Глава 8
Безопасность в сетях

В первые десятилетия своего существования компьютерные сети использовались, в первую очередь, университетскими исследователями для обмена электронной почтой и сотрудниками корпораций для совместного пользования принтеров. В таких условиях вопросы безопасности не привлекали большого внимания. Однако теперь, когда миллионы обычных граждан пользуются сетями для управления своими банковскими счетами, заполнения налоговых деклараций, приобретают товары в интернет-магазинах, и обнаруживаются многие недостатки, проблема сетевой безопасности становится очень серьезной. В этой главе мы рассмотрим вопросы безопасности сетей с различных точек зрения, укажем на подводные камни и обсудим различные алгоритмы и протоколы, позволяющие повысить безопасность сетей.

Тема безопасности является довольно обширной и включает в себя множество вопросов, связанных с различными человеческими грехами. В простейшем виде безопасность — это гарантия, что любопытные личности не смогут читать, или, что еще хуже, изменять сообщения, предназначающиеся другим получателям. Безопасность — это пресечение попыток получения доступа к удаленным службам теми пользователями, которые не имеют на это прав. Система безопасности должна позволять определять, написано ли сообщение «Оплатите счета до пятницы» налоговой службй , или же это фальсификация. Кроме того, системы безопасности решают проблемы, связанные с перехватом и повторным воспроизведением сообщений и с людьми, пытающимися отрицать, что они посылали данные сообщения.

Большинство проблем безопасности возникает из-за злоумышленников, пытающихся извлечь какую-либо пользу для себя или причинить вред другим. Несколько наиболее распространенных типов нарушителей перечислено в табл. 8.1. Из этого списка должно быть ясно, что задача обеспечения безопасности сетей включает в себя значительно больше, нежели просто устранение программных ошибок. Часто стоит задача перехитрить умного, убежденного и иногда хорошо финансируемого противника. Также очевидно, что меры, способные остановить случайного нарушителя, мало повлияют на серьезного преступника. Статистика, собираемая полицией, говорит о том, что большинство атак предпринимается не извне (людьми, прослушивающими линии связи), а изнутри — завистливыми или недовольными чем-либо людьми. Следовательно, системы безопасности должны учитывать и этот факт.

В первом приближении проблемы безопасности сетей могут быть разделены на четыре пересекающиеся области: секретность, аутентификация, обеспечение строгого выполнения обязательств и обеспечение целостности. Секретность (конфиденциаль-

ность) означает предотвращение попадания информации в нечистоплотные руки неавторизованных пользователей. Именно это обычно приходит в голову при упоминании безопасности сетей. Аутентификация позволяет определить, с кем вы разговариваете, прежде чем предоставить собеседнику доступ к секретной информации или вступить с ним в деловые отношения. Проблема обеспечения строгого выполнения обязательств имеет дело с подписями. Как доказать, что ваш клиент действительно прислал электронной почтой заказ на десять миллионов винтиков с левосторонней резьбой по 89 центов за штуку, если впоследствии он утверждает, что цена была 69 центов? Наконец, контроль целостности имеет дело с тем, как можно быть уверенным, что принятое вами сообщение не модифицировано по пути злоумышленником и не подделано?

Таблица 8.1. Типы нарушителей и цели их действий

Нарушитель	Цель
Студент	Прочитать из любопытства чужие письма
Хакер	Проверить на прочность чужую систему безопасности; украсть данные
Торговый агент	Притвориться представителем всей Европы, а не только Андорры
Бизнесмен	Разведать стратегические маркетинговые планы конкурента
Уволенный сотрудник	Отомстить фирме за увольнение
Бухгалтер	Украсть деньги компании
Биржевой брокер	Не выполнить обещание, данное клиенту по электронной почте
Аферист	Украсть номера кредитных карт для продажи
Шпион	Узнать военные или производственные секреты противника
Террорист	Украсть секреты производства бактериологического оружия

Все эти аспекты (секретность, аутентификация, обеспечение строгого выполнения обязательств и обеспечение целостности) встречаются и в традиционных системах, но с некоторыми существенными отличиями. Секретность и целостность достигаются с помощью заказных писем и хранения документов в несгораемых сейфах. Сегодня ограбить почтовый поезд значительно сложнее, чем во времена Джесси Джеймса.

Кроме того, людям обычно несложно отличить на глаз оригинальный бумажный документ от фотокопии. В качестве проверки попробуйте сделать фотокопию настоящего банковского чека. В понедельник попытайтесь обналичить настоящий чек в вашем банке. А во вторник попробуйте сделать то же самое с фотокопией. Проследите за разницей в поведении банковского служащего. Увы, но оригинал и копия электронных чеков неотличимы друг от друга. Понадобится некоторое время, прежде чем банки привыкнут к этому.

Люди опознают друг друга разными способами, например по лицам, голосам и почеркам. Доказательства подлинности бумажных документов обеспечиваются подписями на печатных бланках, печатями, рельефными знаками и т. д. Подделка обычно может быть обнаружена специалистами по почерку, бумаге и чернилам. При работе с электронными документами все это недоступно. Очевидно, требуются другие решения.

Прежде чем перейти к обсуждению сути самих решений, имеет смысл потратить несколько минут и попытаться определить, к какому уровню стека протоколов относится система сетевой безопасности. Вероятно, какое-то одно место для нее найти сложно. Каждый уровень должен внести свой вклад. На физическом уровне с подслушиванием можно бороться за счет помещения передающих кабелей (или, что еще лучше, оптического волокна) в герметичные трубы, наполненные инертным газом под высоким давлением. Любая попытка просверлить трубу приведет к утечке части газа из трубы, в результате давление снизится, и это послужит сигналом тревоги. Подобная техника применяется в некоторых военных системах.

На канальном уровне пакеты, передаваемые по двухточечной линии, могут зашифровываться при передаче в линию и расшифровываться при приеме. Все детали этого могут быть известны только канальному уровню, причем более высокие уровни могут даже не догадываться о том, что там происходит. Однако такое решение перестает работать в том случае, если пакету нужно преодолеть несколько маршрутизаторов, поскольку при этом пакет придется расшифровывать на каждом маршрутизаторе, что сделает его беззащитным перед атаками внутри маршрутизатора. Кроме того, такой метод не позволит защищать отдельные сеансы, требующие защиты (например, осуществление покупок в интернет-магазинах), и при этом не защищать остальные. Тем не менее этот метод, называемый **шифрованием в канале связи (link encryption)**, легко может быть добавлен к любой сети и часто бывает полезен.

На сетевом уровне могут быть установлены межсетевые экраны, позволяющие отвергать подозрительные пакеты, приходящие извне. К этому же уровню относится IP-защита.

На транспортном уровне можно зашифровать соединения целиком, от одного конца до другого. Максимальную защиту может обеспечить только такое сквозное шифрование.

Наконец, проблемы аутентификации и обеспечения строгого выполнения обязательств могут решаться только на прикладном уровне.

Итак, очевидно, что безопасность в сетях — это вопрос, охватывающий все уровни, именно поэтому ему посвящена отдельная глава.

Несмотря на то что эта глава большая, важная и содержит множество технических описаний, в настоящий момент вопрос сетевой безопасности может показаться несколько неуместным. Ведь хорошо известно, что большинство «дыр» в системах безопасности возникают из-за неумелых действий персонала, невнимательного отношения к процедурам защиты информации, недобросовестности самих сотрудников защищаемого объекта, многочисленные ошибки реализации, которые делают возможным проникновение неавторизованных пользователей и атаки с применением методов так называемой «социальной инженерии», когда клиентов обманом заставляют раскрыть свой пароль. Доля проблем, возникающих из-за преступников-интеллектуалов, прослушивающих линии связи и расшифровывающих полученные данные, сравнительно мала. Посудите сами: человек может прийти в совершенно произвольное отделение банка с найденной на улице магнитной кредитной карточкой, посетовать на то, что он забыл свой PIN-код, а бумажку, на которой он написан, съел, и ему тотчас «напомнят» секретный шифр (чего только ни сделаешь ради добрых отношений с клиентами). Ни одна криптографическая система в мире здесь не поможет. В этом

смысле книга Росса Андерсона (Anderson, 2008a) действительно ошеломляет, так как в ней приводятся сотни примеров того, как системы безопасности в самых различных производственных областях не справлялись со своей задачей. И причиной этих неудач была, мягко говоря, неряшливость в ведении дел или простое пренебрежение элементарными правилами безопасности. Тем не менее техническое основание, на котором строится электронная коммерция, когда отсутствуют проблемы с остальными факторами, это криптография.

На всех уровнях, за исключением физического, защита информации в сетях базируется на принципах криптографии. Поэтому мы начнем изучение систем безопасности с детального рассмотрения основ криптографии. В разделе 8.1 «Криптография» мы изучим базовые принципы. Далее до раздела «Управление открытыми ключами» (разделы 8.2–8.5) будут рассмотрены некоторые классические алгоритмы и структуры данных, применяемые в криптографии. После этого мы свяжем теорию с практикой и посмотрим, как все эти концепции применяются в компьютерных сетях. В конце этой главы будут приведены некоторые мысли, касающиеся технологии и социальных вопросов.

Прежде чем приступить к изложению, позвольте назвать вопросы, о которых *не* пойдет речь в этой главе. Подбирая материал, мы старались сконцентрировать внимание на вопросах, связанных именно с компьютерными сетями, а не с операционными системами или приложениями (хотя провести четкую грань зачастую бывает довольно трудно). Например, ничего не говорится об авторизации пользователя с использованием биометрии, защите паролей, атак, связанных с переполнением буферов, вирусах типа «Троянский конь», получении доступа путем обмана, внедрении кода путем межсайтового скриптинга, вирусах, червях и т. п. Обо всем этом очень много говорится в главе 9 книги «Современные операционные системы» (Tanenbaum, 2007). Все желающие узнать о вопросах обеспечения безопасности на уровне системы могут обратиться к этой книге.

Итак, в путь.

8.1. Криптография

Слово **криптография** (**cryptography**) происходит от греческих слов, означающих «скрытное письмо». У криптографии долгая и красочная история, насчитывающая несколько тысяч лет. В данном разделе мы всего лишь кратко упомянем некоторые отдельные моменты в качестве введения к последующей информации. Желающим ознакомиться с полной историей криптографии рекомендуется (Kahn, 1995). Для получения всестороннего представления о текущем положении дел см. (Kaufman и др., 2002). С математическими аспектами криптографии можно познакомиться, прочитав книгу (Stinson, 2002). Менее формальным (с математической точки зрения) языком ведется изложение в (Burnett и Paine, 2001).

С профессиональной точки зрения понятия «шифр» и «код» отличаются друг от друга. **Шифр** (**cipher**) представляет собой посимвольное или побитовое преобразование, не зависящее от лингвистической структуры сообщения. **Код** (**code**), напротив, заменяет целое слово другим словом или символом. Коды в настоящее время

не используются, хотя история у них, конечно, славная. Наилучшим считается код, использовавшийся американскими войсками в Тихом океане во время Второй мировой войны. Просто-напросто для ведения секретных переговоров использовались носители языка индейцев навахо, словами из которого обозначались военные термины. Например, слово *чаи-дагахи-найл-цайди* (*chay-dagahi-nail-tsaidi* — буквально: убийца черепах) означало противотанковое оружие. Язык навахо тоновый (для различения смысла используется повышение или понижение тона), весьма сложный, не имеет письменной формы. Но самое большое его достоинство заключалось в том, что ни один японец не имел о нем ни малейшего представления.

В сентябре 1945 года газета *San Diego Union* так описывала этот код: «В течение трех лет, где бы ни высаживались военные моряки, уши японцев различали лишь странный булькающий шум, перемежающийся с другими звуками. Все это напоминало клич тибетского монаха или звук опустошаемой бутылки с горячей водой». Японцы так и не смогли взломать этот код, и многие носители языка индейцев навахо были удостоены высоких воинских наград за отличную службу и смелость. Тот факт, что США смогли расшифровать японский код, а японцы так и не узнали язык навахо, сыграл важную роль в американской победе в Тихом океане.

8.1.1. Основы криптографии

Исторически использовали и развивали искусство криптографии представители четырех профессий: военные, дипломатический корпус, люди, ведущие дневник, и любовники. Из них наиболее важную роль в развитии этой области сыграли военные. В военных организациях секретные сообщения традиционно отдавались для зашифровки и передачи плохо оплачиваемым шифровальщикам. Сам объем сообщений не позволял выполнить эту работу небольшим количеством элитных специалистов.

До появления компьютеров одним из основных сдерживающих факторов в криптографии была возможность шифровальщика выполнить необходимые преобразования, часто на поле боя, с помощью несложного оборудования. Кроме того, достаточно сложной задачей было быстрое переключение с одного криптографического метода на другой, так как для этого требовалось переобучение большого количества людей. Тем не менее опасность того, что шифровальщик может быть захвачен противником, заставила постоянно развивать способы смены криптографических методов при необходимости. Эти противоречивые требования приводят к модели процесса шифрования—дешифрации[1], показанной на рис. 8.1.

Сообщения, подлежащие зашифровке, называемые **открытым текстом (plaintext)**, преобразуются с помощью функции, вторым входным параметром которой является **ключ (key)**. Результат процесса шифрования, называемый **зашифрованным текстом (ciphertext)**, передается обычно по радио или через связного. Предполагается, что противник или **злоумышленник (intruder)** слышит и аккуратно копирует весь зашифрованный текст. Однако в отличие от получателя, которому предназначается данное

[1] Используемая автором и, соответственно, в переводе терминология несколько отличается от приводимой в российских нормативных документах. Однако она достаточно широко применяется на практике, поэтому сохранена вместе с вводимыми автором определениями. — *Примеч. ред.*

сообщение, злоумышленник не знает ключа дешифрации, и поэтому расшифровка сообщения представляет для него большие трудности, а порой она просто невозможна. Иногда злоумышленник может не только прослушивать канал связи (пассивный злоумышленник), но способен также записывать сообщения и воспроизводить их позднее, вставлять свои сообщения или модифицировать оригинальные сообщения, прежде чем они достигнут получателя (активный злоумышленник). Искусство взлома шифров называется **криптоанализом (cryptanalysis)**. Искусства изобретать шифры (криптография) и взламывать их (криптоанализ) называются вместе **криптологией (cryptology)**.

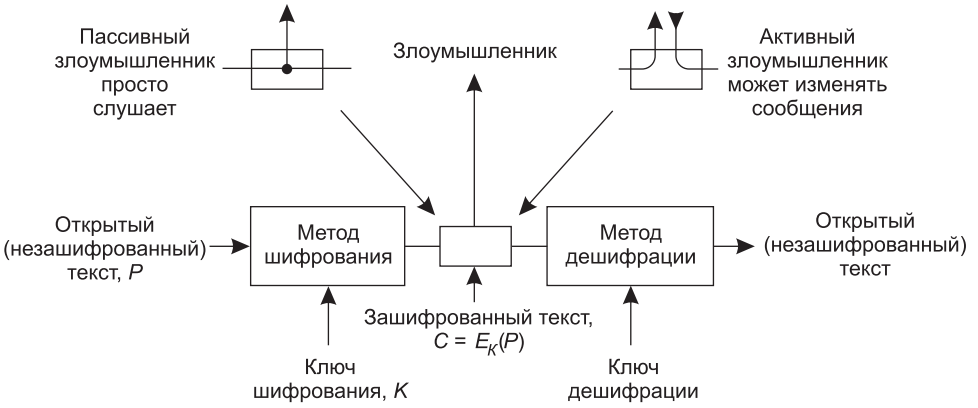

Рис. 8.1. Модель процесса шифрования—дешифрации (для шифра с симметричным ключом)

Обычно для обозначения открытого текста, зашифрованного текста и ключей полезно использовать специальную нотацию. Мы будем использовать формулу $C = EK(P)$, обозначающую, что при зашифровке открытого текста P с помощью ключа K получается зашифрованный текст C. Аналогично формула $P = DK(C)$ означает расшифровку зашифрованного текста C для восстановления открытого текста. Из этих двух формул следует, что

$$D_K(E_K(P)) = P.$$

Такая нотация предполагает, что E и D являются просто математическими функциями. Они в действительности таковыми и являются. Единственная хитрость состоит в том, что обе эти функции имеют по два параметра, один из которых (ключ) мы написали не в виде аргумента, а в виде нижнего индекса, чтобы отличать его от сообщения.

Основное правило криптографии состоит в предположении, что криптоаналитику (взломщику шифра) известен используемый метод шифрования. Другими словами, злоумышленник точно знает, как работают методы шифрования E и дешифрации D на рис. 8.1. Из-за огромных усилий, необходимых для разработки, тестирования и внедрения нового метода, каждый раз, когда старый метод оказывался или считался скомпрометированным, хранить алгоритм шифрования в секрете просто непрактично. А предположение, что метод остается секретным, когда это уже не так, могло бы причинить еще больший вред.

Здесь на помощь приходит ключ шифрования. Ключ состоит из относительно короткой строки, определяющей один из огромного количества вариантов результата шифрования. В отличие от самого метода шифрования, который может изменяться только раз в несколько лет, ключ можно менять так часто, как это нужно. Таким образом, наша базовая модель представляет собой постоянный и известный общий метод, в котором в качестве параметра используется секретный и легко изменяемый ключ. Идея, заключающаяся в предположении о том, что криптоаналитику известен метод и краеугольным камнем секретности является эксклюзивный ключ, называется **принципом Керкгофа (Kerckhoff's principle)**. Его в 1883 году впервые высказал фламандский военный криптограф Аугуст Керкгоф (Auguste Kerckhoff, 1883). Таким образом, принцип Керкгофа гласит:

Алгоритмы шифрования общедоступны; секретны только ключи.

Секретности алгоритма не стоит придавать большого значения. Попытка сохранить алгоритм в тайне, называемая в торговле **безопасностью за счет неясности (security by obscurity)**, обречена на провал. К тому же, опубликовав свой алгоритм, разработчик получает бесплатную консультацию от большого количества ученых-криптоаналитиков, горящих желанием взломать новую систему и тем самым продемонстрировать свой ум и ученость. Если никто не смог взломать алгоритм в течение долгого времени после его опубликования, то, по-видимому, этот алгоритм достаточно прочен.

Поскольку реально в тайне хранится только ключ, основной вопрос заключается в его длине. Рассмотрим простой кодовый замок. Его основной принцип состоит в том, что вы последовательно вводите десятичные цифры. Все это знают, но ключ хранится в секрете. Ключ длиной в две цифры образует 100 вариантов. Ключ длиной в три цифры означает 1000 вариантов, а при длине ключа в шесть цифр число комбинаций достигает миллиона. Чем длиннее ключ, тем выше **показатель трудозатрат (work factor)** взломщика шифра. При увеличении длины ключа показатель трудозатрат для взлома системы путем простого перебора значений ключа растет экспоненциально. Секретность передаваемого сообщения обеспечивается мощным (но все же открытым) алгоритмом и длинным ключом. Чтобы не дать прочитать свою электронную почту младшему брату, достаточно ключа длиной в 64 двоичных разряда. В коммерческих системах имеет смысл использовать ключи длиной 128 бит. Чтобы защитить ваши тексты от правительств развитых государств, потребуются ключи длиной, по меньшей мере, в 256 бит.

С точки зрения криптоаналитика задача криптоанализа имеет три принципиальных варианта. Во-первых, у криптоаналитика может быть некоторое количество зашифрованного текста без соответствующего открытого текста. Задачки, в которых в качестве исходных данных имеется в наличии **только зашифрованный текст (ciphertext-only)**, часто печатаются в различных газетах в разделе ребусов. Во-вторых, у криптоаналитика может оказаться некоторое количество зашифрованного текста и соответствующего ему открытого текста. В этом случае мы имеем дело с **проблемой известного открытого текста (known plaintext)**. Наконец, когда у криптоаналитика есть возможность зашифровать любой кусок открытого текста по своему выбору, мы получаем третий вариант проблемы дешифрации, то есть **проблему произвольного открытого текста (chosen plaintext)**. Если бы криптоаналитикам было позволено

задавать вопросы типа: «Как будет выглядеть зашифрованное ABCDEFGHJKL?», задачки из газет решались бы очень легко.

Новички в деле криптографии часто полагают, что шифр является достаточно надежным, если он может выдержать атаку первого типа (только зашифрованный текст). Такое предположение весьма наивно. Во многих случаях криптоаналитик может угадать часть зашифрованного текста. Например, первое, что говорят многие компьютеры при входе в систему, это `login:`. После того как криптоаналитик получит несколько соответствующих друг другу пар кусков зашифрованного и открытого текста, его работа становится значительно легче. Для обеспечения секретности нужна предусмотрительность криптографа, который должен гарантировать, что система не будет взломана, даже если его оппонент сможет закодировать несколько фрагментов открытого текста по своему выбору.

Исторически методы шифрования разделились на две категории: метод подстановки и метод перестановки. Мы кратко рассмотрим их в качестве введения в современную криптографию.

8.1.2. Метод подстановки

В шифрах, основанных на **методе подстановки (substitution cipher)**, каждый символ или группа символов заменяется другим символом или группой символов. Одним из древнейших шифров является приписываемый Юлию Цезарю (Julius Caesar) **шифр Цезаря (Caesar cipher)**. Этот шифр заменяет все буквы алфавита на другие с помощью циклического сдвига на три позиции. Так буква *a* становится буквой *D*, *b* становится *E*, *c* превращается в *F*, ... , а *z* — в *C*. Например, слово *attack* превращается в *DWWDFN*. В наших примерах открытый текст будет обозначаться строчными символами, а зашифрованный текст — прописными.

Некоторое обобщение шифра Цезаря представляет собой сдвиг алфавита не на три символа, а на произвольное число *k* символов. В этом случае *k* становится ключом к общему методу циклически сдвигаемых алфавитов. Шифр Цезаря, возможно, и сумел обмануть жителей Помпеи, но с тех пор ему более уже никого не удалось ввести в заблуждение.

Следующее усовершенствование состоит в установлении соответствия каждому встречающемуся в открытом тексте символу другого символа. Например,

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

Такая система называется **моноалфавитным подстановочным шифром (monoalphabetic substitution cipher)**, ключом к которому является 26-символьная строка, соответствующая полному алфавиту. В нашем примере слово *attack* будет выглядеть, как *QZZQEA*.

На первый взгляд такая система может показаться надежной, так как, даже если криптоаналитику известна общая система, он не знает, какой из 26! ≈ 4 × 1026 возможных вариантов ключа применить. В отличие от шифра Цезаря применение метода простого перебора в данном случае весьма сомнительно. Даже при затратах 1 нс на проверку одного варианта ключа, чтобы перепробовать все ключи, миллиону компьютерных чипов, работающих одновременно, понадобится около 10 000 лет.

Тем не менее подобный шифр легко взламывается даже при наличии довольно небольших фрагментов зашифрованного текста. Для атаки шифра может быть использовано преимущество статистических характеристик естественных языков. Например, в английском языке буква *e* встречается в тексте чаще всего. Следом за ней по частоте использования идут буквы *t*, *o*, *a*, *n*, *i* и т. д. Наиболее часто встречающимися комбинациями из двух символов (**биграммами** — **digrams**) являются *th*, *in*, *er*, *re* и *an*. Наиболее часто встречающимися комбинациями из трех символов, или **триграммами** (**trigrams**), являются *the*, *ing*, *and* и *ion*.

Криптоаналитик, пытающийся взломать моноалфавитный шифр, начнет с того, что сосчитает относительные частоты всех символов алфавита в зашифрованном тексте. Затем он может попытаться заменить наиболее часто встречающийся символ буквой *e*, а следующий по частоте — буквой *t*. Затем он посмотрит на триграммы и попытается найти что-либо похожее на *tXe*, после чего он сможет предположить, что *X* — это *h*. Аналогично, если последовательность *thYt* встречается достаточно часто, то, вероятно, *Y* обозначает символ *a*. Обладая этой информацией, криптоаналитик может поискать часто встречающуюся триграмму вида *aZW*, что, скорее всего, означает *and*. Продолжая в том же духе, угадывая буквы, биграммы, триграммы и зная, какие последовательности символов являются наиболее вероятными, криптоаналитик побуквенно восстанавливает исходный текст.

Другой метод заключается в угадывании сразу целого слова или фразы. Например, рассмотрим следующий зашифрованный текст, полученный от бухгалтерской фирмы (разбитый на блоки по пять символов):

```
CTBMN BYCTC BTJDS QXBNS GSTJC BTSWX CTQTZ CQVUJ
QJSGS TJQZZ MNQJS VLNSX VSZJU JDSTS JQUUS JUBXJ
DSKSU JSNTK BGAQJ ZBGYQ TLCTZ BNYBN QJSW
```

В сообщении бухгалтерской фирмы, скорее всего, должно встречаться слово *financial* (финансовый). Используя тот факт, что в этом слове буква *i* встречается дважды, разделенная четырьмя другими буквами, мы будем искать в зашифрованном тексте повторяющиеся символы, отстоящие друг от друга на это расстояние. В результате мы найдем 12 таких мест в тексте в позициях 6, 15, 27, 31, 42, 48, 56, 66, 70, 71, 76 и 82. Однако только в двух случаях, в позициях 31 и 42, следующий символ (соответствующий букве *n* в открытом тексте) повторяется в соответствующем месте. Из этих двух вариантов символ *a* будет иметь правильное расположение только для позиции 31. Таким образом, теперь нам известно, что слово *financial* начинается в позиции 30. Далее можно продолжать, применяя лингвистическую статистику английского языка и угадывая целые слова.

8.1.3. Метод перестановки

Шифры, основанные на методе подстановки, сохраняют порядок символов, но подменяют их. Шифры, использующие **метод перестановки** (**transposition ciphers**), меняют порядок следования символов, но не изменяют сами символы. На рис. 8.2 показан простой перестановочный шифр с колоночной перестановкой. Ключом к шифру служит слово или фраза, не содержащая повторяющихся букв. В данном примере в качестве ключа используется слово MEGABUCK. Цель ключа — пронумеровать колонки. Пер-

вой колонкой становится колонка под буквой, расположенной ближе всего к началу алфавита и т. д. Открытый текст записывается горизонтально в строках. Шифрованный текст читается по колонкам, начиная с колонки с младшей ключевой буквой.

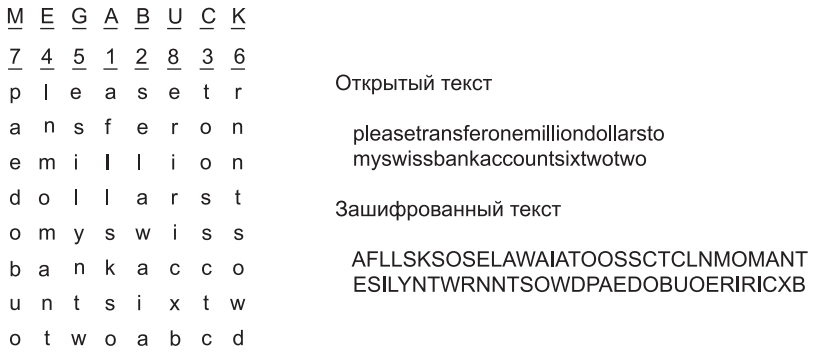

Рис. 8.2. Перестановочный шифр

Чтобы взломать перестановочный шифр, криптоаналитик должен вначале понять, что он имеет дело именно с перестановочным шифром. Если взглянуть на частоту символов *E, T, A, O, I, N* и т. д., легко заметить, что их частоты соответствуют нормальным частотам открытого текста. В таком случае, очевидно, что этот шифр является перестановочным, так как каждая буква в таком шифре представляет сама себя.

Затем нужно угадать число колонок. Во многих случаях по контексту сообщения можно угадать слово или фразу. Например, предположим, что криптоаналитик подозревает, что где-то в сообщении должно встретиться словосочетание *milliondollars*. Обратите внимание, что в результате того, что эти слова присутствуют в исходном тексте, в шифрованном тексте встречаются биграммы *MO*, *IL*, *LL*, *LA*, *IR* и *OS*. Символ *O* следует за символом *M* (то есть они стоят рядом по вертикали в колонке 4), так как они разделены в предполагаемой фразе дистанцией, равной длине ключа. Если бы использовался ключ длиной семь, тогда вместо перечисленных выше биграмм встречались бы следующие: *MD*, *IO*, *LL*, *LL*, *IA*, *OR* и *NS*. Таким образом, для каждой длины ключа в шифрованном тексте образуется различный набор биграмм. Перебрав различные варианты, криптоаналитик часто довольно легко может определить длину ключа.

Остается узнать только порядок колонок. Если число колонок k невелико, можно перебрать все $k(k-1)$ возможных комбинаций пар соседних колонок, сравнивая частоты образующихся биграмм со статистическими характеристиками английского языка. Пара с лучшим соответствием считается правильно позиционированной. Затем все оставшиеся колонки по очереди проверяются в сочетании с уже найденной парой. Колонка, в которой биграммы и триграммы дают максимальное совпадение со статистикой, предполагается правильной. Весь процесс повторяется, пока не будет восстановлен порядок всех колонок. Есть шанс, что на данном этапе текст уже будет распознаваемым (например, если вместо слова *million* мы увидим *milloin*, то сразу станет ясно, где сделана ошибка).

Некоторые перестановочные шифры принимают блок фиксированной длины на входе и выдают блок фиксированной длины на выходе. Такие шифры полностью опре-