ПРЕДИСЛОВИЕ

В последние три-четыре десятилетия криптография из дисциплины, известной только небольшому числу специалистов по «закрытой» передаче информации, превратилась в мощную науку, преобразующую как мировую экономику, так и повседневную жизнь миллионов «обычных» людей. Речь идет не только о недавно появившихся криптовалютах и «идеальной» бухгалтерской книге блокчейн, но и таких операциях, как расчеты по банковским картам, управление счетами, заказ билетов, покупки, совершаемые через Интернет, и т. д. Криптографические методы позволяют защитить от нечестных или чрезмерно любопытных людей все эти операции, как и разговоры по мобильным телефонам, электронную почту и общение в Интернете.

Эта книга предназначена для исследователей, инженеров, аспирантов и студентов, специализирующихся в области информационных технологий. Все необходимые сведения из теории чисел и теории вероятностей приводятся в книге, но не в виде отдельных разделов, а по мере необходимости. Такой стиль, как мы надеемся, будет удобен читателям книги.

При изложении материала мы старались следовать принципу А. Эйнштейна «Все должно делаться настолько просто, насколько это возможно, но не проще» и соблюдать правило «Кратко и подробно», сформулированное одним из героев известной поэмы А. Твардовского. Поэтому мы не пытались описать всю современную криптографию и стеганографию на строгом математическом уровне и во всей общности, но, как нам кажется, рассмотрели основные идеи и методы без их вульгаризации. При этом в книге содержится точное описание целого ряда практически используемых методов, в том числе криптовалюты биткоин, системы блокчейн и стандартов на криптографические алгоритмы.

Мы надеемся, что эта книга поможет читателям не только понять основные задачи и методы современной криптографии и стеганографии, но и оценить красоту и изящество идей и результатов, лежащих в основе этих наук.

Глава 1. ВВЕДЕНИЕ

Мы начинаем изложение основ криптографии с классической задачи передачи секретных сообщений от некоторого отправителя A к получателю B.

Отправитель сообщений и их получатель могут быть физическими лицами, организациями, какими-либо техническими системами. Иногда об A и B говорят как об абонентах некоторой сети, о пользователях некоторой компьютерной системы или, еще более формально, как об абстрактных «сторонах» (англоязычный термин «party») или «сущностях» (entity), участвующих в информационном взаимодействии. Но чаще бывает удобно отождествлять участников обмена с некоторыми людьми и заменить формальные обозначения A и B на Алиса и Боб.

Предполагается, что сообщения передаются по так называемому «открытому» каналу связи, в принципе доступному для прослушивания некоторым другим лицам, отличным от получателя и отправителя. Такая ситуация возникает при радиопередаче сообщений (например, посредством мобильного телефона) и возможна при использовании даже таких «проверенных» каналов связи, как проволочный телефон, телеграф, да и обычная почта. Особый интерес как средство передачи данных, стремительно завоевывающее лидирующие позиции во всем мире и в то же время чрезвычайно уязвимое с точки зрения возможности несанкционированного доступа третьих лиц, представляет Интернет. В этой среде легко реализуется не только копирование, но и подмена передаваемых сообщений.

В криптографии обычно предполагается, что у лица, передающего сообщения и (или) их принимающего, есть некоторый противник E, который может быть конкурентом в бизнесе, членом преступной группировки, представителем иностранной разведки или даже чрезмерно ревнивой женой, и этот противник может перехватывать сообщения, передаваемые по открытому каналу, и анализировать их. Часто удобно рассматривать противника как некую особу по имени Ева, которая имеет в своем распоряжении мощную вычислительную

технику и владеет методами криптоанализа. Естественно, Алиса и Боб хотят, чтобы их сообщения были непонятны Еве, и используют для этого специальные шифры.

Перед тем как передать сообщение по открытому каналу связи от A к B, A шифрует сообщение, а B, приняв зашифрованное сообщение, расшифровывает его, восстанавливая исходный текст. Важно то, что в рассматриваемой нами в этой главе задаче Алиса и Боб могут договариваться об используемом ими шифре (или, скорее, о некоторых его параметрах) не по открытому каналу, а по специальному «закрытому» каналу, недоступному для прослушивания противником. Такой «закрытый канал» может быть организован при помощи курьеров, или же Алиса и Боб могут обмениваться шифрами во время личной встречи и т.п. При этом надо учитывать, что обычно организация такого закрытого канала и передача по нему сообщений слишком дороги по сравнению с открытым каналом и (или) закрытый канал не может быть использован в любое время. Например, курьерская почта намного дороже обычной, передача сообщений с ее помощью происходит намного медленнее, чем, скажем, по телеграфу, да и использовать ее можно не в любое время суток и не в любой ситуации.

Чтобы быть более конкретными, рассмотрим пример шифра. Так как проблема шифрования сообщений возникла еще в глубокой древности, некоторые шифры связаны с именами известных исторических личностей и в качестве первых примеров обычно используют именно такие шифры. Мы также будем придерживаться этой традиции. Начнем с известного шифра Гая Юлия Цезаря (см., например, [2, 68]), адаптировав его к русскому языку. В этом шифре каждая буква сообщения заменяется на другую, номер которой в алфавите на три больше. Например, А заменяется на Г, Б на Д и т.д. Три последние буквы русского алфавита — Э, Ю, Я — шифруются буквами А, Б, В соответственно. Например, слово ПЕРЕМЕНА после применения к нему шифра Цезаря превращается в ТИУИПИРГ (если исключить букву Ё и считать, что в алфавите 32 буквы).

Последующие римские цезари модифицировали шифр, используя смещение в алфавите на четыре, пять и более букв. Мы можем описать их шифр в общем виде, если пронумеруем (закодируем) буквы русского алфавита числами от 0 до 31 (исключив букву Ё). Тогда

правило шифрования запишется следующим образом:

$$c = (m + k) \bmod 32, \tag{1.1}$$

где m и c — номера букв соответственно сообщения и шифротекста, а k — некоторое целое число, называемое ключом шифра (в рассмотренном выше шифре Цезаря $k = 3$). (Здесь и в дальнейшем $a \bmod b$ обозначает остаток от деления целого числа a на целое число b, причем остаток берется из множества $\{0, 1, \ldots, b-1\}$. Например, $13 \bmod 5 = 3$.)

Чтобы расшифровать зашифрованный текст, нужно применить «обратный» алгоритм

$$m = (c - k) \bmod 32. \tag{1.2}$$

Можно представить себе ситуацию, когда источник и получатель сообщений договорились использовать шифр (1.1), но для того чтобы усложнить задачу противника, решили иногда менять ключ шифра. Для этого Алиса каким-либо образом генерирует число k, передает его Бобу по закрытому каналу связи, и после этого они обмениваются сообщениями, зашифрованными с помощью этого ключа k. Замену ключа можно проводить, например, перед каждым сеансом связи или после передачи фиксированного числа букв (скажем, каждую десятку символов шифровать со своим k) и т.п. В таком случае говорят, что ключ порождается источником ключа. Схема рассмотренной криптосистемы с секретным ключом приведена на рис. 1.1.

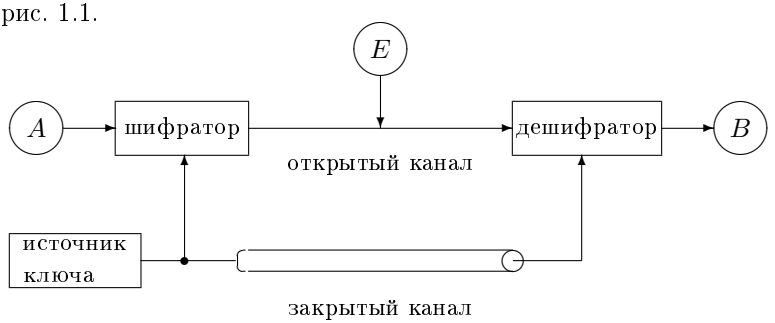

Рис. 1.1. Классическая система секретной связи

Обратимся теперь к анализу действий противника, пытающегося дешифровать сообщение и узнать секретный ключ, иными словами, вскрыть, или взломать шифр. Каждая попытка вскрытия шифра называется атакой на шифр (или на криптосистему). В криптографии принято считать, что противник может знать использованный алгоритм шифрования, характер передаваемых сообщений и перехваченный шифротекст, но не знает секретный ключ. Это называется «правилом Керкхоффса» (см. [68]) в честь ученого, впервые сформулировавшего основные требования к шифрам (A. Kerckhoffs, 1883). Иногда это правило кажется «перестраховкой», но такая «перестраховка» отнюдь не лишняя, если, скажем, передается распоряжение о переводе миллиона долларов с одного счета на другой.

В нашем примере Ева знает, что шифр был построен в соответствии с (1.1), что исходное сообщение было на русском языке и что был передан шифротекст ТИУИПИРГ, но ключ Еве не известен.

Наиболее очевидная попытка дешифровки — последовательный перебор всех возможных ключей (это так называемый метод «грубой силы» (brute-force attack)). Итак, Ева перебирает последовательно все возможные ключи $k = 1, 2, \ldots$, подставляя их в алгоритм дешифрования и оценивая получающиеся результаты. Попробуем и мы использовать этот метод. Результаты дешифрования по (1.2) при различных ключах и шифротексте ТИУИПИРГ сведены в табл. 1.1. В большинстве случаев нам достаточно было дешифровать две–три буквы, чтобы отвергнуть соответствующий ключ (из-за отсутствия слова в русском языке, начинающегося с такого фрагмента).

Таблица 1.1. **Дешифровка слова ТИУИПИРГ путем перебора ключей**

k	m	k	m	k	m	k	m
1	СЗТ	9	ЙЯ	17	БЧ	25	ЩП
2	РЖС	10	ИЮЙ	18	АЦБ	26	ШОЩ
3	ПЕРЕМЕНА	11	ЗЭИ	19	ЯХА	27	ЧН
4	ОДП	12	ЖЬ	20	ЮФ	28	ЦМ
5	НГ	13	ЕЫ	21	ЭУ	29	ХЛЦ
6	МВ	14	ДЪ	22	Ь	30	ФК
7	ЛБМ	15	ГЩ	23	Ы	31	УЙ
8	КАЛАЗ	16	ВШГ	24	Ъ	32	ТИУИПИРГ

Из табл. 1.1 мы видим, что был использован ключ $k = 3$ и зашифровано сообщение ПЕРЕМЕНА. Причем для того чтобы про-

верить остальные возможные значения ключа, нам не требовалось дешифровать все восемь букв, а в большинстве случаев после анализа двух–трех букв ключ отвергался (только при $k = 8$ надо было дешифровать пять букв, зато при $k = 22, 23, 24$ хватало и одной, т.к. в русском языке нет слов, начинающихся с Ь, Ъ, Ы).

Из этого примера мы видим, что рассмотренный шифр совершенно нестоек: для его вскрытия достаточно проанализировать несколько первых букв сообщения и после этого ключ k однозначно определяется (и, следовательно, однозначно расшифровывается все сообщение).

В чем же причины нестойкости рассмотренного шифра и как можно было бы увеличить его стойкость? Рассмотрим еще один пример. Алиса спрятала важные документы в ячейке камеры хранения, снабженной пятидекадным кодовым замком. Теперь она хотела бы сообщить Бобу комбинацию цифр, открывающую ячейку. Она решила использовать аналог шифра Цезаря, адаптированный к алфавиту, состоящему из десятичных цифр:

$$c = (m + k) \bmod 10. \tag{1.3}$$

Допустим, Алиса послала Бобу шифротекст 26047. Ева пытается дешифровать его, последовательно перебирая все возможные ключи. Результаты ее попыток сведены в табл. 1.2.

Таблица 1.2. **Дешифровка сообщения 26047 путем перебора ключей**

k	m	k	m
1	15936	6	60481
2	04825	7	59370
3	93714	8	48269
4	82603	9	37158
5	71592	0	26047

Мы видим, что все полученные варианты равнозначны и Ева не может понять, какая именно комбинация истинна. Анализируя шифротекст, она не может найти значения секретного ключа. Конечно, до перехвата сообщения у Евы было 10^5 возможных значений кодовой комбинации, а после — только 10. Однако важно отметить то, что в данном случае всего 10 значений ключа. Поэтому при таком ключе

(одна десятичная цифра) Алиса и Боб и не могли рассчитывать на бóльшую секретность.

В первом примере сообщение — текст на русском языке, поэтому оно подчиняется многочисленным правилам, различные буквы и их сочетания имеют различные вероятности и, в частности, многие наборы букв вообще запрещены (это свойство называется избыточностью текста). Поэтому-то и удалось легко подобрать ключ и дешифровать сообщение, т.е. избыточность позволила «взломать» шифр. В противоположность этому, во втором примере все комбинации цифр допустимы. «Язык» кодового замка не содержит избыточности. Поэтому даже простой шифр, примененный к сообщениям этого языка, становится невскрываемым. В классической работе К. Шеннона [28] построена глубокая и изящная теория шифров с секретным ключом и, в частности, предложена «правильная» количественная мера избыточности. Мы кратко коснемся этих вопросов в главе 7, а в главе 8 будут описаны современные шифры с секретным ключом.

Описанная в приведенных примерах атака называется атакой по *шифротексту*. Но часто на шифр может быть проведена атака по *известному тексту*. Это происходит, если Ева получает в свое распоряжение какие-либо открытые тексты, соответствующие ранее переданным зашифрованным. Сопоставляя пары «текст–шифротекст», Ева пытается узнать секретный ключ, чтобы с его помощью расшифровывать все последующие сообщения от Алисы к Бобу.

Можно представить себе и более «серьезную» атаку — атаку по *выбранному тексту*, когда противник пользуется не только предоставленными ему парами «текст–шифротекст», но может и сам формировать нужные ему тексты и шифровать их с помощью того ключа, который он хочет узнать. Например, во время Второй мировой войны американцы, подкупив охрану, выкрали шифровальную машину в японском посольстве на два дня и имели возможность подавать ей на вход различные тексты и получать соответствующие шифровки. (Они не могли взломать машину с целью непосредственного определения заложенного в нее секретного ключа, т.к. это было бы замечено и повлекло бы за собой смену всех ключей.)

Может показаться, что атаки по известному и выбранному тексту надуманы и далеко не всегда возможны. Отчасти это так. Но разработчики современных криптосистем стремятся сделать их неуязвимыми даже и по отношению к атакам по выбранному тексту, и

на этом пути достигнуты значительные успехи. Иногда считается, что более надежно использовать шифр, противостоящий атаке по выбранному тексту, чем организационно обеспечивать неосуществимость такой атаки, хотя наиболее осторожные пользователи делают и то, и другое.

Итак, мы познакомились с основными героями криптографии — Алисой, Бобом и Евой и с важными понятиями этой науки — шифром, ключом, атакой, открытым и защищенным каналом. Заметим, что с последним понятием связан один интригующий факт — возможно построение надежных криптосистем без защищенного канала! В таких системах Алиса и Боб вычисляют секретный ключ так, что Ева не может этого сделать. Это открытие было сделано в основополагающих работах Диффи, Хеллмана и Меркля (см., например, [48]) в 1976 году и открыло новую эру в современной криптографии. Большая часть этой книги будет связана именно с такими системами, называемыми криптосистемами с открытым ключом.

Задачи и упражнения

1.1. Определить ключи шифра Цезаря, если известны следующие пары открытый текст – шифротекст:

а. АПЕЛЬСИН – САЦЬНВЩЮ,

б. АБРИКОС – ЫЬЛГЕЙМ.

1.2. Расшифровать следующие сообщения, зашифрованные шифром Цезаря с неизвестным ключом k, $0 < k < 32$:

а. ФХНЗКЧ,

б. ЦЩЕБФ.