если бы нашел другое, безболезненное решение анаграммы. Однако надежность такого рода перестановки целиком основывается на одноразовой перестановке символов, что означает индивидуальный ключ. Если схема шифрования с помощью перестановки используется несколько раз, то может быть применена простая атака из разд. 21.1.1 и специальные методы, обсуждаемые в разд 21.3.

21.1.4. Шаблоны кодовой группы. Даже когда код перешифровывается простой перестановкой, к нему можно применить упомянутый выше метод, если его кодовые группы имеют определенные шаблоны. Например, в случае, если «удобопроизносимые» кодовые группы строятся при помощи гласно-согласных шаблонов, подобно СГСГС для кода GREEN госдепартамента США (разд. 4.4.2). Еще во Второй мировой войне госдепартамент использовал коды типа СГСГС и СГССГ, некоторые свойства которых помогали немецкой криптоаналитической группе Pers Z (разд. 19.3.2) отслаивать добавки.

21.1.5. Иллюзорная сложность. Более того, «контактный метод» работает также для смешанной строчно-колонной перестановки и смешанной строчно-блочной перестановки (разд. 6.2.3), поскольку контакт прерывается лишь случайно. Отсюда получается промежуточный криптотекст с переставленными строками, например, в 8×8 квадрате

```
MICHZWAN ALSICHIN ONDUNKEL NEANKUNF
GMICHMEI ESWARSCH TNICHTMI BONNANKA
```

Как Живарж так и Эйрауд подчеркивали, что *двойная перестановка* из смешанной строчно-колонной перестановки и смешанной строчно-блочной перестановки, включающей перестановку Нигилиста, является не более трудным для дешифрования, чем простейшая колонная перестановка. *Двойная перестановка* обладает лишь *иллюзорной сложностью*.

21.2. Двойная колонная перестановка

Двойная колонная перестановка (разд. 6.2.4) — исключая особые случаи вроде рассмотренных в разд. 6.2.5 — требует от незаконного дешифровальщика гораздо более трудной работы. Причина в том, что после первой перестановки все контакты полностью разорваны. Эйрауд рассмотрел эту причину довольно подробно, но не смог дать полный метод. Кан писал: «...в теории криптоаналитик просто должен строить столбцы второго блока по два и по три, чтобы их биграфы и триграфы могли быть, в свою очередь, присоединены к хорошим фрагментам открытого текста. Но это гораздо проще сказать, чем сделать. Даже одаренный криптоаналитик может добиться этого лишь изредка; и даже с помощью вероятного слова это никогда не бывает легким».

Действительно мощным средством атаки является кратное анаграммирование.

21.3. Кратное анаграммирование

Для наиболее общего случая перестановки, даже с шириной порядка величины всего текста, включая решетки и маршрутные транскрипции, существует один общий метод, не требующий ничего другого кроме двух открытых текстов одной и той же длины, зашифрованных одной и той же схемой шифрования, т. е. шифрование должно быть перестановкой и должно быть повторено по крайней мере один раз. Такой компромисс открытый текст — открытый текст наводит на параллель с методом наложения Керкхоффса.

21.3.1. Пример. Предлагаемый метод основан на том простом факте, что одинаковые схемы шифрования производят одну и ту же перестановку открытого текста. Поэтому надо один криптотекст подписать под другим и образованные таким образом столбцы сохранить. Предположим, что мы имеем (в фазе) фрагменты криптотекста (Кан) GHINT и OWLCN. Это означает, что должны быть анаграммированы пять пар:

G H I N T
O W L C N

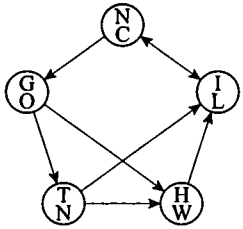
Среди $5 \times 4 = 20$ комбинаций по два столбца лишь следующие 12 (расположенные по убыванию контакта)

TH NG GT IN TI NI GH HI IG TN HT GI
NW CO ON LC NL CL OW WL LO NC WN OL

имеют достаточно большой контакт на обоих уровнях. Используя только первые четыре комбинации, получаем лишь бессмысленное решение:

I N G T H
L C O N W

И даже с первыми восемью комбинациями получаются только циклические сдвиги этого решения, как показывает следующий граф с пятью вершинами и восемью дугами:



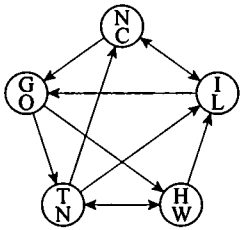
Используя девятую комбинацию, находим новое решение

N I G T H
C L O N W

которое снова бессмысленно. И только с первыми одиннадцатью комбинациями получаем осмысленное решение

N I G T H
C L O W N

(и его циклические сдвиги), которые можно видеть на следующем графе с пятью вершинами и одиннадцатью дугами:



Таким образом, NIGHT и CLOWN — решения, полученные с помощью кратного анаграммирования. Холден впервые сказал в 1879 г.: «Будет существовать один порядок — и только один — в котором два сообщения будут одновременно получать смысл».

21.3.2. Практическая польза. Этот пример показывает, что кратное анаграммирование двух и более криптотекстов можно трактовать как теоретико-графовую проблему, где число вершин эквивалентно длине текстов. Возможна механизация подсчета допустимых комбинаций. Комбинаторная сложность задачи нахождения пути, проходящего все вершины и не проходящего дважды через одну и ту же вершину, ограничивает длину текстов, которые можно рассматривать таким образом. Если даны шесть текстов с длинами, скажем, 25, 36, 49, 64, 81 или 100, как часто бывало в случае полевых шифров, то кратное анаграммирование может быть быстро проведено при помощи компьютера. Кратное анаграммирование, в частности, важно для перестановок, полученных с помощью решеток и маршрутных транскрипций, так как эти устройства изготовлены фабричным способом и, как правило,

предназначены для краткого использования. Перешифование многоалфавитной подстановки перестановкой (*chiffre à triple clef* (шифр тройного ключа, *фр.*) Керкхоффса) выдерживает контактный метод анаграммирования. Но это не значит, что не существует других способов атаки.

21.3.3. Хассард, Гросвенор, Холден. Перестановку можно также производить не с буквами, а со словами; тогда средством атаки служит кратное анаграммирование слов. Кратное анаграммирование было изобретено (или, по крайней мере, опубликовано) в 1878 г. — за пять лет до Керкхоффса — Хассардом и Гросвенором, двумя издателями газеты Нью-Йорк Трибюн, и — независимо — уже упоминавшимся Холденом из американской военно-морской обсерватории в Вашингтоне. Причиной столь значительных усилий был скандал в американском сенате, вызванный несколькими сотнями зашифрованных телеграмм. Были использованы любительские системы: открытый текст с зашифрованными собственными именами был записан в решетке, причем были использованы четыре такие решетки — с 15, 20, 25 и 30 словами. Телеграммы были независимо дешифрованы, и были получены совпадающие решения, которые гарантировали их правильность и подлинность. Разоблачение этого скандала имело глубокие политические последствия, кроме того, американская публика стала интенсивно снабжаться информацией о тайных кодах и о том, как их взламывать. Возможно, предпочтение, которое американцы отдают криптограммам в качестве приятного времяпрепровождения возникло именно в это время.

В 1914 г. французы под командой полковника Картье получили приятный опыт, когда они противостояли армии немецкого кайзера, применявшей в качестве полевого кода колонную перестановку. Это не было новостью для французов, так как немцы тупо применяли этот метод в огромном объеме для тренировочных сообщений еще в мирное время. Чтобы никто в этом не сомневался, немцы все сигналы маркировали кодовой группой ÜBCHI (Üdingschiffre — тренировочный шифр, *нем.*); поэтому французы называли ее ubchi. Они могли дешифровать такие сообщения кратным анаграммированием по крайней мере, в больших фрагментах (типичная ситуация). Это позволяло им восстанавливать пароль. Первого октября 1914 г. Картье и его помощники (Оливари, Шваб и Фрейс) разослали метод дешифрования этого шифра в различные французские штабы, дав им возможность читать немецкие радиограммы так же быстро, как и самим немцам. Такая ситуация продолжалась до середины ноября 1914 г.

Генералы *кайзеровской армии* тогда допустили ужасную ошибку: они перешли от упрямой, поглощающей время двойной колонной перестановки к простой колонной перестановке, дополнительно перешифрованной системой ВИЖЕНЕРа с ключом ABC, который мог быть дан в заголовке. Это — *иллюзорная сложность* — для отслаивания добавки требовалось лишь взглянуть на частотный профиль — позволяла французам использовать контакты при решении отдельной колонной перестановки, что было легким делом. Такая ситуация продолжалась до мая 1915 г. и сберегла французам много труда.

Несмотря на то, что в газете *Le Matin* была опубликована история успеха французов в октябре 1914 г., немецкая армия снова вернулась к перестановкам в конце 1916 г., применяя в этот раз поворотную решетку. Это продолжалось четыре месяца и, конечно, не создавало французам проблем.

Глава 22

Заключительные замечания

Недостаточная координация в развитии собственных технологических процессов, несовершенные генерация и распределение ключей, недооценка возможных компромиссов при использовании ключей и многие другие причины могут дать удобный случай незаконному дешифровальщику.

Хюттенхайн[1], 1978 *г.*

История криптологии показывает, что незаконный дешифровальщик пирует на ошибках противника (разд. 11.2.5). Ошибки шифрования делаются шифровальщиками. Тактические и стратегические криптографические ошибки встречаются в разведке и органах информации на всех уровнях вплоть до генералов и директоров. Это включает даже вопросы политической организации. Раскол между службами в Германии перед и в течение Второй мировой войны в конечном счете (но не только) был следствием соперничества между Риббентропом, Герингом и Гиммлером. Деление в Особой службе: отдел Pers Z в МИДе (*Auswärtiges Amt*), шифротдел Хи главного командования Вермахта (*Oberkommando der Wehrmacht*), служба B-Dienst ВМФ (*Kriegsmarine*), служба исследований ВВС (*Reichsluftfahrtministerium*) и отдел Amt VI Службы безопасности (*Reichssicherheitshauptamt*) было в высшей степени непродуктивно. Англичане концентрировали свои особые службы с самого начала обострения политической обстановки в Европе под крылом МИДа (*Foreign Office*) в Школе кодов и шифров. В Блетчли Парк даже военные службы не казались плохо работающими, это относится и к секретным службам, МИ.6 (под началом Мензиса) и Американскому отделу O.S.S. (под началом Брюса); все партнеры сидели рядом в Лондонской секции управления Уинстона Черчилля.

Но как немцы, так и англичане (с их доктриной «необходимо знать») сохраняли внутренние барьеры по причинам соблюдения безопасности; эффектом

[1] Доктор Хюттенхайн (26.1.1905 г.–1.12.1990 г.) изучал математику и астрономию в Мюнстере. В 1936 г. начал работать в отделе шифров (*Xu*) Главного Командования Вермахта (OKW); под конец он был главой группы IV (аналитический криптоанализ). После войны с 1956 г. по 1973 г. он руководил Немецкой шифрколлегией.

было то, что ни одно подразделение не могло узнать от других достаточно, чтобы быть полезным, а также позволяло им иногда замалчивать неудачи и провалы.

Конечно, это больше дело историков, чем криптологов — судить, в какой мере криптологические результаты повлияли на войну и мир. Многотомные журналистские документы заключают в себе серьезные обсуждения сенсационных откровений.

22.1. Успехи во взломе шифров

У криптографии есть враги. Генералы и послы не всегда считаются с требованиями криптографов. Они могут чувствовать, что зависимость от шифровальщиков является потерей времени и унизительна, могут сомневаться в честности кого-либо из них. Великий Вольтер пошел так далеко, что назвал взломщиков кодов шарлатанами: «...ceux qui se vantent de déchiffrer une lettre sans être instruit des affaires qu'on y traite ... sont de plus grands charlatans que ceuxqui se vanteraient d'entendre une langue qu'ils n'ont point apprise» (те, кто хвастает, что способны дешифровать письмо, не имея никакой информации об обсуждаемых в нем делах, — бóльшие шарлатаны, чем те, кто хвастались бы языком, которого они не учили, *фр.*).

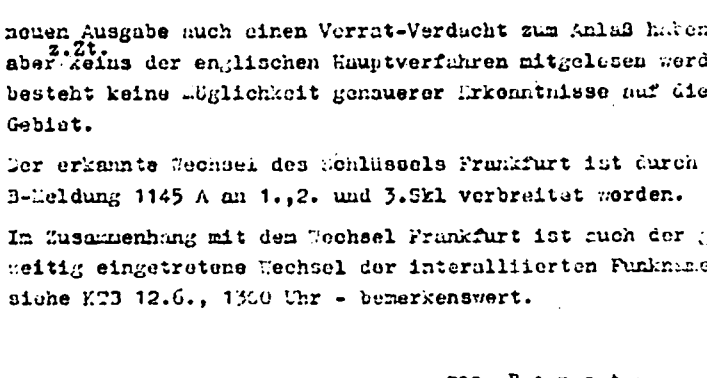
Стали известны имена некоторых удачных взломщиков шифров: в Первую мировую войну англичане Холл, де Грей, Хэй, Хитчингс, Брук-Хант, французы Пэнви, Картье, Судар, американцы Хитт, Чайлдс, Мурмен, Моборн, Ярдли, Мендельсон, итальянец Сакко, австрийцы Ронге, Фигль, Покорны и пруссак Дюбнер. Гораздо больше имен забыто и никогда не было опубликовано. Во Второй мировой войне имеется даже большее число лиц, вовлеченных во взлом кодов (многие лишь в военное время). В недавно изданной книге Хинсли и Стриппа собраны мемуары около 30 бывших сотрудников Блетчли Парк. Некоторые имена становятся известными совершенно случайно, например, криптоаналитиков Новопашенны из Chi-Stelle, Шрёдера из Forschungsamt и Нееба из армии группы Центр. А вот Рейв (см. разд. 19.3.2) остался совершенно незамеченным.

22.1.1. Немецкие службы B-Dienst, Chi-Stelle, Sonderdienst Dahlem. Один из людей, остававшихся в тени, пока Дэвид Кан не открыл его имя широкой публике, был Транов. В прошлом радиотехник военно-морского флота (*Krigsmarine*), он дешифровывал сообщения английского военно-морского флота во время Первой мировой войны и вновь преуспел в этом в 1935 г. Во время Второй мировой войны он был Главным криптоаналитиком немецкой военно-морской службы B-Dienst[2]. Возможно, в результате поражения Германии во Второй мировой войне, так мало было известно о немецких взломщиках кодов (кроме успехов Рорбаха), однако было бы ошибкой полагать, что их не было. В качестве беспристрастного источника Дэвид Кан описал ситуацию се-

[2]Сокращение слова Beobachtungsdienst (служба наблюдения, *нем.*). Эта служба образована из Beobachtungs- и Entzifferungsdienst (шифровальной службы) кайзеровского флота.

редины 1943 г. в службе B-Dienst под управлением опытного и энергичного Транова следующим образом:

«...Служба B-Dienst была в расцвете своих сил, своевременно вскрывая от 5 до 10% своих перехватов для Деница, которые в дальнейшем использовались в тактических целях. Ранняя информация иногда давала ему возможность так передвинуть подводные лодки, чтобы встретить конвой противника посередине группы».


neuen Ausgabe auch einen Verrat-Verdacht zum Anlaß haben. Da
z.Zt.
aber keine der englischen Hauptverfahren mitgelesen werden kann,
besteht keine Möglichkeit genauerer Erkenntnisse auf diesem
Gebiet.

Der erkannte Wechsel des Schlüssels Frankfurt ist durch X-
B-Meldung 1145 A an 1.,2. und 3.Skl verbreitet worden.

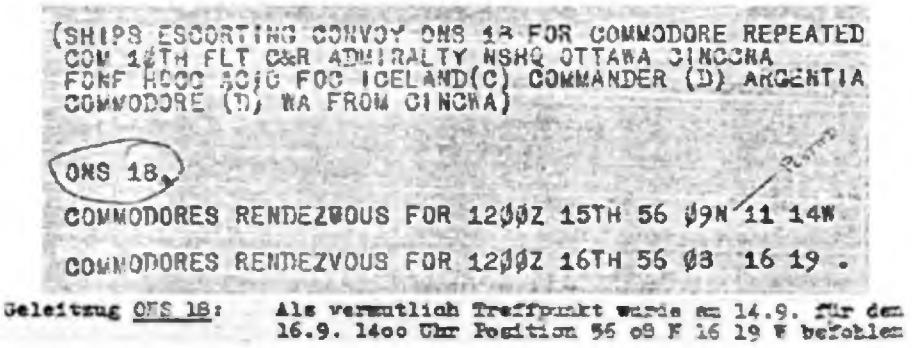
Im Zusammenhang mit dem Wechsel Frankfurt ist auch der gleich-
zeitig eingetretene Wechsel der interalliierten Funknamen -
siehe KTB 12.6., 1300 Uhr - bemerkenswert.

gez. B o n a t z
Kapt.z.See und Mat.Chef

Рис. 161. Фрагмент из военного дневника (15 июня 1943 г.) шефа немецкой военно-морской радиоразведки: «Так как в настоящее время ни один из главных английских шифров не может быть прочитан... Идентифицированное изменение ключа «Франкфурт» было передано X-B-сообщением 1145 A к...»

Действительно, с апреля 1940 г. служба B-Dienst взломала от трети до половины текущих военно-морских шифров, включая Торговый морской код. Когда 20 августа 1940 г. англичане ввели военно-морской шифр № 2 (немецкое кодовое название «Кёльн»), он был частично взломан к концу 1940 г. и полностью в феврале 1941 г., и так продолжалось более двух лет. Таким образом, во время кульминации противолодочной войны военно-морской шифр № 3 (немецкое кодовое название «Франкфурт»), введенный в июне 1941 г. для союзных атлантических конвоев, был скомпрометирован. Эта работа включала отслаивание перешифрования от того, что действительно было кодом. Для этой цели использовались 6 табуляционных машин Холлерита, находящих параллели. К концу 1942 г. 80% сигналов были дешифрованы, но лишь 10% были своевременно использованы. Дениц допускал, что более половины всей его информации пришло из этого источника. Источник пересох лишь тогда, когда возникли подозрения командующего (позднее вице-адмирала) Деннинга, вызванные ENIGMA-источниками, дешифрованными в Блетчли Парк, после чего с 10 июня 1943 г. (рис. 161) два военно-морских флота прекратили использование военно-морского шифра № 3 и начали использование военно-морского шифра № 5. Тем не менее, немцы продолжали дешифровывать некоторые сообщения (рис. 162). Английское командование, как и немецкое, не хотело

верить, что их коды могли быть ненадежными. Бисли возложил ответственность за этот провал на Блетчли Парк, который оскорбительно возразил, но мало что сделал для защиты собственных методов шифрования. Бурке сообщил, что в 1942 г. между Соединенным Королевством и США существовало соперничество, затруднявшее передачу информации. На флоте это прекратилось в октябре 1942 г., а в армии продолжалось до сентября 1943 г.; пока, наконец, не удалось достичь сотрудничества и доверия.


(SHIPS ESCORTING CONVOY ONS 18 FOR COMMODORE REPEATED
COM 12TH FLT C&R ADMIRALTY NSHQ OTTAWA CINCWA
FOKF HQOC ACIC FOC ICELAND(C) COMMANDER (D) ARGENTIA
COMMODORE (D) WA FROM CINCWA)

ONS 18

COMMODORES RENDEZVOUS FOR 1200Z 15TH 56 09N 11 14W

COMMODORES RENDEZVOUS FOR 1200Z 16TH 56 08 16 19 .

Geleitzug ONS 18: Als vermutlich Treffpunkt wurde am 14.9. für den 16.9. 1400 Uhr Position 56 08 N 16 19 W befohlen

Рис. 162. Сравнение радиограммы CINCWA, переданной 14 сентября 1943 г. конвою ONS 18, с расшифровкой немецкого военно-морского флота, осуществленной силами B-Dienst: «...вероятное место встречи... 16.9.1400 ч., позиция 56 08 N 16 19 W»

Был задан вопрос: криптоанализ ли решил Битву в Атлантике? Ровэр и Хинсли сошлись на том, что пока англичане были вынуждены вести оборонительную войну, ситуация была достаточно сбалансированной. И только после середины 1943 г., когда Союзники стали достаточно сильны, чтобы переломить ход противолодочной войны в свою сторону, немецкое подводное командование было парализовано.

Успех службы B-Dienst имел предысторию: когда разразилась война, они уже могли читать английский военно-морской шифр № 1, четырехсимвольный код перешифрования; это стало возможно в результате одного компромисса 1935 г. в абиссинской войне с широким использованием пятисимвольного морского кода, который уже был взломан. Поэтому во время захвата Норвегии в апреле–мае 1940 г. немецкое военно-морское командование всегда представляло себе ситуацию в Британском Адмиралтействе. Это может объяснить удивительно благоприятный для немцев ход событий. Большой успех Транова пришел с опытом Первой мировой войны. Кан цитирует один анонимный источник: «Если кто-то из немецкой интеллигенции когда-либо держал ключи победы во Второй мировой войне, то это был Вильгельм Транов». Но с Гитлером ключей для победы не было.

Криптографы Вермахта, согласно Хюттенхайну, взломали линию связи между французским военным министром и военными округами еще в 1930-х гг. Французские шифры были плохими: числовой код, который оставался неизменным в течение многих лет, перешифровался периодическим

шифром ВИЖЕНЕРа по mod 10 с периодом, варьировавшимся от 7 до 31. Все сообщения могли читаться. И лишь между Парижем и Савойей применялся другой метод, использующий для перешифрования перестановку. В 1938 г. Хи-отдел OKW добился успеха и здесь. Когда началась война 3 сентября 1939 г., французское военное министерство приказало, чтобы этот метод использовался повсеместно. Таким образом, немцы могли читать французские радиосообщения с первого дня войны без задержки, и это объясняет то преимущество, которое они имели в битве за Францию в июне 1940 г. Франция допустила ошибку (разд. 11.1.3), приняв в качестве главного метода шифрования метод, уже применявшийся в течение некоторого времени.

Устойчивый успех Рорбаха между 1942 г. и 1944 г. против ленточного метода шифрования американской дипломатии уже обсуждался в разд. 14.3.6. Так называемые CQ радиосообщения госдепартамента в Вашингтоне, посылавшиеся всем дипломатическим представителям, играли важную роль в создании криптотекст—криптотекст компромисса и даже криптотекст—открытый текст компромисса.

Меньшее значение имел успех, упомянутый в разд. 19.4.1, против румынского военного атташе. Но военные атташе, как правило, являются перспективными целями, обладающими военной и дипломатической информацией, часть которой может быть получена из других источников. В то время как фельдмаршал Роммель в Северной Африке сражался против британской 8-й армии под командованием фельдмаршала Монтгомери осенью 1941 г., немцы из Службы исследований в Берлине проникли в линию связи американского военного атташе в Каире полковника Феллерса (позднее бригадный генерал и секретарь генерала МакАртура) — частично потому, что Феллерс упорствовал в привычке начинать свои сообщения со стереотипов, а частично потому, что итальянцы, тогда еще бывшие в мире с США, имели свою службу *squadra penetrazione*, позаимствовавшую кодовую книгу из американского посольства в Риме, чтобы скопировать ее. Феллерс изо дня в день сообщал используя код BLACK среди прочих вещей планы 8-й армии на следующий день, которые всегда могли быть доставлены Роммелю в течение нескольких часов. Америка была криптологически ненадежным союзником Великобритании. Италия была более осторожной, или, по крайней мере, она так думала: шеф итальянской секретной службы, генерал Аме не снабдил немцев кодовой книгой, ограничившись только дешифрованием. А поскольку немцы тоже записывали криптотекстовые сообщения, они имели совершенный криптотекст—криптотекст компромисс, и могли не только сами восстановить кодовую книгу, но и проверить, насколько их союзники заслуживают доверия. Этот взлом шифра имел в июне 1942 г. катастрофические последствия для конвоя союзников, направлявшегося на Мальту. Блетчли Парк еще раз заподозрил своих друзей, и Феллерс был отправлен в отставку. Тем не менее он был награжден медалью «За отличную службу». На поприще криптографии «соперничать» с ним мог только Мэрфи.

При работе с шифромашинами Hagelin M-209 шифровальщики американской армии показали себя не более дисциплинированными, чем их немецкие

коллеги: для ключей сообщений они часто выбирали шесть начальных букв из имен их подруг, и таким образом одни и те же ключи использовались целый день. В результате Хюттенхайн смог взломать не слишком надежный ежедневный шифр БОФОРТа. Пользу от этого получал Фельдмаршал Эрвин Роммель (Лейберих).

22.1.2. Японские службы. Angō Kenkyū Han. Япония пыталась в 1930-х гг. взломать не только китайские коды, но главным образом американские. Это было не слишком трудно, так как вопреки предостережению Ярдли (разд. 8.5.5), американские дипломатические криптологи были все еще безответственными. При Рузвельте был введен новый код BROWN, но он попал в руки взломщиков сейфов в Загребе и таким образом, вероятно, был скомпрометирован, однако он не был выведен из использования, поскольку в скандал были вовлечены «лишь» криминальные элементы. И Хорнбек писал своему шефу, государственному секретарю Стимсону: «Господин Секретарь, я чувствую, что вполне вероятно, японцы взламывают каждую конфиденциальную телеграмму, которая приходит к нам или идет от нас». Ненадежность американских дипломатических кодов была принята как неизбежность. В этой ситуации неудивительно, что служба дешифрования японского Министерства иностранных дел, возглавляемая Анго Кенкиу Ханом иногда добивалась успеха при дешифровании простейших кодов, например, кода GRAY. Но она потерпела неудачу с кодом BROWN, шифровальная служба адмиралтейства тоже не имела успеха. Ничего кроме налета не могло помочь, и вот под командой капитана Морикава к концу 1937 г. код BROWN и ленточное шифровальное устройство M-138, о появлении которых ничего не было известно японцам, были сфотографированы в американском консульстве в Кобе. Несмотря на это, японцы не могли читать сообщения, зашифрованные машиной M-138. Тогда морская шифровальная служба сосредоточилась на родственном ленточном устройстве CSP 642 американского военного флота (разд. 14.1). Здесь они получали результаты, однако происходило это слишком медленно. Японцы получили код BAMS (разд. 4.4.5) от немцев, которые 10 июля 1940 г. захватили рейдер «Атлантис». В результате японцам осталось только отслоить перешифрование, что они, конечно, сумели сделать.

В противоположном направлении был достигнут гораздо больший успех. Соединенные Штаты, несущие главную тяжесть военных усилий союзников на Тихоокеанском театре боевых действий, добились также наибольших успехов в раскрытии японских радиосообщений. Япония могла чувствовать, что она защищена своим языком, который был чужим и непроницаемым для жителей Запада. Но это был не тот случай. Американцы взламывали японские коды и шифры в 1920-х гг. (Ярдли), в 1930-х гг. (Холтвик) и в 1940-х гг. (Розен). Имелись достоверные сообщения (см. разд. 19.3.2), что также и немецкая сторона неоднократно взламывала линию связи PURPLE.

Немногое известно об американских успехах во взломе советских шифров во время холодной войны. В 1972 г. в течение переговоров об ограничении стратегических вооружений АНБ добилась большого успеха: «но это решение было счастливой случайностью, полученной, возможно, в результате совет-

ской ошибки шифрования» (Кан). Такие вещи могут случаться всегда. И если это случалось более часто, то были веские причины не хвастаться этим.

22.1.3. Главное Разведывательное Управление СССР. Главное Разведывательное Управление Советского Союза имело криптографические успехи, например, против швейцарской дипломатии, работавшей на машинах HAGELIN, а также против Италии, не говоря уже о более малых нациях.

К концу Второй мировой войны число документов, зашифрованных машиной ENIGMA и захваченных Красной армией, выросло до такой степени, что процент советских успехов против ENIGMA-сообщений Вермахта стал значительным. Однако, кажется, в СССР не было машин по взлому кодов, сравнимых с польскими, английскими и американскими бомбами.

В эру холодной войны, согласно Торделлу, Советский Союз успешно работал против роторной машины KW-7, используемой НАТО. В 1992 г. Дэвид Кан нашел русского, жившего тогда в Англии, Виктора Макарова, который работал переводчиком в 16-м управлении КГБ (директор — генерал Андрей Николаевич Андреев) и был знаком с этой работой. От него, а позднее из контактов с Андреевым Кан узнал некоторые детали, среди которых было утверждение, что к концу 1941 г. советский криптоаналитик Сергей Толстой имел успех против японской машины PURPLE. Однако технически полная картина советского криптоанализа пока отсутствует.

22.2. Образ действия незаконного дешифровальщика

> Естественно, что математический интерес представляет только работа незаконного дешифровальщика, обладающего большим опытом. Опыт в дешифровании должен быть завоеван многими годами практики. Поэтому, криптоаналитик, в зависимости от ситуации и наклонностей будет развивать либо более лингвистическое, либо более математическое направление. Создание достаточно сложных методов является результатом коллективной работы разных криптоаналитиков обоих направлений, в каждом из которых, в свою очередь, есть узкие специалисты. В частности, математики нуждаются в специалистах по машинным методам.
>
> *Ханс Рорбах*, 1949 *г.*

> Дешифрование является делом времени, мастерства и настойчивости.
>
> *Чарльз Бэббидж*, 1864 *г.*

> Главное, что необходимо криптографу, это, наверное, настойчивость, аккуратность, стойкость, разумная свобода, некоторый опыт и умение работать с другими.
>
> *Кристофер Моррис*, 1992 *г.*

Поскольку я не являюсь профессиональным дешифровальщиком, говорить о работе незаконного дешифровальщика мне в одно и то же время и трудно,

и легко. Трудно — потому, что я собрал свои впечатления без давления профессиональной среды, а также без трудового пота и слез. Легко — потому, что мне не грозит опасность вскружить себе голову в случае успеха или озлобиться в случае неудачи. Однако, мой математический подход помогал мне систематизировать методы криптологической атаки и защиты.

Как бы то ни было, литература и мои личные контакты показали, что заслуженные криптоаналитики ведут ныне нелегкую жизнь. Например, Режевски после войны вернулся обратно в коммунистическую Польшу и предпочел административную работу университетской карьере. Деннистон своему сыну Робину дал такой совет: «Делай то, что тебе нравится, но не то, что делал я». Робин Деннистон стал издателем. Иногда бывает трудно сохранять абсолютное молчание в течение двадцати, тридцати, сорока лет. Особенно в ситуации, случившейся с Гудом: он проживал в отеле неподалеку от Блетчли Парк, и однажды был вовлечен одним отставным банковским клерком в живой разговор, где тот описывал коммерческую машину ENIGMA, которую его банк использовал в минувшие дни.

22.2.1. Очарование и мука. Работа профессионального криптолога неблагодарна; он не может отметить свой успех публично или с друзьями, и даже его семье нельзя знать, чем он занимается. Он находится в постоянной опасности похищения или шантажа. И такие ограничения сохраняются даже после окончания активной службы.

С другой стороны, Андерсон, который в 1940 г. впервые вошел в шифровальную комнату военно-морского департамента в Вашингтоне, а с 1946 г. работал криптографом в Государственном Департаменте США, где он прослужил более 20 лет, признавался: «Если бы я мог выбирать любую должность, какую захочу, я снова выбрал бы ту, которую имел».

22.2.2. Индивидуальность. Кажется, труднее всего дать общее правило или совет, что следует применять в незаконном дешифровании. Базерье, который был очень удачливым криптоаналитиком, как известно, рекомендовал *changer son fusil d'épaule* (изменить положение ружья на плече, *фр.*), т. е. попытаться найти новую линию атаки, что может помочь лишь людям с достаточно сильным воображением. «Не надо надевать шоры», «не следуйте избитой колеей» — это гораздо легче сказать, чем сделать. Помогут лишь свежие идеи. Пример Алана Тьюринга и Гордона Уэлчмена показывает это: в их неопытности скрывалась их сила. Поэтому они оказались лучше, чем Кнокс, который был гораздо опытнее их, но в то же время и менее отважен. Как команда, Тьюринг и Кнокс были непревзойденными, и даже тандем Тьюринга и Уэлчмена достиг гораздо большего, чем они могли достичь поодиночке.

Одна вещь не случится с незаконным дешифровальщиком: он не будет обескуражен заявлением о сложности работы. Работа поляков оказалась столь успешной, потому что они автоматизировали любой анализ, который съедал слишком много времени при ручной работе после того, как они находили идею. Время, требуемое *Xu* для взлома машины ENIGMA, удалось сократить в шесть раз за счет параллелизации и по крайней мере в 20 раз за счет механизации. BOMBE Уэлчмена могла перебирать триллионы возможных коммутаций

(как гордо отмечал Уэлчмен), потому что лавинное размножение напряжения в относительно простой схеме с обратной связью делается «меньше, чем за тысячную долю секунды».

22.2.3. Стратегии. В принципе, имеется бесконечно много путей криптоаналитической атаки. Далее мы даем лишь грубый обзор стратегий криптоанализа.

22.2.3.1. Незаконное дешифрование в своей чистейшей форме не использует никаких предположений. Чистый криптоанализ не пользуется и не нуждается в лингвисте, так как он математичен по своей природе. Как сказал Дэвид Кан, он функционирует даже для языка, которого незаконный дешифровальщик не знает, например, последний промежуточный текст композиции двух или более шифрований x, скажем, перешифрованный код, кодовая книга которого неизвестна. Чистый криптоанализ является самым подходящим для его выполнения машиной, и может быть представлен в виде компьютерной программы.

Для чистого криптоанализа требуется, как правило, больше текстов, чем для любой атаки, рассмотренной ниже. В некоторых случаях компромисса открытый текст—открытый текст, например, при определении периода многоалфавитного шифра (гл. 18) или согласовании в фазе и наложении различных многоалфавитных шифров с различными начальными установками ключей, так же, как и в случае компромисса криптотекст—криптотекст (гл. 19), чистый криптоанализ выполняет сведение к некоторому промежуточному языку, который является одноалфавитным и, возможно, многосимвольно зашифрованным языком открытого текста.

22.2.3.2. Чистый криптоанализ является частным случаем «атаки одного криптотекста» (или «атаки известного криптотекста»), которая допускает только размышления и предположения о типе языка открытого текста. Конечно, при этом исследуется распределение частот отдельных символов криптотекста. Если оно достаточно близко к одному из нескольких рассматриваемых естественных языков, то можно исключить все методы шифрования, которые выравнивают частоты, в частности, собственно многосимвольные (если они не подделывают частот, как мы видели в разд. 4.1.2) и собственно многоалфавитные шифры; среди оставшихся одноалфавитных шифров имеются функциональные простые подстановки, перестановки и их композиции. Если даже индивидуальные частоты букв близки к частотам букв какого-либо естественного языка, то собственные простые подстановки могут быть исключены. Среди оставшихся шифров будут перестановки, а также многосимвольные шифры, напоминающие перестановки.

Однако если распределение частот отдельных символов в криптотексте выравнено, то (если при этом использование полифонов можно исключить) подтверждается подозрение о многоалфавитном и/или полиграфическом шифре.

В первом случае может помочь чистый криптоанализ — сведением к простой одноалфавитной или к собственно многосимвольной подстановке, которая может быть подвергнута частотному исследованию (гл. 15) отдельных

символов или полиграмм. Последние исследования являются уже лингвистическими по своей природе.

22.2.3.3. Гораздо более лингвистическими являются методы, основанные на частичном или полном компромиссе открытый текст—криптотекст. В качестве исходных точек для отыскания шаблонов они используют вероятные слова или фразы (гл. 13, 14). Существует «атака известного открытого текста» и «атака выбранного открытого текста», различие которых состоит лишь в том, активно или пассивно достигается компромисс. Атака известного открытого текста нуждается в хитром, умном угадывании фрагментов открытого текста, при этом требуется тонкое понимание чувств противника, его образа мышления, знание его идиом и фразеологии, и этому помогает знание не только языка противника, но также знание его окружающей среды. Англичане в Блетчли Парк имели чемпионов по изготовлению материалов для сравнения фрагментов открытого текста и криптотекста, «шпаргалок» (разд. 19.7.1), отметим здесь хотя бы лингвиста Хинсли, урожденную Брет-Смит, и сведущего в лингвистике математика Вайли; были также люди с определенного рода абстрактной способностью находить шаблоны, которые поддерживали Банберизм в Блетчли Парк (разд. 19.6.4.3): шахматный чемпион Александер и разносторонне одаренный филолог-германист Бэйти, урожденная Левер. Ее способности иллюстрирует такой факт: однажды она заметила отсутствие буквы L в длинном фрагменте шифротекста машины ENIGMA. Уже заметить такой факт было неслыханно, но она вывела отсюда существование длинной последовательности букв /l/ в открытом тексте. Это привело к определению установок, взлому шифра и в конечном счете к победе британского флота над итальянцами 28 марта 1941 г. около греческого острова Капе Матапан.

Успех в атаке известного открытого текста требует, однако, чтобы незаконный дешифровальщик обладал всеми возможными разведывательными данными: полученными при помощи допроса пленных, опроса штатских, прослушивания телефонных разговоров, шпионских донесений и особенно дешифрований, выполненных другими дешифровальщиками. Это требование находится в очень большом противоречии с мерами безопасности (доктрина «необходимо знать»), и потому нереалистично — в противном случае во время Второй мировой войны Черчилль сам бы готовил себе «шпаргалки».

Атака выбранного открытого текста, наоборот, нуждается в ловкости для создания компромисса. Неистощимой ловкости. События, о которых надо сообщить, меняются под действием определенных военных действий, подобно артиллерийскому огню в Первой мировой войне и хитрости «erlöschen ist leuchttonne» (погасла светящаяся бочка, *нем.*) (разд. 11.1.3) во Второй мировой войне, чтобы всучить сообщение противнику, подобно яйцу японской кукушки или газетной информации Фигля (разд. 11.1.2).

Третий случай, атака полученного открытого текста, возникает из криптотекст-криптотекст компромисса, если одна из шифросистем уже взломана, и таким образом уже из одного криптотекста можно получить открытый текст. Такое «продолжение взлома» было довольно частой уловкой в Блетчли

Парк, где критические ситуации, приводящие к компромиссу криптотекст—криптотекст, порождались намеренно (разд. 19.4.1).

22.2.3.4. Специальный вид атаки — атака по выбранному шифротексту — может быть использован в случае асимметричных методов шифрования с открытым ключом, когда нужно определить секретный ключ.

22.2.4. Скрытые опасности. Компромисс криптотекст—криптотекст особенно коварен, потому что его легко проглядеть. Его причиной может быть использование многих ключевых сетей при недостаточной дисциплине шифрования (см. разд. 19.4.1) или криптологическая беспечность (дублирование индикатора при шифровании машиной ENIGMA до мая 1940 г., разд. 19.6.1). Особые методы такой атаки рассматривались также в разд. 19.4 и разд. 19.5. Компромисс криптотекст—криптотекст допускает и чистый криптоанализ, его можно провести с помощью мощных компьютеров. Поскольку компромисс криптотекст—криптотекст присущ также и системам с открытым ключом, опасность напоминает об увеличении осторожности.

22.2.5. Дешифрование по слоям. При композиции методов шифрования естественно добиваться отслоения одного шифрования от другого. Это облегчается в случае, если перешифрование производилось над текстом, зашифрованным ранее использовавшимся и в свое время взломанным методом: в таком случае промежуточный текст является известным языком. Это особенно просто, если метод перешифрования уже взломан, так как тогда композиция не более стойка, чем заново введенный метод (перешифрование, разд. 19.6.3.1). Вообще говоря, можно констатировать, что немецкие вооруженные силы не могли лучше обучить своих противников в отношении машины ENIGMA: они вводили усовершенствования малыми дозами, каждый раз через промежуток времени вполне достаточный, чтобы поляки и англичане освоили предыдущий шаг. Например, когда в апреле 1944 г. в Миконосе документы по резервным ручным ключам попали к противнику, метод не был изменен полностью, а изменялся лишь слегка и постепенно, позволяя англичанам приспосабливаться к этим изменениям.

22.2.6. Сила. Криптоанализ в собственном смысле не включает добычу шифродокументов противника и шифроаппаратуры (вплоть до целых машин) путем нелегальной покупки, шпионажа за обычными оффисами, воровством и грабежом или боевыми действиями и рейдами (разд. 11.1.10). Опыт Второй мировой войны полностью подтверждает предостережение Керкхоффса и принцип Шеннона «врагу известна используемая система». Машина SIGABA (ECM Mark II) американской армии была одним из немногих устройств Второй мировой войны, которые не попали в руки врага, и то, возможно, лишь потому, что после высадки союзников в Нормандии война в Европе окончилась меньше, чем через год.

Кроме того, разрушение кабельных каналов связи противника, которое проводили союзники до и в течение высадки в Нормандии, способствовало криптоанализу: его цель заключалась в том, чтобы увеличить объем информации, передаваемой по радиоканалам.

22.2.7. Предотвращение. Что можно предпринять для предотвращения криптоанализа, для защиты каналов связи? Наиболее важным оружием защиты кажется наличие воображения. Необходимо полностью войти в криптоаналитические мысли гипотетического незаконного дешифровальщика и суметь сделать это психологически. Запрещения столь же неуместны, как и высокомерие. Обороняющийся должен не просто обладать воображением, он должен обладать достаточным воображением, чтобы ощутить воображение атакующего.

Мы дадим три примера серьезной беспечности из богатой истории дешифрования машины ENIGMA.

1. Не было абсолютно никакой необходимости строго воздерживаться от применения одного и того же ротора в одной и той же позиции в течение двух последовательных дней, как это делалось в Люфтваффе, а также от применения одного и того же порядка колес дважды в течение одного месяца. Эта искусственная случайность спасала англичан от большой работы по нахождению порядка колес, так как было установлено непрерывное изменение шифров.

2. Было неправильно во что бы то ни стало избегать двух последовательных букв, например, /a/ и /b/ при штекеринге, так как это уменьшает число штепсельных коммутаций, которые нужно было опробовать в бомбах, и даже позволило построить специальную улавливающую схему, названную CSKO («consecutive stecker knock-out», последовательный штекерный нокаут, *англ.*).

3. Совершенно не было необходимости делать входную подстановку (выполняемую штепсельным коммутатором) инволютивной. Ни английская имитация TYPEX, ни японская PURPLE не имели этого «упрощения». Фактически, немцы редко использовали *Uhr*-бокс — искусственное и неуклюжее приспособление (Вклейка M), которое делает подстановку штепсельного коммутатора[3)] (но не шифр ENIGMA) неинволютивной. Кроме того, *Uhr*-бокс должен был часто меняться, вероятно, раз в час (признак, говорящий, что некоторые немецкие криптологи имели серьезные сомнения в надежности ENIGMA, но помочь больше ничем не могли).

Уэлчман назвал ошибки немцев, связанные с машиной ENIGMA, «комедией ошибок». Он писал: «Немецкие ошибки... возникали из-за недостаточного понимания теории шифромашины ENIGMA, из-за слабости машинных операционных процедур, процедур ручного управления сообщениями и процедур радиосети; и прежде всего — от недостатка контроля над всеми процедурами». Далее он упомянул о дублировании индикатора, о «Cillis» и «совете Херивела» (разд. 19.7), о «паркеризме» (обычай немецких составителей оперативных инструкций, который расцвел в 1942 г., и состоял в повторении полных ежемесячных последовательностей дискриминаторов, кольцевых установок, порядка колес или штекеров, например, установки SCORPION были копия-

[3)] *Uhr*-бокс заключал до 10 штекерных пар и имел 40 позиций, 10 из которых (00, 04, 08, ... , 36) сохраняли инволюцию. Скрамблер внутри совершал перестановку с циклическим представлением (01 31 05 39 09 23 17 27 33 19 21 03 29 35 13 11) (00 06 16 26) (02 04 18 24) (12 38 32 22) (14 36 34 20) (07 25) (08 30) (10 28) (15 37).

ми установок PRIMROSE за предыдущий месяц), а также о «неумышленном содействии» немецких штабных работников в составлении шпаргалок. И в заключение он отмечает, что во всех этих ошибках можно обвинять только людей: «сама машина (ENIGMA) могла бы быть неуязвимой, если бы она использовалась надлежащим образом».

22.3. Иллюзорная надежность

Уэлчмен мог бы иронически добавить, — в соответствии с принципом Рорбаха (разд. 11.2.5) — что ни одна машина и ни одна криптосистема не могут все время использоваться надлежащим образом. Криптоаналитик в годы войны и математик в мирное время Ханс Рорбах знал это. Адольф Пашке, советник МИДа и номинальный глава группы Pers Z, знал это так же хорошо. Он решительно возражал против использования машины ENIGMA в дипломатических каналах для передачи не слишком важной информации, например, для визового регулирования. К шифромашинам в МИДе относились с неодобрением. Шифромашина T 52a рассматривалась как ненадежная; фактически, было известно, как ее можно было взломать без большого труда. Кстати, это объясняет успех Бьюрлинга. Сообщения машины T 52e, передаваемые военным атташе по каналам Форин Офис, тоже были дешифрованы людьми Pers Z. Машину T 52 можно было использовать лишь для шифрования несекретных сообщений, передаваемых внутри Германии по кабельным линиям. Одно исключение было сделано в 1944 г. на радиолинии между Берлином и Мадридом, где использовалось SZ 42 для сообщений, имеющих гриф *Geheim* (секретно, *нем.*), но не для сообщений с грифом *Geheim Reichssache* (секретно государственной важности, *нем.*). Это свидетельствует об осторожности, принятой Pers Z. И шеф связи OKW Фельгебель выразил это следующими словами: «*Funken ist Landesverrat*» (радиопередача — это государственная измена, *нем.*).

Но в других отношениях и в другом месте дух иллюзорной надежности процветал. Где бы ни возник шанс более быстрого и менее надежного шифрометода, он имел хорошие перспективы. Мнение жаждущих преобладало. Не считая тех редких случаев, когда индивидуальные ключи вообще использовались, были сделаны надлежащим образом и аккуратно применялись, — очень малое число криптологических систем остались невзломанными между 1900 г. и 1950 г. В случае ENIGMA невзломанными оказались военно-морские ключевые сети «Neptun», «Thetis», «Aegir» и «Sleipnir», правда некоторые из них имели очень мало линий связи или по ним передавалось мало сообщений, и потому они не рассматривались союзниками как достаточно важные.

Надзор за сообщениями своей стороны велся изредка, например, Роулетт обнаружил слабость Фридмановского Конвертора SIGCUM, когда он вступил в действие в январе 1943 г. (разд. 8.8.6). Последствия были бы менее тяжелыми, если бы, например, специальная группа OKW проверяла ENIGMA сообщения недисциплинированной ключевой сети RED Люфтваффе; они нашли бы утечку, полезную англичанам, достаточно рано, чтобы остановить дальнейшую катастрофу.

Но даже надзор не помогает, если он производится неудовлетворительно. Пашке знал, конечно, что одноразовые блокноты AA были сделаны механически с помощью набора из 48 пятизначных счетчиков; после каждого шага печати большинство из них двигалось вперед с нерегулярными интервалами («дополнительный толчок»). В качестве специальной предосторожности последовательно печатаемые листы никогда не брошюровались в один и тот же блок. Казалось, что этого было вполне достаточно, но, как показывает история FLORADORA (см. разд. 8.8.7), это было не так.

22.4. Важность криптологии

Тому, кто читает эту книгу глава за главой с самого начала, вероятно, трудно будет время от времени сдерживать улыбку. История криптологии полна волнующих, забавных, личных историй. Это делает ее притягательной даже для неспециалиста.

Однако мало-помалу на сцене возникают мрачные тени. Сражение под Танненбергом дает первый пример. Вступление в Первую мировую войну Соединенных Штатов Америки было следствием телеграммы от 16 января 1917 г. от немецкого министра Иностранных дел Циммермана своему послу в Мехико фон Эккарту, которая была дешифрована в лондонской «комнате 40» де Греем. Ее содержание — предложение раздуть ссору Мексики с ее северным соседом, — привело к заявлению президента Вильсона, которое заканчивалось словами: «право более драгоценно, чем мир». События Второй мировой войны развивались также на отвратительном фоне. Десятилетия холодной войны выставили на показ жестокость, которую не мог стереть романтизм шпионских романов.

Разговор с криптоаналитиком прошлого о его официальном положении всегда требует такта и осмотрительности. Иногда встречаешься с заносчивостью профессионала, который хочет показать, что он что-то знает, но не показывает, что знает. Как бы то ни было, быть благоразумным — хороший совет для профессионала, как показывает пример Уэлчмена, который подвергся нападкам после опубликования своей книги «The Hut Six Story» (история шестого корпуса, *англ.*).

22.4.1. Сомнения. Криптоанализ воспринимался многими людьми, вовлеченными в него, как тяжелое бремя, не из-за нервного стресса, а из-за конфликта с совестью. Криптология разделяет это бремя не только с другими разделами математики и информатики, которые находятся в опасности злоупотреблений, но и в большой степени с другими науками, такими как физика, химия и биология — достаточно упомянуть ключевые слова ядерная энергия, отравляющий газ, генетическая манипуляция. Это цена, которую наше столетие должно платить за чрезмерный прогресс науки, которую никто не хочет потерять; должны расплачиваться также и сами ученые. Они должны соответствовать высоким требованиям гуманизма. Падение некоторых коммунистических режимов и возрастающее замешательство людей, стоящих перед неограниченными возможностями, внушает надежду, что ученые проявят

проницательность и укажут правильное направление. Таким образом, криптология заслуживает осуждения не более, чем естественные науки. Текст, напечатанный на задней стороне книги Мейера и Матиаша «Криптография» (Wiley, New York, 1981), гласит: «Криптография является единственным известным практическим средством для защиты информации, передаваемой через крупные сети связи, такие как телефонные линии, радио или спутники связи». В другом месте мы читаем: «Криптология прошла путь от тайного искусства до уважаемого раздела информатики». В просторечии это выражается так: «Нынче кодирование и декодирование — игры для всякого, кто может играть». Действительно, оригинальные научные статьи на криптологические темы сегодня находятся не только в немногих специальных журналах или в трудах конференций, но повсюду в компьютерной литературе, особенно в теоретической. Контакт и взаимное оплодотворение происходят главным образом в связи с рождающейся теорией сложности и теорией формальных языков; а кроме того, — из объединенных в единое целое ветвей математики, — теории чисел и комбинаторики.

22.4.2. Новые идеи. Криптология развивает новые направления связанные с теорией информации Шеннона и Реньи, а создание систем с открытыми ключами подтолкнуло и к новым понятиям, таким как асимметрические криптосистемы и аутентификация. Центральным понятием аутентификации является *протокол*, согласованный метод и процедура связи; криптографический протокол между двумя партнерами включает меры, вызванные не только подозрением против некоторого третьего лица, но и взаимным недоверием друг к другу.

Проблема может состоять и в том, как два партнера должны сообщать друг другу секретную информацию так, чтобы не выдавать другие секреты. Другой проблемой является следующее: как два партнера могут шаг за шагом построить отношения доверия без риска обнаружить некоторые секреты. Приложения к ежедневной частной, публичной, политической и экономической жизни очевидны: они могут касаться поведения супругов, держав или фирм. Повседневными примерами являются сертификация держателя пластиковой карточки законным обладателем, а значит, и юридическим собственником которой он является, или лицензионная торговля, где изобретатель должен убедить предполагаемого владельца лицензии в полезности и эффективности его метода, не раскрывая его до подписания контракта.

Такова идея *скрытого доказательства* или «доказательства с нулевым знанием»: партнеру не будет сообщено ничего, что он не мог найти сам.

Это очень старая проблема. Еще во времена Тартальи и Кардано (XVI в.) математики старались держать свои методы в секрете.Они соглашались применить метод (скажем, решения алгебраического уравнения в радикалах) секретно к примерам, предложенным им оппонентом, и потом, спустя короткое время, как кролика из шляпы, представить решение, правильность которого каждый легко мог проверить. Мало-помалу доверие зрителей к правильности и эффективности секретного метода возрастало, и в конце-концов он признавался верным без представления доказательства. Как мы знаем, бедный

Никколо Тарталья не добился успеха в этой игре, и Кардано сумел обмануть его с помощью его же метода решения кубического уравнения.

22.4.3. Дешифрованные тайны природы. Криптоанализ в широком смысле выходит за рамки систем связи. Научное исследование природы часто является криптоанализом ее тайн.

Например, рентгеновская кристаллография протеинов является настоящей криптографической работой. Чтобы определить фазовую функцию, которая принадлежит некоторой заданной амплитудной функции (в трехмерном пространстве), в рентгеновских лучах измеряется рефракционное изображение. Когда фазовая функция (ключевой текст) и амплитудная функция (криптотекст) подобраны так, что они описывают некоторую физическую реальность (открытый текст) с положительной плотностью электронов и правильным числом атомов, то дешифрование (обычно единственное) является успешным. Предположение о структуре молекул, т. е. двойная спираль ДНК, так удачно угаданная Уотсоном и Криком, играют роль вероятных слов. Эта точка зрения уже обсуждалась Тьюрингом и Сэйром в 1950-х гг.

Нахождение иголки в стоге сена — это пример исследования на совпадение. Сглаживание сигналов, рассматриваемое Норбертом Винером в 1950-х годах, особенно отслаивание шума от сигнала имеет параллель в отслаивании добавки от кода в перешифрованном тексте. Еще более серьезной является работа по нахождению образца некоторого до сих пор неизвестного сорта в массе данных, что соответствует продвинутым методам криптоанализа, вроде анализа Фридмана или анализа Кульбака.

Остается узнать, до какой степени здесь уместны понятия, превосходящие теорию информации Шеннона, например, использующие энтропию Кульбака, упомянутую здесь. Первые понятия этой области даны в Приложении *Аксиоматическая теория информации*.

Наконец, существует главное занятие думающего человека: распознавание ситуаций, формирование понятий, разработка абстракций. Это тоже в широком смысле является криптоаналитической работой: это означает отыскание какой-нибудь тайны, чего-нибудь уже тайно существующего. Требуется умение читать между строк и ум. Чистый криптоанализ пытается сделать это без дальнейших знаний и без помощи интуиции, но его результаты ограничены, в то время как имеется богатство Искусственного интеллекта. Там, где он работает, проявляется преимущество автоматической работы. Широкий инструментарий криптоанализа, однако, использует также интуицию, коварство и хитрость. Показать эту взаимосвязь — главная цель этой книги. Криптоанализ как прототип методов, используемых в науке, — это было моим руководящим принципом при написании этой книги. Чарльз Бэббидж сказал (отрывки из книги «Жизнь философа»): «Дешифрование, по моему мнению, это одно из наиболее очаровательных искусств, и я боюсь, что я потратил на него больше времени, чем оно заслуживает». Я не жалел своих усилий, но я надеюсь, что я не потерял свое время.

Приложение

Аксиоматическая теория информации

Логика секретного была зеркальным изображением логики информации.

Колин Бурке, 1994 *г.*

Совершенная надежность обещалась изобретателями криптосистем, особенно криптомашин, во все времена (Базерье: «je suis indéchiffrable» — я недешифруем, *фр.*). В 1949 г. Клод Шеннон дал в рамках своей теории информации ясное определение того, что можно понимать под совершенной надежностью. Мы покажем далее, что можно ввести относящуюся к делу часть теории информации аксиоматически.

Шеннон был знаком с криптоанализом, так как он работал в 1936–1938 гг. в команде Буша, занимавшегося развитием COMPARATORa — машины для определения символьных совпадений. Его работа в лаборатории Белла, начавшаяся в 1940 г., привела к написанию конфиденциального доклада (*A Mathematical Theory of Communication* — математическая теория связи), датированному 1 сентября 1945 г. и содержащему помимо определения энтропии Шеннона (разд. 16.5), основные соотношения, которые рассматриваются в этом Приложении. Опубликован этот доклад был тремя годами позже: *Communication Theory of Secrecy Systems*, Bell System Technical Journal 28, 656–715 (1949).

A.1. Аксиомы аксиоматической теории информации

Целесообразно начать с событий, т. е. множеств $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, ... элементарных событий и с неопределенности на событиях — множествах действительных чисел. Более точно, пусть

$H_{\mathcal{Y}}(\mathcal{X})$ обозначает неопределенность на $\mathcal{X}$, если $\mathcal{Y}$ известно,

$H(\mathcal{X}) = H_{\emptyset}(\mathcal{X})$ обозначает неопределенность на $\mathcal{X}$, если ничего не известно.

А.1.1. Интуитивно очевидные аксиомы для действительно-значной бинарной функции H множеств:

$$0 \leqslant H_{\mathcal{Y}}(\mathcal{X}) \quad \text{(неопределенность неотрицательна).} \tag{0}$$

Если $H_{\mathcal{Y}}(\mathcal{X}) = 0$, мы говорим «$\mathcal{Y}$ однозначно определяет $\mathcal{X}$».

$$H_{\mathcal{Y}\cup\mathcal{Z}}(\mathcal{X}) \leqslant H_{\mathcal{Z}}(\mathcal{X}) \text{ (неопределенность убывает, если известно больше).} \tag{1}$$

При $H_{\mathcal{Y}\cup\mathcal{Z}}(\mathcal{X}) = H_{\mathcal{Z}}(\mathcal{X})$ мы говорим «$\mathcal{Y}$ ничего не говорит об $\mathcal{X}$».

Основная аксиома об аддитивности гласит:

$$H_{\mathcal{Z}}(\mathcal{X}\cup\mathcal{Y}) = H_{\mathcal{Y}\cup\mathcal{Z}}(\mathcal{X}) + H_{\mathcal{Z}}(\mathcal{Y}). \tag{2}$$

Это означает, что неопределенность может быть определена аддитивно, исходя из элементарных событий.

Классическая стохастическая модель для этой аксиоматической теории информации основана на

$$p_X(a) = \Pr\,[X = a],$$

т. е. вероятности того, что случайная величина X принимает значение a, и определяется функциями

$$H_{\emptyset}(\{X\}) = -\sum_{s:p_X(s)>0} p_X(s)\cdot \operatorname{ld} p_X(s),$$

$$H_{\emptyset}(\{X\}\cup\{Y\}) = -\sum_{s,t:p_{X,Y}(s,t)>0} p_{X,Y}(s,t)\cdot \operatorname{ld} p_{X,Y}(s,t),$$

$$H_{\{Y\}}(\{X\}) = -\sum_{s,t:p_{X/Y}(s/t)>0} p_{X,Y}(s,t)\cdot \operatorname{ld} p_{X/Y}(s/t),$$

где по определению $p_{X,Y}(a,b) \stackrel{def}{=} \Pr[(X=a)\wedge(Y=b)]$ и $p_{X/Y}(a/b)$, удовлетворяет правилу Байеса для условных вероятностей:

$$p_{X,Y}(s,t) = p_Y(t)\cdot p_{X/Y}(s/t).$$

Таким образом,

$$-\operatorname{ld} p_{X,Y}(s,t) = -\operatorname{ld} p_Y(t) - \operatorname{ld} p_{X/Y}(s/t).$$

А.1.2. Из аксиом (0), (1) и (2) можно получить следующие свойства, обычно выводимые для классической модели.

Для $\mathcal{Y} = \emptyset$ из (2) следует

$$H_{\mathcal{Z}}(\emptyset) = 0 \qquad \begin{array}{l}\text{(не существует неопределенности}\\ \text{на пустом множестве событий).}\end{array} \tag{2a}$$

Из (1) и (2) следует

$$H_Z(\mathcal{X}\cup\mathcal{Y}) \leqslant H_Z(\mathcal{X}) + H_Z(\mathcal{Y}) \quad \text{(неопределенность субаддитивна).} \tag{3a}$$

Из (0) и (2) выводим, что

$$H_Z(\mathcal{Y}) \leqslant H_Z(\mathcal{X}\cup\mathcal{Y}) \quad \begin{array}{l}\text{(неопределенность увеличивается}\\ \text{с увеличением множества событий).}\end{array} \tag{3b}$$

Из (2) и коммутативности объединения $\cup$ вытекает

$$H_Z(\mathcal{X}) - H_{\mathcal{Y}\cup Z}(\mathcal{X}) = H_Z(\mathcal{Y}) - H_{\mathcal{X}\cup Z}(\mathcal{Y}). \tag{4}$$

Свойство (4) приводит к следующему определению.

Определение. *Взаимная информация от событий $\mathcal{X}$ и $\mathcal{Y}$ при известном событии $\mathcal{Z}$ определяется равенством*

$$I_Z(\mathcal{X}, \mathcal{Y}) \stackrel{\text{def}}{=} H_Z(\mathcal{X}) - H_{\mathcal{Y}\cup Z}(\mathcal{X}).$$

Таким образом, взаимная информация является симметричной и (ввиду (1)) неотрицательной функцией событий $\mathcal{X}$ и $\mathcal{Y}$. Из (2) имеем

$$I_Z(\mathcal{X}, \mathcal{Y}) = H_Z(\mathcal{X}) + H_Z(\mathcal{Y}) - H_Z(\mathcal{X}\cup\mathcal{Y}).$$

Ввиду (4) высказывания «$\mathcal{Y}$ ничего не говорит об $\mathcal{X}$» и «$\mathcal{X}$ ничего не говорит об $\mathcal{Y}$» являются эквивалентными и это выражается равенством

$$I_Z(\mathcal{X}, \mathcal{Y}) = 0.$$

Другими словами это можно выразить так: при известном $\mathcal{Z}$ события $\mathcal{X}$ и $\mathcal{Y}$ являются взаимно независимыми.

В классической стохастической модели эта ситуация описывается следующим образом: равенство

$$p_{X,Y}(s, t) = p_X(s)\cdot p_Y(t)$$

выполняется тогда и только тогда, когда $\mathcal{X}$ и $\mathcal{Y}$ — независимые случайные величины.

Равенство $I_Z(\mathcal{X}, \mathcal{Y}) = 0$ эквивалентно аддитивности H при известном $\mathcal{Z}$.

$$I_Z(\mathcal{X}, \mathcal{Y}) = 0 \text{ тогда и только тогда, когда } H_Z(\mathcal{X}) + H_Z(\mathcal{Y}) = H_Z(\mathcal{X}\cup\mathcal{Y}). \tag{5}$$

А.2. Аксиоматическая теория информации криптосистем

Для криптосистемы **X** событиями в смысле абстрактной теории информации являются множества конечных текстов над алфавитом Z_m. Пусть P — открытый текст (событие), C — криптотекст (событие) и K — ключевой текст (событие)[1]. Неопределенности $H(K)$, $H_C(K)$, $H_P(K)$, $H(C)$, $H_P(C)$, $H_K(C)$, $H(P)$, $H_K(P)$, $H_C(P)$ будем называть эквивокациями (двусмысленностями).

[1] Следуя широко распространенным неправильным обозначениям, в дальнейшем вместо $\{X\}$ будем писать X, вместо $\{X\}\cup\{Y\}$ будем писать X, Y; мы также будем опускать символ $\emptyset$, используемый в качестве нижнего индекса.

А.2.1. Прежде всего, из (1) получаем

$$\begin{aligned} H(K) &\geqslant H_P(K), \qquad H(C) \geqslant H_P(C), \\ H(C) &\geqslant H_K(C), \qquad H(P) \geqslant H_K(P), \\ H(P) &\geqslant H_C(P), \qquad H(K) \geqslant H_C(K). \end{aligned}$$

А.2.1.1. Если криптосистема **X** функциональна, то криптотекст C однозначно определяется открытым текстом P и ключевым текстом K, так что

$$\text{(CRYPT)}\ H_{P,K}(C) = 0, \quad \text{т. е.}\ I_K(P, C) = H_K(C), \quad I_P(K, C) = H_P(C)$$

(«открытый текст и ключевой текст вместе не допускают неопределенности в криптотексте»).

А.2.1.2. Если криптосистема **X** инъективна, то открытый текст P однозначно определяется криптотекстом C и ключевым текстом K, так что

$$\text{(DECRYPT)}\ H_{C,K}(P) = 0, \quad \text{т. е.}\ I_C(K, P) = H_C(P), \quad I_K(C, P) = H_K(P)$$

(«криптотекст и ключевой текст вместе не допускают неопределенности в открытом тексте»).

А.2.1.3. Если криптосистема **X** шеннонова, то ключевой текст K однозначно определяется криптотекстом C и открытым текстом P, так что

$$\text{(SHANN)}\ H_{C,P}(K) = 0, \quad \text{т. е.}\ I_P(C, K) = H_P(K), \quad I_C(P, K) = H_C(K)$$

(«криптотекст и открытый текст не допускают неопределенности для ключевого текста»).

А.2.2. Из (4) сразу вытекают равенства

$$\begin{aligned} H_K(C) + H_{K,C}(P) &= H_K(P), \qquad H_P(C) + H_{P,C}(K) = H_P(K), \\ H_C(P) + H_{C,P}(K) &= H_C(K), \qquad H_K(P) + H_{K,P}(C) = H_K(C), \\ H_P(K) + H_{P,K}(C) &= H_P(C), \qquad H_C(K) + H_{C,K}(P) = H_C(P). \end{aligned}$$

Учитывая (1), получаем следующую теорему.

Теорема 1.

$$\begin{aligned} \text{(CRYPT)}\ &\textit{влечет}\ H_K(C) \leqslant H_K(P), \ H_P(C) \leqslant H_P(K), \\ \text{(DECRYPT)}\ &\textit{влечет}\ H_C(P) \leqslant H_C(K), \ H_K(P) \leqslant H_K(C), \\ \text{(SHANN)}\ &\textit{влечет}\ H_P(K) \leqslant H_P(C), \ H_C(K) \leqslant H_C(P). \end{aligned}$$

А.2.3. Обычно криптоситема **X** является инъективной, так что условие (DECRYPT) выполняется. На рис. 163 графически показаны получающиеся соотношения.

В классических профессиональных криптосистемах обычно нет омофонов, так что условие Шеннона (разд. 2.6.4) выполняется. Одноалфавитные

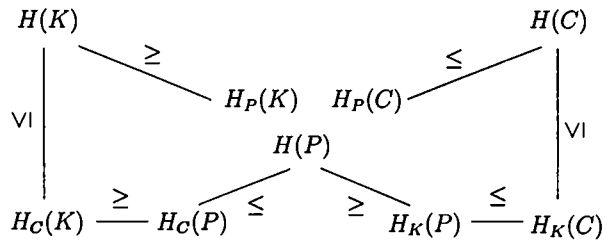


Рис. 163. Эквивокационные соотношения для инъективных криптосистем

простые подстановки и транспозиции, конечно, тривиальны, но ВИЖЕНЕР, БОФОРТ и особенно ВЕРНАМ — серьезные примеры таких классических криптосистем.

Одновременное выполнение любых двух из трех условий (DECRYPT), (CRYPT), (SHANN) имеет далеко идущие последствия ввиду симметрии таких числовых соотношений.

Теорема 2.

$$(\text{CRYPT}) \wedge (\text{DECRYPT}) \quad \textit{влечет} \quad H_K(C) = H_K(P).$$

(«*Неопределенность криптотекста при известном ключевом тексте равна неопределенности открытого текста при известном ключевом тексте.*»)

$$(\text{DECRYPT}) \wedge (\text{SHANN}) \quad \textit{влечет} \quad H_C(P) = H_C(K).$$

(«*Неопределенность открытого текста при известном криптотексте равна неопределенности ключевого текста при известном криптотексте*».)

$$(\text{SHANN}) \wedge (\text{CRYPT}) \quad \textit{влечет} \quad H_P(K) = H_P(C).$$

(«*Неопределенность ключевого текста при известном открытом тексте равна неопределенности криптотекста при известном открытом тексте*».)

На рис. 164 графически показаны соотношения для классических криптотекстов с условиями (CRYPT), (DECRYPT) и (SHANN).

А.3. Совершенные криптосистемы и криптосистемы с независимым ключом

А.3.1. Криптосистема называется совершенной криптосистемой, если открытый текст и криптотекст не зависят друг от друга:

$$I(P, C) = 0.$$

Это равносильно равенствам

$$H(P) = H_C(P), \qquad H(C) = H_P(C)$$

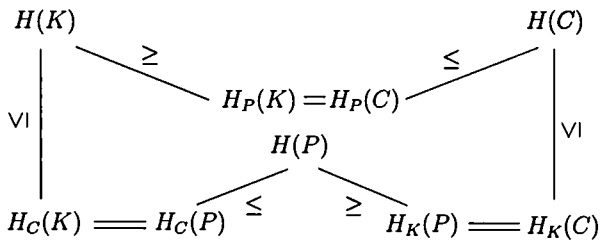


Рис. 164. Эквивокационные соотношения для классических криптосистем

(«без знания ключевого текста знание криптотекста не изменяет неопределенности открытого текста, а знание открытого текста не изменяет неопределенности криптотекста») и, в соответствии с (5), эквивалентно равенству $H(P, C) = H(P) + H(C)$.

А.3.2. Криптосистема называется криптосистемой с независимым ключом, если открытый текст не зависит от ключевого текста:

$$I(P, K) = 0.$$

Это эквивалентно равенствам

$$H(P) = H_K(P), \qquad H(K) = H_P(K)$$

(«без знания криптотекста знание ключевого текста не изменяет неопределенности на открытом тексте, а знание открытого текста не изменяет неопределенности на ключевом тексте») и, в соответствии с (5), эквивалентно равенству $H(K, P) = H(K) + H(P)$.

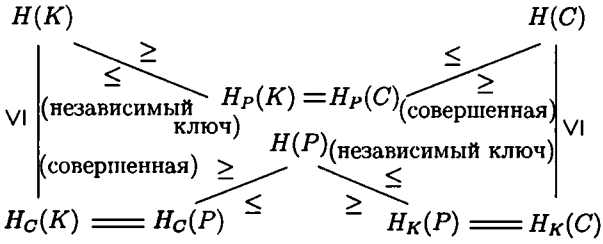


Рис. 165. Эквивокационные соотношения для классических криптосистем с добавочными свойствами *совершенности* и/или *с независимыми ключами*

А.3.3. Шеннон доказал также следующие неравенства:

Теорема 3^K. *В совершенной классической криптосистеме* (*рис.* 165)

$$H(P) \leqslant H(K) \quad \textit{и} \quad H(C) \leqslant H(K).$$

Доказательство. Имеем $H(P) \leqslant H_C(P)$ (так как криптосистема совершенная), $H_C(P) \leqslant H_C(K)$ (из (DECRYPT)), поэтому из теоремы 1

$$H_C(K) \leqslant H(K). \tag{1}$$

Неравенство для $H(C)$ получается аналогичным образом с использованием (CRYPT).

Таким образом, совершенная классическая криптосистема неопределена относительно ключа не меньше, чем неопределена относительно открытого текста и не меньше, чем неопределена относительно криптотекста.

Из (SHANN) $\wedge$ (DECRYPT) с помощью теоремы 1 находим, что $H_C(P) = H_C(K)$. После прибавления $H(C)$ к обеим частям равенства, в соответствии с (2), получаем $H(P, C) = H(K, C)$. Далее, в соответствии с (2), имеем

$$H(K, C) = H(K) + H_K(C).$$

Таким образом,

$$H_K(C) = H(P) - (H(K) - H(C)) = H(C) - (H(K) - H(P)).$$

На рис. 166 этот результат показан графически.

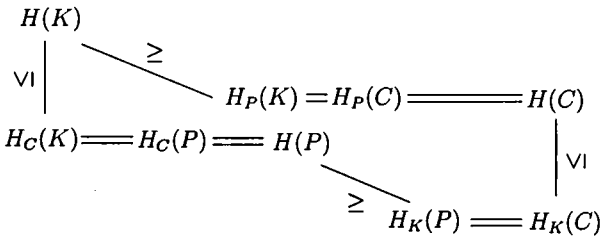


Рис. 166. Числовые эквивокационные соотношения для совершенных классических криптосистем

А.3.4. Циклический сдвиг K, C, P приводит к следующей теореме.

Теорема 3^C. *В классической криптосистеме с независимым ключом*

$$H(K) \leqslant H(C) \quad \textit{и} \quad H(P) \leqslant H(C),$$

так же, как и

$$H_C(P) = H(K) - (H(C) - H(P)) = H(P) - (H(C) - H(K)).$$

А.4. Главная теорема Шеннона

А.4.1. Для классической совершенной криптосистемы с независимым ключом, из теорем 3^K и 3^C сразу получаем, что $H(K) = H(C)$.

А.4.2. Криптосистему с совпадающими $H(K)$ и $H(C)$ будем называть криптосистемой ВЕРНАМа. Примерами являются шифрования со схемами ВИЖЕНЕР, БОФОРТ и, естественно, ВЕРНАМ, а также шифрования, выполняемые с помощью линейных многосимвольных блочных шифров.

В стохастической модели это условие, в частности, выполняется, если криптотекст C и ключевой текст K одновременно состоят из k символов с максимальными $H(K)$ и $H(C)$:

$$H(K) = H(C) = k \cdot \mathrm{ld}\ N.$$

Главная теорема (Шеннон, 1949). *В классической криптосистеме любые два из трех свойств:*

совершенная,

независимого ключа,

типа ВЕРНАМа,

влекут третье.

Доказательство очевидно из рис. 165.

А.4.3. Достаточное условие для того, чтобы классическая криптосистема была совершенной, состоит в том, чтобы она была системой с независимым ключом и была системой типа ВЕРНАМа; эти условия могут быть гарантированы извне. Тогда $H(P) \leqslant H(C) = H(K)$.

В стохастической модели совершенная безопасность требует наряду с $H(P) \leqslant H(K)$, чтобы ключ обладал, по крайней мере, не меньшим числом символов, чем открытый текст. Это означает, что длина каждого сообщения не меньше длины ключа (требование Шайтина, разд. 8.8.4).

Таким образом, совершенная безопасность требует надежного распределения независимого ключа, который обеспечивает каждому символу открытого текста собственный ключевой символ — экстремальное требование, которое часто невозможно выполнить на практике. Несовершенная практическая безопасность гарантируется только временем, требуемым для взлома ключа.

А.4.4. Шеннон рассматривал и дальнейшие свойства криптосистем. Назовем криптосистему идеальной (по Шеннону — строго идеальной), если криптосистема не зависит от ключевого текста:

$$I(K, C) = 0.$$

Это эквивалентно условию $H(K) = H_C(K)$ или $H(C) = H_K(C)$.

В соответствии с Шенноном, идеальные криптосистемы имеют практические дефекты: для совершенной криптосистемы должно выполняться условие $H(K) = H(P)$. Совершенные идеальные криптосистемы обязательно приспособлены к языку открытого текста, который обычно является некоторым естественным языком. В таком случае необходимы достаточно сложные алгоритмы шифрования. К тому же ошибки передачи неизбежно вызывают лавинный эффект. Таким образом, здесь мы действительно имеем практически недостижимый идеал.

А.5. Расстояние единственности

Условие $H_C(P) > 0$ выражает факт, что в известном криптотексте имеется некоторая неопределенность открытого текста. Для классической криптосистемы с независимым ключом (не обязательно совершенной) это означает, по теореме 3^C, что

$$H(K) > H(C) - H(P).$$

Теперь мы используем стохастическую модель со словами открытого текста V^* и со словами криптотекста W^* над некоторым N-элементным множеством символов $V = W$. Мы сосредоточим свое внимание на словах длины k.

Следуя Хеллману (1975 г.), предположим, что N_P и N_C — такие числа, что среди N^k слов длины k число осмысленных, т. е. возможно встречающихся в открытом тексте, только $(N_P)^k$, а число встречающихся криптотекстов лишь $(N_C)^k$. Тогда $N_P \leqslant N$ и $N_C \leqslant N$. Если все эти тексты встречаются равновероятно, то в нашей стохастической модели

$$H(P) = k \cdot \operatorname{ld} N_P, \quad H(C) = k \cdot \operatorname{ld} N_C.$$

Пусть Z обозначает мощность класса методов, т. е. число ключевых слов. Допустим, что все эти ключевые слова встречаются с равной вероятностью. Тогда

$$H(K) = \operatorname{ld} Z.$$

Неравенство в начале параграфа, означающее существование неопределенности, превращается в

$$\operatorname{ld} Z > k \cdot (\operatorname{ld} N_C - \operatorname{ld} N_P),$$

или, при условии, что $\operatorname{ld} N_C > \operatorname{ld} N_P$, в неравенство

$$k < U,$$

где

$$U = \frac{1}{\operatorname{ld} N_C - \operatorname{ld} N_P} \cdot \operatorname{ld} Z.$$

Таким образом, если $k \geqslant U$, то не существует никакой неопределенности. В таком случае U является расстоянием единственности (разд. 12.6).

Если N_C максимально, $N_C = N$, т. е. если все возможные криптотексты встречаются с равной вероятностью, и если $N_P < N$, т. е. открытые тексты принадлежат естественному языку, то условие $\operatorname{ld} N_C > \operatorname{ld} N_P$ обязательно выполняется, и расстояние единственности равно

$$U = \frac{1}{\operatorname{ld} N - \operatorname{ld} N_P} \cdot \operatorname{ld} Z;$$

оно определяется только энтропией Шеннона $\operatorname{ld} N_P$ слов открытого текста. А это, в свою очередь, зависит от криптоаналитической процедуры. Если анализ ограничивается лишь однобуквенными последовательностями, то должна

рассматриваться энтропия Шеннона ld $N_P^{(1)}$, значения которой не сильно различаются в английском, французском и немецком языках и составляет в подсчете Мейера—Матиаша ld $N_P^{(1)} \approx 4.17$, где $N = 26$ и ld $N =$ ld $26 \approx 4.70$. Кроме того, с ld $N_P^{(2)} \approx 3.5$ для частот биграмм и ld $N_P^{(3)} \approx 3.2$ для частот триграмм мы находим

$$U \approx \frac{1}{0.53} \text{ ld } Z \quad \text{для шифров с односимвольным дешифрованием,} \tag{1}$$

$$U \approx \frac{1}{1.2} \text{ ld } Z \quad \text{при биграммном дешифровании,} \tag{2}$$

$$U \approx \frac{1}{1.5} \text{ ld } Z \quad \text{при триграммном дешифровании.} \tag{3}$$

Средняя длина слова открытого текста приблизительно равна 4.5, а соответствующая энтропия Шеннона — примерно ld $N_P^{(w)} \approx 2.6$, так что

$$U \approx \frac{1}{2.1} \text{ ld } Z \text{ при пословном дешифровании.} \tag{w}$$

Энтропия Шеннона английского языка при учете всех, даже грамматических и семантических аспектов, значительно меньше; величина порядка ld $N_P^{(*)} \approx 1.2$ представляется примерно правильной. Это дает расстояние единственности

$$U \approx \frac{1}{3.5} \text{ ld } Z \quad \text{при произвольном дешифровании (см. разд. 12.6).}$$

Для простых (односимвольных) подстановок с $Z = 26!$ мы имеем ld $Z =$ $= 88.38$ (разд. 12.1.1.1). Это приводит к значениям 167, 74, 59, 42 и 25 расстояния единственности, которые подтверждаются практическим опытом. Совершенно аналогичная ситуация имеет место для французского, немецкого, итальянского, русского и родственных им индоевропейских языков.

А.6. Кодовое сжатие

Хотя Шеннон пришел к своей теории информации через занятие криптологическими вопросами во время Второй мировой войны, теория информации в форме, подходящей и полезной для техники связи, не имеет секретных аспектов. Ее практическое значение заключается больше в том, чтобы показать, как увеличить скорость передачи с помощью подходящего кодирования, вплоть до некоторого предела, который соответствует сообщению без какой бы то ни было избыточности — таким будет, например, сообщение P из k символов с максимальной неопределенностью $H(P) = k \cdot \text{ld } N$.

Криптологические результаты, полученные выше, применяются непосредственно к каналам связи. Теоретически, сообщение, стоимость передачи которого равна ld $26 = 4.70$ [бит/сим], может быть сжато кодированием в сообщение, стоимость передачи которого ≈ 1.2 [бит/сим]. Для достижения этого

результата нужны очень сложные схемы кодирования. Простейший случай метода Хаффмана, применяемого для отдельных символов, уменьшает стоимость передачи лишь до 4.17 [бит/сим]. Кодирование Хаффмана для биграмм и триграмм, требующее большого объема памяти, не дает значительного снижения стоимости. В будущем, однако, должно быть достигнуто существенное сокращение избыточности с помощью кодирования Хаффмана для тетраграмм при помощи специализированных интегральных схем.

Совершенно иная ситуация имеет место при передаче изображений. Сжатие, получаемое сравнительно простыми методами, дает замечательные результаты и находит все более широкое практическое применение.

Подобные приложения особенно характерны для после-шенноновской криптологии, так как устранение избыточности из открытого текста можно расценить как шаг к усилению надежности криптосистем.

А.7. Невозможность полного беспорядка

Когда в 1920-х гг. было рекомендовано использование независимых («индивидуальных») ключей, их изготовление не казалось большой проблемой. То что индивидуальный ключ должен быть случайной последовательностью ключевых символов, было интуитивно ясно. После работы Шеннона и особенно после работы Шайтина в 1974 г. все попытки произвести случайную последовательность алгоритмически должны были быть прекращены. Дело в том, что если ключи произведены алгоритмически, то настоящая случайность в ключевых текстах не недостижима, так как в них должен оставаться какой-то порядок — вот в чем вопрос.

Как следствие, стали возникать подозрения, что «псевдослучайные последовательности» с длинным периодом имеют скрытые регулярности, которые должны помогать криптоанализу, хотя конкретных примеров в открытой литературе пока недостаточно. Профессионалы, ответственные за надежность их собственных систем, все больше и больше сталкивались с головной болью, в то время как честолюбивые взломщики кодов всегда могли таить надежду на неожиданные решения.

Странно, но примерно в то же время подобное развитие имело место и в математике. Мирский писал: «Существуют многочисленные теоремы в математике, которые утверждают, грубо говоря, что каждая система некоторого класса обладает большой подсистемой с более высокой степенью организации, чем исходная система».

Приведем несколько примеров.

1. Каждый граф с n вершинами содержит либо полный подграф на k вершинах, либо пустой (без ребер) подграф на k вершинах (k — число Рамсея, например, $k = 6$ для $n = 102$ (Рамсей, 1930 г.)).

2. Каждая ограниченная бесконечная последовательность комплексных чисел включает в себя сходящуюся бесконечную подпоследовательность (К. Вейерштрасс, 1865 г.).

3. Если натуральные числа разбиты на два класса, то по крайней мере один из них содержит арифметическую прогрессию произвольно большой длины (Шур, около 1925 г., Ван дер Варден, 1927 г.).

4. Любое частично упорядоченное множиство из n^2+1 элементов содержит либо цепь длины $n+1$, либо множество из $n+1$ несравнимых элементов (Дилуорс, 1950 г.).

5. Каждая последовательность из n^2+1 натуральных чисел содержит либо монотонно возрастающую, либо монотонно убывающую подпоследовательность длины $n+1$ (Эрдёш, Секереш, 1950 г.).

Между этими и некоторыми другими примерами, казалось, не было никакой связи, пока Эрдёш в 1950 г. не рассмотрел большое число аналогичных задач. В результате он нашел общую теорему, частными случаями которой являются многие из рассмотренных им примеров. Дальнейшие исследования привели к созданию теории, называемой теорией Рамсея. Использование этой теории позволило получить после 1970 г. много тонких математических работ о неупорядоченных системах с упорядоченными подсистемами. Среди них работа Семереди (1975 г.) «О множествах целых чисел, не содержащих k-элементной арифметической прогрессии».

Теорему Рамсея можно интерпретировать как фундаментальную невозможность полного беспорядка и как предостережение криптологам быть осторожнее с применением ключей, выработанных машиной. В данный момент это лишь теоретическая опасность, но все же серьезная.

Мариан Режевски, польский герой дешифрования, выразил это предостережение в 1978 г. в такой форме: «Всякий раз, когда возникает произвольность, возникает также и некоторая регулярность».

Список литературы

Хорошие введения в классическую криптографию для любителей:

Gaines Helen Fouché. *Cryptanalysis*. Dover, New York, 1956 (new ed.).
Smith Laurence Dwight. *Cryptography*. Dover, New York, 1955 (new ed.).
Millikin Donald D. *Elementary Cryptography and Cryptanalysis*. New York, 1943 (3rd ed.).

Введение в криптографию для математически ориентированных читателей:

Sinkov Abraham. *Elementary Cryptanalysis*. Mathematical Association of America, Washington, 1966. (Эта книга, написанная профессиональным криптологом, содержит, однако, не все знания автора.)

Классика криптоанализа:

Friedman William Frederick. *Military Cryptanalysis*. Part I, II, III, IV. Washington, 1938, 1938, 1938, 1942. (Доступна в виде препринта.)

Исчерпывающее историческое исследование криптологии, основанное на литературе, доступной к 1967:

Kahn David. *The Codebreakers*. Macmillan, New York, 1967. (Книга, написанная профессиональным историком с яркостью журналиста, содержит также ссылки на специальную, труднодоступную литературу, в частности, на литературу XIX столетия.)

Работой «первостепенной важности для знания современной криптологии» (Кан) является статья

Rohrbach Hans. *Mathematische und Maschinelle Methoden beim Chiffrieren und Dechiffrieren*. FIAT Review of German Science, 1939–1946: Applied Mathematics, V. 3, Part I, pp. 233–257, Wiesbaden: Office of Military Government for Germany, Field Information Agencies, 1948. (Эту статью в английском переводе Брадфорда Харди можно найти в *Cryptologia* II, 20–37, 101–121 (1978).)

Результаты работ британских взломщиков кодов, а также упоминание о машинах BOMBE и COLOSSUS можно найти в следующих книгах:

Bertrand Gustave. *Enigma ou la plus grande énigme de la guerre* 1939–1945. Librairie Plon, Paris, 1973.
Winterbotham Frederick W. *The Ultra Secret*. Weidenfeld and Nicolson, London, 1974.
Beesly Patrick. *Very Special Intelligence*. Hamish Hamilton, London, 1977.

Lewin Ronald. *Ultra Goes to War*. Hutchinson, London, 1978.
Johnson Brian. *The Secret War*. Methuen, London, 1978.
Rohwer Jürgen, Jäckel Eberhard. *Die Funkaufklärung und ihre Rolle im Zweiten Weltkrieg*. Motorbuch-Verlag, Stuttgart, 1979.
Randell Brian. *COLOSSUS*. In: N. Metropolis et al., *A History of Computing in the Twentieth Century*. Academic Press, New York, 1980.

Детальная и достоверная информация о взломе ENIGMA:
Welchman Gordon. *The Hut Six Story: Breaking the Enigma Codes*. McGraw-Hill, New York, 1982.
Garliński Józef. *Intercept. The Enigma War*. Scribner, New York, 1980.
Kozaczuk Władysław. *Enigma*. Arms and Armour Press, London, 1984 (Polish original edition: W kregu Enigmy, 1979).
Hinsley Francis H. et al., *British Intelligence in the Second World War*. Volumes I–IV, Cambridge University Press, 1979–1988.
Bloch Gilbert. *Enigma avant Ultra*. «Texte definitive», September, 1988. (Английский перевод Cipher A. Deavours содержится в *Cryptologia* XI, 142–155, 227–234 (1987), XII, 178–184 (1988).)
Kahn David. *Seizing the Enigma*. Houghton-Mifflin, Boston, 1991.
Hinsley Francis H., Stripp Alan (eds.). *Codebreakers. The inside story of Bletchley Park*. Oxford University Press, 1993.

Описание жизни и работы Алана Тьюринга:
Hodges Andrew. *Alan Turing: The Enigma*. Simon and Schuster, New York, 1983.

Информация о статистических методах:
Kullback Solomon. *Statistical Methods in Cryptanalysis*. Aegean Park Press, Laguna Hills, CA, 1976.

Криптоанализ в США во время Второй мировой войны:
Rowlett Frank B. *The Story of Magic*. Aegean Park Press, Laguna Hills, CA, 1998.

Описание криптологических устройств и машин:
Türkel Siegfried. *Chiffrieren mit Geräten und Maschinen*. Graz, 1927.
Deavours Cipher A. and Kruh Louis. *Machine Cryptography and Modern Cryptanalysis*. Artech House, Dedham, MA, 1985.

Работы для специалистов, посвященные современной криптографии:
Konheim Alan G. *Cryptography*. Wiley, New York, 1981.
Meyer C. H., Matyas St. M. *Cryptography*. Wiley, New York, 1982.
Brassard G. *Modern Cryptology*. Lecture Notes in Computer Science, V. 325, Springer, Berlin, 1988.
Beker H. and Piper F. *Cipher Systems*. Northwood Books, London, 1982.

Salomaa Arto. *Public-Key Cryptography*. Springer, Berlin, 1990. (Последние две работы затрагивают также криптоанализ машин Хагелина.)

Schneier Bruce. *Applied Cryptography*. Wiley, New York, 1993, 1995 (2nd ed.). (Эта книга содержит протоколы, алгоритмы и тексты программ на языке *C*.)

Goldreich Oded. *Modern Cryptography, Probabilistic Proofs and Pseudo-randomness*. Springer, Berlin, 1999. (Книга, типичная для современной части научной криптологии.)

Вопросы криптологии и гражданских прав обсуждаются в книгах:

Hoffman Lance J. (ed.) *Building in Big Brother*. Springer, New York, 1995.

Denning Dorothy E. R. *Cryptography and Data Security*. Addison-Wesley, Reading, MA, 1983.

Современные элементарные книги:

Bamford James. *The Puzzle Palace*. Penguine Books, New York, 1983.

Kippenhahn Rudolf. *Verschlsselte Botschaften*. Rowohlt, Reinbek, 1997.

Sebag-Montefiore Hugh. *ENIGMA. The Battle for the Code*. Weidenfeld & Nicolson, London, 2000.

Budiansky Stephen. *Battle of Wits*. Simon & Schuster, New York, 2000.

Среди журналов для специалистов можно указать следующие:

Cryptologia. A Quarterly Journal Devoted to Cryptology. Editors: David Kahn, Louis Kruh, Cipher A. Deavours, Brian J. Winkel, Greg Mellen. ISSN 0161-1194. Terre Haute, Indiana.

Journal of Cryptology. The Journal of the International Association for Cryptologic Research. Editor-in-Chief: Gilles Brassard. ISSN 0933-2790. Springer, New York.

В сборнике the Lecture Notes in Computer Science series (Springer, Berlin) под заглавием *Advances in Cryptology* встречаются труды ежегодных конференций *International Cryptology Conference* (CRYPTO), *International Conference on the Theory and Application of Cryptographic Techniques* (EUROCRYPT) и *International Conference on the Theory and Application of Cryptology and Information Security* (ASIACRYPT), организованных при содействии International Association for Cryptologic Research (IACR).

Следующие работы представляют интерес в основном для историков:

Breithaupt 1737; Buck 1772; Hindenburg 1795, 1796; Andres 1799; Klüber 1809; Lindenfels 1819; Vesin de Romanini 1838, 1844; Kasiski 1863; Myer 1866; Koehl 1876; Fleißner von Wostrowitz 1881; Kerckhoffs 1883; Josse 1885; de Viaris 1888, 1893; Valério 1892; Carmona 1894; Gioppi di Türkheim 1897; Bazeries 1901; Myszkowski 1902; Delastelle 1902; Meister 1902, 1906; Schneickert 1900, 1905, 1913; Hitt 1916; Langie 1918; Friedman 1918, 1922, 1924, 1925; Givierge 1925; Lange-Soudart 1925; Sacco 1925, 1947; Figl 1926; Gyldén 1931; Yardley 1931; Ohaver 1933; Baudouin 1939; d'Agapayeff 1939; Pratt 1939; Eyraud 1953; Weiss 1956; Callimahos 1962, Muller 1971; Konheim 1981; Meyer-Matyas 1982.

Более полную библиографию можно найти в книге:

Shulman David. *An Annotated Bibliography of Cryptography*. Garland, New York, 1976.

Работы таких авторов, как Hitt, Langie, Friedman, Givierge, Lange-Soudart, Sacco, Gyldén, Ohaver, Callimahos, Kullback доступны в виде репринтов:

Aegean Park Press, P. O. Box 2837, Laguna Hills, CA 92654–0837, USA.

Книги, переведенные на русский язык:

Брассар Ж. *Современная криптология*. М.: Полимед, 1999.

Кан Д. *Взломщики кодов*. М.: Центрполиграф, 2000.

Саломаа А. *Криптография с открытым ключом*. М.: Мир, 1995.

Шеннон К. *Работы по теории информации и кибернетеке*. М.: ИЛ, 1963.

Шнайдер Б. *Прикладная криптография*. М.: Триумф, 2002.

У. Фридман (сидит) и (стоят слева направо) С. Кульбак, Ф. Роулет, А. Синкоф (Арлингтон Холл, 1944)

Д. Кнокс

Х. Рорбах

Н. Грей

Рис. 167. Знаменитые криптографы XX века

Предметный указатель

Б

В

Г

Д

Л

М

Н

О

П

Р

С

Т

У

Ф

Х

Ц

Ч

Ш

Э

Я

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

R

S

T

U

W

Y

Z

Источники иллюстраций

Kahn, David, The Codebreakers. Macmillan, New York 1967:
Рис. 1, 4, 5, 12, 23, 30, 31, 33, 34, 35, 36, 37, 38, 40, 57

Smith, Laurence Dwight, Cryptography. Dover, New York 1955:
Рис. 3, 16, 24, 53

Bayerische Staatsbibliothek München Рис. 10, 52

Lange, André and E.-A. Soudart. Traité de cryptographie. Paris 1925:
Рис. 26

Crypto AG, Zug, Switzerland:
Рис. 48, 54, 55, 60a, 60b, 62, 64, 65, Вклейки L, O, P

Deavours, Cipher A. and Kruh, Louis, Machine Cryptography and Modern Cryptanalysis. Artech House, Dedham, MA 1985:
Рис. 63, 66, 67, 68

Bundesamt für Sicherheit in der Informationstechnik. Bonn:
Рис. 63r

Public Record Office, London:
Рис. 143, 160a

FRA, Bromma, Sweden:
Рис. 145

National Museum of American History. Washington:
Рис. 160b

Deutsches Museum (Reinhard Krause), Munich, Germany:
Вклейки A, B, C, D, F, G, I, K, N, Q

Russell, Francis, The Secret War. Time-Life Books. Chicago, IL 1981:
Вклейки E, H, M

Решение второй криптографической головоломки на рис 87

$I = m$. Поиск шаблонов приводит к двум шаблонам: 1211234 (KRKKRLH) и 53675 (ULZIU). Среди примерно двух десятков возможных вариантов лишь несколько не выглядят слишком причудливо. Среди этих вариантов /peppery/ и следующий за ним (с 5r6m5) /aroma/ кажутся подходящими и приводят к дальнейшему успеху:

KRKKRLH	PLRU I	OZGK	AYMMGORA U	LYPQ,	QRUAH	ULZ I U
peppery	re m	p	e a r	,	e y	r m
peppery	ream	o p	e a r	,	ea y	aroma
peppery	cream	soup	use a r c	,	ea y	aroma
peppery	cream	soup i	use a	rich,	hea y	aroma
peppery	cream	soup di	used a	rich,	heady	aroma
peppery	cream	soup	diffused a	rich,	heady	aroma

Оглавление

Лист А.

Фестский диск, Крито-Минойский глиняный диск диаметром около 160 мм (XVII до н. э.), покрытый графемами с ясно выраженными пробелами между словами. В настоящее время нет общепринятой дешифровки надписей на этом диске. Определить значение «разорванного и короткого текста невозможно без дополнительной информации» (Фридерихс).

Лист В.

Два шифродиска предположительно XVIII–XIX веков. Верхний диск типа классификатора. На стороне открытого текста вместе с буквами алфавита содержатся разные слоги и часто встречающиеся слова; на стороне криптотекста используются двузначные десятичные числа.

Лист С.

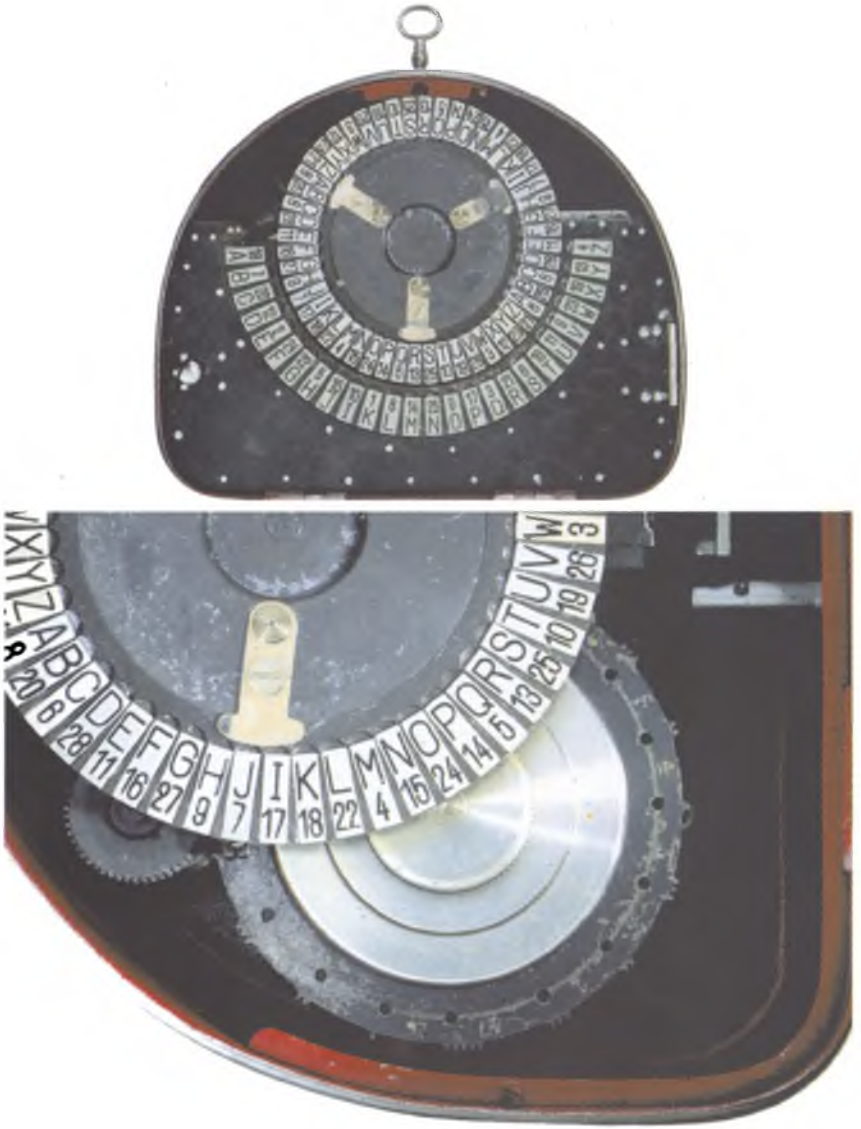
«Криптограф» Вейтстоуна, устройство в виде часов, впервые было показано на Парижской международной выставке в 1867 г. Это многоалфавитное шифровальное устройство: указатель движется по часовой стрелке каждый раз до следующей буквы открытого текста, это движение, в свою очередь, медленно передвигает диск со смешанным алфавитом криптотекста.

Лист D.

Шифровальное устройство американской армии M-94 цилиндрической формы с 25 алюминиевыми дисками диаметром 35 мм с буквами алфавита, выгравированными на ободе, восходит к моделям Джефферсона и Базерье. Введенное в 1922 г. под влиянием Фридмана для военных линий связи низшего уровня, оно широко использовалось до 1942 г.

Лист E.

Ленточный шифратор M-138-T4, использовавшийся армией и флотом США во Второй мировой войне, основанный на проекте Паркера Хитта 1914 г. Двадцать пять подвижных бумажных лент пронумерованы и используются в заранее подготовленном порядке. Такое шифрование криптографически эквивалентно машине M-94.



Лист F.

Шифромашина «Крыха», изобретенная фон Крыхой, Берлин-Шарлоттенбург, около 1926 г., является многоалфавитным шифровальным устройством с фиксированным периодическим ключом длины 442. Нерегулярное движение дисков получается в результате вращения колеса с изменяемым числом зубьев. Несмотря на свою криптографическую простоту, эта изящная машина хорошо продавалась в разных странах.

Лист G.

Шифромашина Хагелина «Криптограф» C-36, построенная фирмой Актиболагет Криптотекник, Стокгольм, в 1936 г., имела взаимообратное шифрование со схемой шифрования БОФОРТа, выполняемое пошагово с помощью магазина, изобретенного Борисом Хагелином. Нерегулярное движение, основанное на использовании ключевых колес с разной калибровкой, а именно с 17, 19, 21, 23 и 25 зубцами, которые дают ключ с периодом 3 900 225. Для чисто механической машины это было большим достижением.

Лист H.

Машина M-209 была усовершенствованием машины Хагелина C-36. По лицензии Хагелина она производилась фабрикой Смит-Корона для армии США; она имела добавочное ключевое колесо с 26 зубцами, которое увеличивало период ключа до 101 405 950. Когда кривошип поворачивался на один оборот, литерные колеса сдвигали зубцы и выступы, которые сдвигали стержни в цилиндрическом магазине; стержни действовали подобно зубцам, которые поворачивают колесо, чтобы напечатать букву криптотекста на рулоне бумаги позади выступа.

Лист I.

Роторная машина ENIGMA, изобретенная Артуром Шербиусом в 1919 г., со световым дисплеем и штепсельным коммутатором (спереди); четырехроторная версия М-4 для немецкого ВМФ, 1944 г. Шифрование производится тремя (из 8) нормальных роторов и одним (из двух) отражающим ротором («греческим» ротором β или γ), введение которых немцами в феврале 1942 г. остановило чтение англичанами сообщений немецких подводных лодок до декабря 1942 г.

Лист K.

Роторы машины ENIGMA: Внутренняя коммутация имеет 26 электрических соединений между контактами на одной стороне и контактами на другой стороне.

Вверху: Ротор I с видимым установочным кольцом.

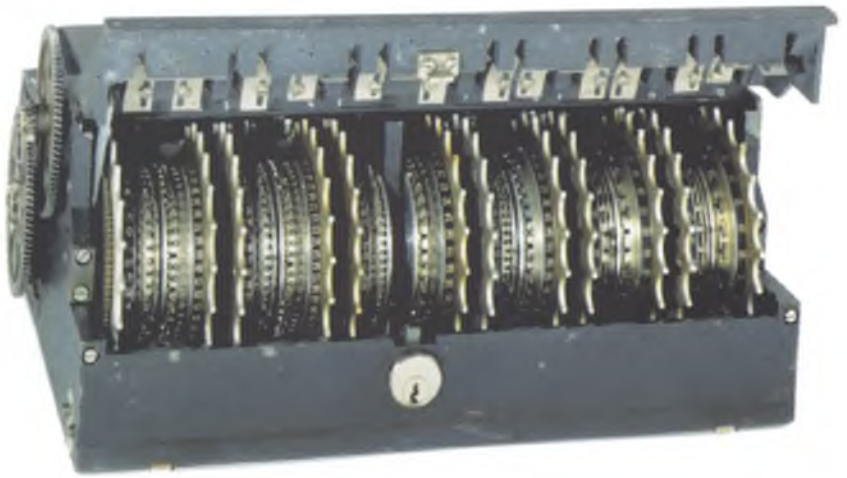
Внизу: Ротор VIII с двумя пазами.

Лист L.

Британская шифромашина TYPEX была усовершенствованной копией немецкой трехроторной машины ENIGMA; она имела два дополнительных ротора (не движущихся во время работы), что делает взлом гораздо более трудным. Эта машина активно применялась на британских линиях связи, а также помогала дешифровать немецкие сообщения после того, как был взломан их ключ. На рисунке изображена машина TYPEX Mark III, серия № 376.

Лист М.

Блок *Uhr* был использован, чтобы штепсельную коммутацию машины ENIGMA немецкого Вермахта заменить на необратимую подстановку, которая легко меняется простым поворотом ручки (предположительно каждый час) выбором одной из 40 позиций. Несмотря на дополнительную надежность, которая при этом гарантируется, такое приспособление широко не использовалось.

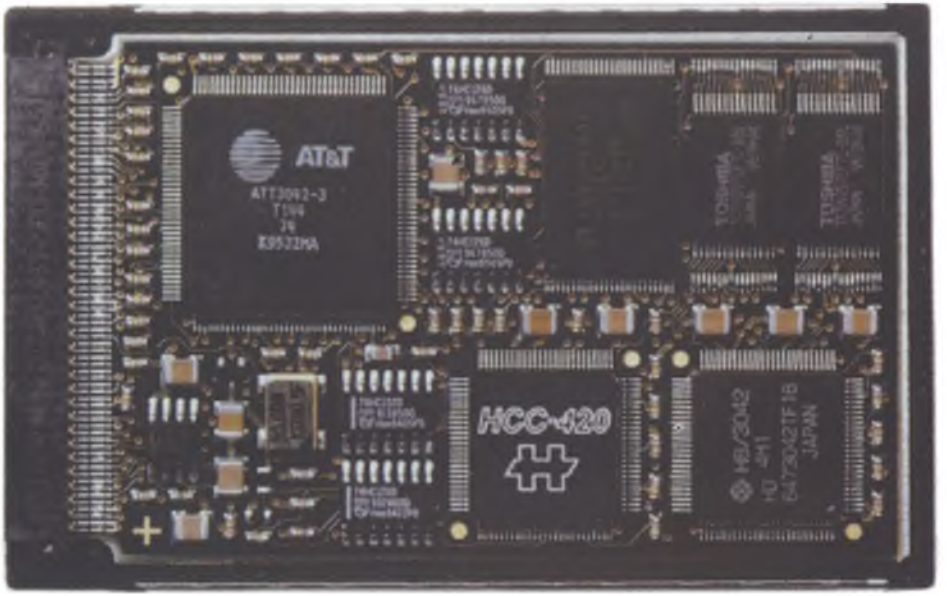


Лист N.

Шифровальная телетайпная машина Лоренца SZ-42, созданная Лоренцом, Берлин, около 1943 г. Шифромашина для телетайпных сигналов Бодо. Англичане называли ее tunny (тунец). Она использовалась на стратегическом уровне — для армейских штабов. 12 ключевых колес с разной калибровкой, имеющих (слева направо) 43, 47, 51, 53, 59, 37, 61, 41, 31, 29, 26 и 23 зубцов с нерегулярными пропусками, производят ключ очень большого периода. Пять пар колес, каждая из которых управляет пятью ВЕРНАМ подстановками 5-битного кода; два колеса («движущие колеса») служат только для нерегулярного движения. Шифры SZ 40/SZ 42 были взломаны англичанами благодаря ошибке шифрования немецкой стороны, а потом стали читаться регулярно с помощью электронных машин COLOSSUS.

Лист О.

Русский одноразовый бумажный блокнот, достаточно маленький, чтобы его можно было вложить в ладонь. Форма цифр в блокноте совпадает с формой цифр в русских пишущих машинках.



Лист Р.

Криптоплата, выпущенная в 1996 г. фирмой Crypto AG, Цуг, Швейцария, используемая в отдельных и работающих в сети компьютерах для обеспечения защиты доступа, секретности информации, целостности информации и защиты от вирусов. Это надежный прибор с очень большим средним временем между сбоями, который может работать без дополнительного питания.

Лист Q.

CRAY-1 S (1979 г.). Суперкомпьютеры CRAY являются потомками известного CRAY-1, спроектированного Крэем (1928–1996 гг.) и использовавшегося с 1976 г., когда его рыночная стоимость состовляла около $ 8 000 000. Суперкомпьютеры содержат большое число интегральных схем, обеспечивающих высокий параллелизм в работе. Они работают на экстремально высокой скорости и нуждаются в сильном охлаждении. Впервые использован для криптоаналитических работ; гражданские версии появились в продаже с некоторыми ограничениями, начиная с 1979 г. Эти серии, продолжаемые CRAY-2, CRAY X-MP, CRAY Y-MP, CRAY C 90, CRAY J 90, привели к CRAY T 90, чья конфигурация T932 включает 32 процессора. Первой массивно-параллельной моделью был CRAY T3D. Более новая модель CRAY T3E (июль 1996) с жидким охлаждением имеет до 2048 процессоров DEC Alpha EV-5 (21164), производительность каждого из которых равна 600 megaflops, а пиковая производительность 1.2 teraflops (1998 г.: T3E-1200T 2.4 teraflops).

Учебное издание

Фридрих Л. Бауэр

Расшифрованные секреты
Методы и принципы криптологии

Заведующий редакцией академик В. И. Арнольд
Зам. зав. редакцией А. С. Попов
Ведущий редактор М. С. Стригунова
Художник М. М. Иванов
Технический редактор Е. В. Денюкова
Оригинал-макет подготовлен А. С. Протопоповым

Подписано к печати 30.10.06. Формат $70 \times 100\ {}^{1}/_{16}$.
Печать офсетная. Объем 17,75 бум. л.
Печ. л. 34,50+1,0 п.л. вкл. Усл.-печ. л. 46,15.
Изд. № 1/9861. Тираж 1500 экз. Заказ 12145

Издательство «Мир»
Министерства культуры и массовых коммуникаций РФ
107996, ГСП-6, Москва, 1-й Рижский пер., 2.

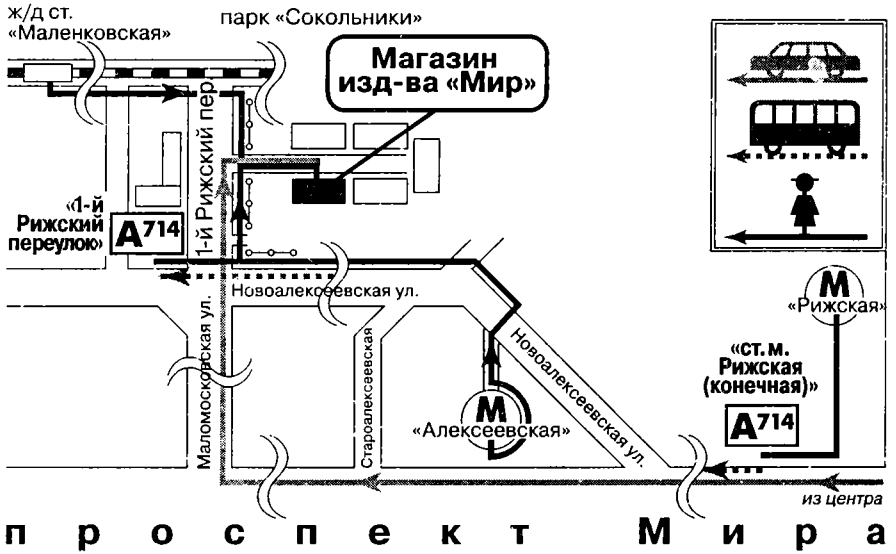
Диапозитивы изготовлены в издательстве «Мир»

Отпечатано с готовых диапозитивов в ОАО «ИПК «Южный Урал»,
г. Оренбург, пер. Свободина, 4.

Книги издательства «Мир»

можно приобрести по издательским ценам по адресу:

Москва, 1-й Рижский пер., д. 2
Тел.: (495) 686-84-44, 686-84-55
Проезд: метро «Рижская», далее авт. 714 до остановки
«1-й Рижский переулок»



ЗАКАЗЫ НАПРАВЛЯТЬ:

обычной почтой: 107996, ГСП-6, Москва, 1-й Рижский пер., д. 2.

факсом: (495) 686-84-44, 686-84-55

по электронной почте: mir-info@mail.ru

по сети Internet: http://www. mir-publishers.net

Ф. БАУЭР

РАСШИФРОВАННЫЕ СЕКРЕТЫ

Методы и принципы криптологии

В отличие от большинства авторов современных книг по криптографии Ф. Бауэр не ограничивается описанием тех или иных методов шифрования. Задача предотвращения несанкционированного доступа к информации рассматривается комплексно – помимо методов шифрования и анализа их стойкости обсуждаются также способы предварительной подготовки информации перед шифрованием и вопросы криптографической дисциплины.

Большим достоинством изложения является обилие фактического материала, иллюстрирующего создание и применение криптографических методов в реальных исторических условиях.

Книга не имеет аналогов на русском языке и может быть рекомендована любителям математики, криптографии и истории, студентам, преподавателям и учащимся старших классов.

ISBN 5-03-003551-6
9 785030 035512