Дориченко С.А., Ященко В.В. 25 этюдов о шифрах

Предисловие

Дорогой читатель! Перед Вами — первая книга новой серии «Математические основы криптологии». Серия задумана как популярное изложение вопросов и задач, связанных с защитой информации, шифрованием и дешифрованием, цифровой подписью, компьютерной безопасностью и т.п.

Такие задачи в настоящее время часто приходится решать с целью обеспечения определенных интересов (государственных, коммерческих, личных и др.). Наиболее надежные средства их решения дает криптография — наука о методах преобразования (шифрования) информации с целью ее защиты от незаконных пользователей. Криптография базируется на самых последних достижениях фундаментальных наук и в первую очередь математики.

Книга «25 этюдов о шифрах» дает представление о математических проблемах современной криптографии. В ней вводятся и поясняются на примерах все основные понятия криптографии. В строгой, но общедоступной форме даются необходимые математические определения и на их основе разъясняются многие идеи криптографии. Для иллюстрации используются шифры, широко известные из исторической и приключенческой литературы.

Книга ликвидирует имеющийся в настоящее время пробел в литературе на русском языке о криптографии. Она может быть интересна широкому кругу читателей. В силу своей строгости и лаконичности может использоваться и для учебных целей в качестве популярного справочника основных понятий криптографии.

Н.Н. Андреев

Президент Академии криптографии Российской Федерации

Введение Как читать эту книгу

Настоящая книга является популярным изложением основных понятии и идеи современной криптографии, а точнее — математических вопросов криптографии. Она открывает новую серию «Математические проблемы криптологии».

Авторы избрали для книги жанр этюдов или зарисовок, чтобы при небольшом объеме дать как можно больше сведений. Полное систематическое изложение тех же вопросов требует сотен страниц. Любознательных читателей отсылаем к изданным за рубежом учебникам и к изданиям на русском языке, список которых приведен в конце книги.

Одна из целей книги — введение в математическую проблематику криптографии. Поэтому вторая и особенно третья главы рассчитаны на читателя, склонного к математическим размышлениям. Вместе с тем для понимания книги не требуются какие-либо углубленные математические знания.

Основные термины и понятия при первом своем появлении в тексте выделяются курсивом. Иногда в текст мелким шрифтом «врезается» формальное определение этого понятия.

Некоторые этюды являются ответом на один вопрос, вынесенный в заголовок этюда. Поэтому они построены так: сначала дается короткий и формальный ответ на вопрос, а потом более подробные комментарии и примеры. После прочтения такого этюда полезно вернуться к его началу и уже с новыми знаниями вновь прочитать ответ на основной вопрос этюда.

Часть читателей пожелает не просто прочитать, а разобраться, глубже обдумать каждый этюд. Для них некоторые этюды дополняются разделом «*Подумайте сами*», содержащим контрольные вопросы и задачи. Кроме того, в Приложении приведены наиболее интересные задачи олимпиад по криптографии.

Если при прочтении книги возникнут трудности в понимании некоторых фрагментов, не надо огорчаться: скорее всего данный фрагмент — это «окно» в большую науку.

В настоящее время в теоретической криптографии используются понятия и результаты многих разделов математики: алгебры, теории чисел, теории сложности алгоритмов и вычислений, теории кодирования и др. Поэтому для современного криптографа прежде всего важна хорошая математическая подготовка.

Школьникам, которые решили избрать криптографию своей профессией, рекомендуем один из трех вузов:

— институт криптографии, связи и информатики (ИКСИ) Академии безопасности ФСК Российской Федерации;

— механико-математический факультет Московского государственного университета им. М.В. Ломоносова (МГУ);

— факультет защиты информации Российского государственного гуманитарного университета (РГГУ).

Авторы благодарят сотрудников лаборатории МГУ по математическим проблемам криптографии, которые участвовали в обсуждении этой книги на всех этапах ее написания.

С.А. Дориченко

В.В. Ященко

Глава 1 Основные понятия

1.1. Защита информации

Когда? В тех случаях, когда есть опасения, что информация станет доступной посторонним, которые могут обратить её во вред законному пользователю.

Зачем? Чтобы предотвратить возможный вред от разглашения информации.

Информация — основное понятие научных направлений, изучающих процессы передачи, переработки и хранения различных данных. Суть понятия информации обычно поясняется на примерах. Формальное определение не дается, поскольку понятие информации относится к таким же фундаментальным понятиям, как материя, люди уже давно поняли, что информация может быть настоящим сокровищем, и поэтому часто много усилий затрачивалось как на ее охрану, так и на ее добывание. Вообще говоря, совершенно не обязательно это связано с какими-то «шпионскими» делами. Информация, которая нуждается в защите, возникает в самых разных жизненных ситуациях. Обычно в таких случаях говорят, что информация содержит тайну или является *защищаемой*, *приватной*, *конфиденциальной*, *секретной*. Для наиболее типичных, часто встречающихся ситуаций такого типа введены даже специальные понятия:

— *государственная тайна*;

— *военная тайна*;

— *коммерческая тайна*;

— *юридическая тайна*;

— *врачебная тайна* и т.д.

Далее в этой книге мы будем говорить о защищаемой информации, имея в виду следующие признаки такой информации:

— имеется какой-то определенный круг *законных пользователей*, которые имеют право владеть этой информацией;

— имеются *незаконные пользователи*, которые стремятся овладеть этой информацией с тем, чтобы обратить ее себе во благо, а законным пользователям во вред.

Для простоты мы здесь ограничиваемся рассмотрением только одной *угрозы* — угрозы *разглашения информации*. Существуют и другие угрозы для защищаемой информации со стороны незаконных пользователей: подмена, имитация и др. Заинтересованному читателю рекомендуем аналогично продумать вопросы, связанные с подменой и имитацией информации.

Сейчас жизнь устроена так, что между людьми происходит интенсивный обмен информацией, причем часто на громадные расстояния. Для этого земной шар опутали

различными видами *технических средств связи*: телеграф, телефон, радио, телевидение и др. Но часто возникает необходимость в обмене между удаленными законными пользователями не просто информацией, а защищаемой информацией. В этом случае незаконный пользователь может попытаться перехватить информацию из общедоступного технического *канала связи*. Опасаясь этого, законные пользователи должны принять дополнительные меры для защиты своей информации. Разработкой таких мер защиты занимаются *криптография*и *стеганография*.

Криптография — наука о методах преобразования (шифрования) информации с целью ее защиты от незаконных пользователей.

Стеганография — набор средств и методов скрытия факта передачи сообщения.

Шифр — способ, метод преобразования информации с целью ее защиты от незаконных пользователей.

В заключение данного этюда подчеркнем, что есть еще одна важная проблема: проблема соотношения цены информации, затрат на ее защиту и затрат на ее добывание. Подробное обсуждение этого вопроса выходит за рамки настоящей книги, но любознательный читатель может сам обдумать различные возникающие здесь ситуации. Отметим только, что при современном уровне развития техники сами средства связи, а также разработка средств перехвата информации из них и средств защиты информации требуют очень больших затрат.

Подумайте сами:

1. Приведите примеры упомянутых в этюде видов тайны.

2. Для ваших примеров опишите законный пользователей, незаконных пользователей, возможный вред от разглашения защищаемой информации.

1.2. Чем криптография отличается от стеганографии

Стеганография скрывает сам факт передачи сообщения, а криптография считает, что сообщение (в шифрованном виде!) доступно незаконному пользователю, но он не может извлечь из этого сообщения защищаемую информацию.

Первые следы стеганографических методов теряются в глубокой древности. Например, известен такой способ скрытия письменного сообщения: голову раба брили, на коже головы писали сообщение и после отрастания волос раба отправляли к адресату.

Из детективных произведений хорошо известны различные способы скрытого письма между строк обычного, незащищаемого письма: от молока до сложных химических реактивов с последующей обработкой.

Также из детективов известен современный метод «*микроточки*»: сообщение записывается с помощью современной техники на очень маленький носитель — «микроточку», которая пересылается с обычным письмом, например, под маркой или где-нибудь в другом заранее обусловленном месте.

Один типично стеганографический прием тайнописи — *акростих* — хорошо известен знатокам поэзии. Акростих — это такая организация стихотворного текста, при которой, например, начальные буквы каждой строки образуют скрываемое сообщение.

В настоящее время в связи с широким распространением компьютеров известно много тонких методов «запрятывания» защищаемой информации внутри больших объемов информации, хранящейся в компьютере.

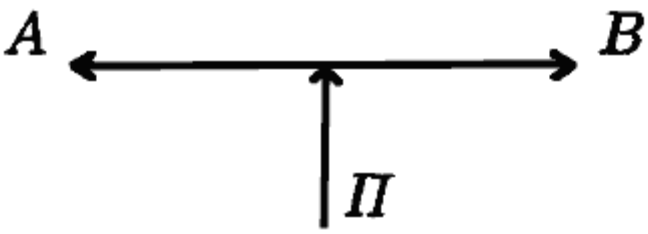
Даже из приведенного небольшого количества примеров видно, что при использовании стеганографии в отличие от криптографии защищаемая информация не преобразуется, а скрывается сам факт ее передачи.

Подумайте сами:

1. Разработайте какой-нибудь стеганографический метод защиты информации, хранящейся в компьютере.

1.3. Как можно представить основной объект криптографии?

Можно представить так:



Здесь *А* и *В* — удаленные законные пользователи защищаемой информации; они хотят обмениваться информацией по общедоступному каналу связи, а *П* — незаконный пользователь (*противник*), который может перехватывать передаваемые по каналу связи сообщения и пытаться извлечь из них интересующую его информацию

Приведенную формальную схему можно также считать моделью типичной ситуации, в которой применяются криптографические методы защиты информации.

Отметим, что исторически в криптографии закрепились некоторые чисто военные слова (противник, атака на шифр и др.) Они наиболее точно отражают смысл соответствующих криптографических понятий. Вместе с тем широко известная военная терминология, основанная на понятии кода (военно-морские коды, коды Генерального штаба, кодовые книги, код обозначения и т.п.) уже уходит из теоретической криптографии. Дело в том, что за последние десятилетия сформировалась *теория кодирования* — новое большое научное направление, которое разрабатывает и изучает методы защиты информации от случайных искажений в каналах связи. И если ранее термины *кодирование* и *шифрование*употреблялись в некотором

смысле как синонимы, то теперь это недопустимо. Так, например, очень распространенное выражение «кодирование — разновидность шифрования» становится просто неправильным.

Криптография занимается методами преобразования информации, которые бы не позволили противнику извлечь ее из перехватываемых сообщений. При этом по каналу связи передается уже не сама защищаемая информация, а результат ее преобразования с помощью шифра, и для противника возникает сложная задача *вскрытия шифра*.

Вскрытие (*взламывание*) *шифра* — процесс получения защищаемой информации (*открытого текста*) из шифрованного сообщения (*шифртекста*) без знания примененного шифра.

Шифрование (*зашифровывание*) — процесс применения шифра к защищаемой информации, т.е. преобразование защищаемой информации в шифрованное сообщение с помощью определенных правил, содержащихся в шифре.

Дешифрование — процесс, обратный шифрованию, т.е. преобразование шифрованного сообщения в защищаемую информацию с помощью определенных правил, содержащихся в шифре.

Однако помимо перехвата и вскрытия шифра противник может пытаться получить защищаемую информацию многими другими способами.

Наиболее известным из таких способов является агентурный, когда противник каким-либо путем склоняет к сотрудничеству одного из законных пользователей и с помощью этого агента получает доступ к защищаемой информации. В такой ситуации криптография бессильна.

Противник может пытаться не получить, а уничтожить или модифицировать защищаемую информацию в процессе ее передачи. Это — совсем другой тип угроз для информации, отличный от перехвата и вскрытия шифра. Для защиты от таких угроз разрабатываются свои специфические методы. Среди многочисленных угроз для защищаемой информации криптография противостоит только некоторым. Поэтому естественно сочетать криптографию с мерами по защите информации от других угроз.

В заключение этого этюда отметим, что чаще всего обмен защищаемой информацией происходит не только между двумя *абонентами* — законными пользователями, а в сети абонентов, и тогда возникают новые задачи. Сети могут быть разных размеров — от единиц до тысяч абонентов. Тем не менее, основные понятия и идеи криптографии можно понять на примере описанного основного объекта криптографии.

1.4. Криптография, как искусство. Немного теории

Долгое время занятие криптографией было уделом чудаков-одиночек. Среди них были одаренные ученые, дипломаты, священнослужители. Известны случаи, когда криптография считалась даже черной магией. Этот период развития криптографии как искусства длился с незапамятных времен до начала XX века, когда появились первые шифровальные машины. Понимание математического характера решаемых криптографией задач пришло только в середине XX века — после работ выдающегося американского ученого К. Шеннона.

История криптографии связана с большим количеством дипломатических и военных тайн и поэтому окутана туманом легенд. Наиболее полная книга по истории криптографии содержит более тысячи страниц. Она опубликована в 1967 году в Нью-Йорке и на русский язык еще не переведена[1]. На русском языке недавно вышел в свет фундаментальный труд по истории криптографии в России[2].

Свой след в истории криптографии оставили многие хорошо известные исторические личности. Приведем несколько наиболее ярких примеров.

Первые сведения об использовании шифров в военном деле связаны с именем спартанского полководца Лисандра (шифр «Сциталь», V век до нашей эры). Цезарь использовал в переписке шифр, который вошел в историю как «шифр Цезаря». В древней Греции был изобретен вид шифра, который в дальнейшем стал называться «квадрат Полибия». Братство франкмасонов с

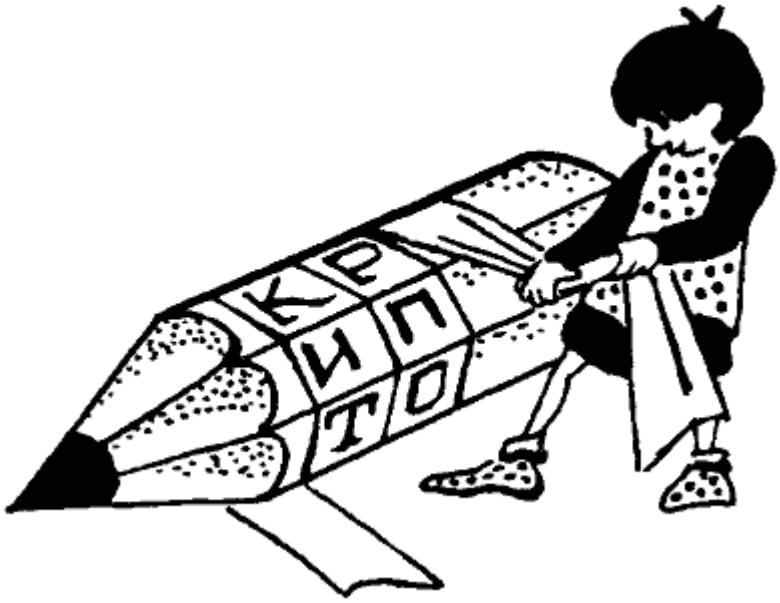
момента своего возникновения (VIII век) разработало и использовало целую систему особых шифров. Одну из первых книг по криптографии написал аббат И. Трителий (1462–1516), живший в Германии. В 1566 году известный механик и математик Д. Кардано опубликовал работу с описанием изобретенной им системы шифрования («решетка Кардано»). Франция XVI века оставила в истории криптографии шифры короля Генриха IV и Ришелье. В упомянутой книге Т.А. Соболевой подробно описано много российских шифров, в том числе и «цифирная азбука» 1700 года, автором которой был Петр Великий.

Некоторые сведения о свойствах шифров и их применениях можно найти и в художественной литературе, особенно в приключенческой, детективной и военной. Хорошее подробное объяснение особенностей одного из простейших шифров - *шифра замены* и методов его вскрытия содержится в двух известных рассказах: «Золотой жук» Э. По и «Пляшущие человечки» А. Конан-Дойля.

Много занимательной информации по криптографии публикуется в издаваемом в США научно-популярном журнале «Cryptology». Обширный библиографический список (111 названий) зарубежной литературы по криптографии содержится в очень полезной и важной статье Диффи и Хеллмэна[3], которая переведена на русский язык и общедоступна (о революционном вкладе авторов этой статьи в криптографию будет рассказано в главе 3).

Рассмотрим более подробно три примера.

Шифр «Сциталь». Этот шифр известен со времен войны Спарты и Персии против Афин. Спартанский полководец Лисандр подозревал персов в возможной измене, но не знал их тайных планов. Его агент в стане персов прислал шифрованное сообщение, которое позволило Лисандру опередить персов и разгромить их. Шифрованное сообщение было написано на поясе официального гонца от персов следующим образом: агент намотал пояс на сциталь (деревянный цилиндр определенного диаметра) и написал на поясе сообщение вдоль сциталя; потом он размотал пояс, и получилось, что поперек пояса в беспорядке написаны буквы. Гонец не догадывался, что узор на его красивом поясе на самом деле содержит зашифрованную информацию. Лисандр взял сциталь такого же диаметра, аккуратно намотал на него пояс и вдоль сциталя прочитал сообщение от своего агента.



Например, если роль сциталя выполняет карандаш с шестью гранями, то открытый текст КРИПТОГРАФИЯ может быть преобразован в шифртекст РПОРФЯКИТГАИ. Шифртекст может быть и другим, так как он зависит не только от диаметра карандаша. Поэкспериментируйте!

Отметим, что в этом шифре преобразование открытого текста в шифрованный заключается в определенной перестановке букв открытого текста. Поэтому класс шифров, к которым относится и шифр «Сциталь», называется *шифрами перестановки*. Математическому описанию таких шифров посвящен этюд 2.4.

Шифр Цезаря. Этот шифр реализует следующее преобразование открытого текста: каждая буква открытого текста заменяется третьей после нее буквой в алфавите, который считается написанным по кругу, т.е. после буквы «я» следует буква «а». Поэтому класс шифров, к которым относится и шифр Цезаря, называется *шифрами замены*. Математическому описанию таких шифров посвящен этюд 2.4.

Например, открытый текст КРИПТОГРАФИЯ при таком способе шифрования преобразуется в шифртекст НУЛТХСЁУГЧЛВ. Отметим, что Цезарь заменял букву третьей после нее буквой, но можно заменять и пятой, и какой-нибудь другой. Главное, чтобы тот, кому посылается шифрованное сообщение, знал эту величину сдвига.

Шифр Виженера. Этот шифр удобнее всего представлять себе как шифр Цезаря с переменной величиной сдвига. Чтобы знать, на сколько сдвигать очередную букву открытого текста, заранее договариваются о способе запоминания сдвигов. Сам Виженер предлагал запоминать *ключевое слово*, каждая буква которого своим номером в алфавите указывает величину сдвига. Ключевое слово повторяется столько раз, сколько нужно для замены всех букв открытого текста. Например, ключевое слово ВАЗА означает следующую последовательность сдвигов букв открытого текста: 3191319131913191... Например, открытый текст КРИПТОГРАФИЯ при таком способе шифрования преобразуется в шифртекст НССРХПЛСГХСА.

Дальнейшее развитие идеи ключевого слова, а именно, идея запоминать способ преобразования открытого текста с помощью какой-либо книги, привело к возникновению различных видов так называемых *книжных шифров*. Они хорошо известны любителям детективной и приключенческой литературы.

Подумайте сами:

1. Поэкспериментируйте с шифрами Цезаря и Виженера.
2. Попробуйте найти способ вскрытия шифра «Сциталь» (не зная диаметра сциталя).

1.5. Что такое ключ?

Под *ключом* в криптографии понимают сменный элемент шифра, который применен для шифрования конкретного сообщения.

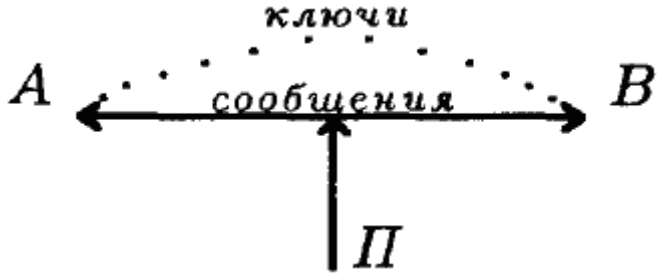
В древнейшем шифре «Сциталь», описанном в этюде 1.4, ключом является диаметр сциталя. При этом не меняя принцип построения шифра, можно для шифрования разных сообщений пользоваться сциталями разных диаметров.

В шифрах типа шифра Цезаря ключом является величина сдвига букв шифртекста относительно букв открытого текста.

Зачем же нужен ключ? Из предыдущего изложения понятно, что придумывание хорошего шифра — дело трудоемкое. Поэтому желательно увеличить «время жизни» хорошего шифра и использовать его для шифрования как можно большего количества сообщений. Но при этом возникает опасность, что противник уже разгадал (вскрыл) шифр и читает защищаемую информацию. Если же в шифре есть сменный ключ, то, заменив ключ, можно сделать так, что разработанные противником методы уже не дают эффекта. Этот принцип особенно полезен и важен в тех случаях, когда применимы дорогостоящие *шифрующие машины* (*шифрмашины*) в больших сетях связи.

Описанные соображения привели к тому, что безопасность защищаемой информации стала определяться в первую очередь ключом. Сам шифр, шифрмашина или принцип шифрования стали считать известными противнику и доступными для предварительного изучения. Но применяемые в шифрах преобразования информации стали сильно зависеть от ключа. А для противника появились новая задача — определить ключ, после чего можно легко прочитать зашифрованные на этом ключе сообщения. Законные пользователи, прежде чем обмениваться шифрованными сообщениями, должны тайно от противника обменяться ключами или установить одинаковый ключ на обоих концах канала связи.

Вернёмся к формальному описанию основного объекта криптографии (этюд 1.3). Теперь в него необходимо внести существенное изменение — добавить недоступный для противника секретный канал связи для обмена ключами:

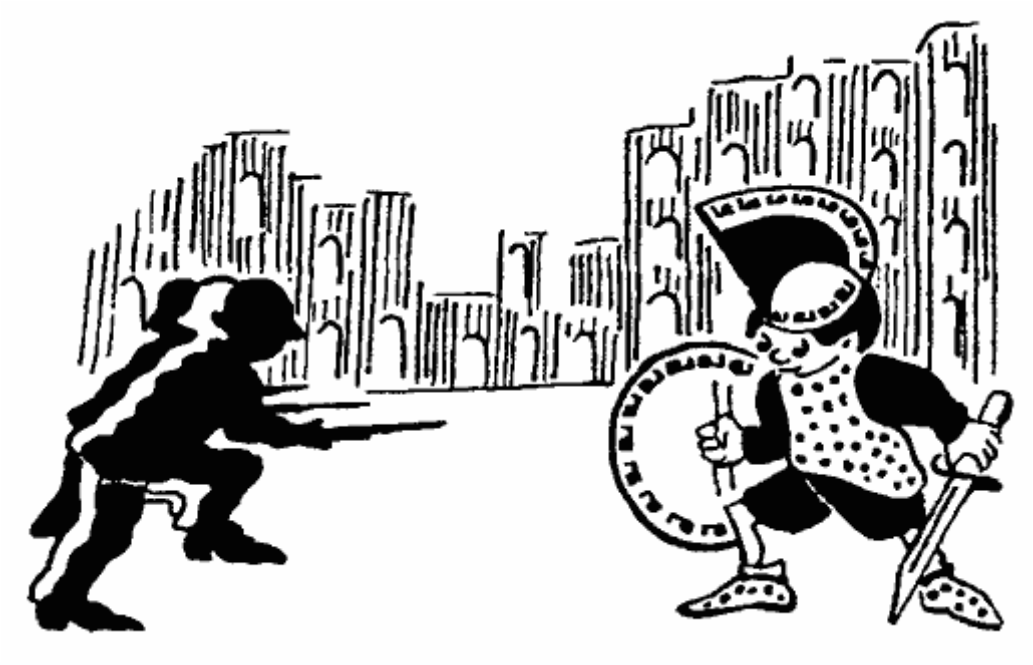


Практическое построение таких сетей связи для большого обмена шифрованными сообщениями стало ещё более дорогостоящим мероприятием.

Подумайте сами:

1. Что является ключом в шифре Виженера.

1.6. Атака на шифр. Стойкость шифра



Под *атакой на шифр* понимают попытку вскрытия этого шифра.

Под *стойкостью шифра* понимают способность шифра противостоять всевозможным атакам на него.

Понятие стойкости шифра является центральным для криптографии. Хотя качественно понять его довольно легко, но получение строгих доказуемых оценок стойкости для каждого конкретного шифра — проблема нерешённая. Это объясняется тем, что до сих пор нет необходимых для решения такой проблемы математических результатов. (Мы вернемся к обсуждению этого вопроса в этюде 2.6.) Поэтому стойкость конкретного шифра оценивается только путем всевозможных попыток его вскрытия и зависит от квалификации *криптоаналитиков*, атакующих шифр. Последнюю процедуру иногда называют *проверкой стойкости*.

Важным подготовительным этапом для проверки стойкости шифра является продумывание различных предполагаемых возможностей, с помощью которых противник может атаковать шифр. Появление таких возможностей у противника обычно не зависит от криптографии, это является некоторой внешней подсказкой и существенно влияет на стойкость шифра. Поэтому оценки стойкости шифра всегда содержат те предположения о противнике, в условиях которых эти оценки получены.

Прежде всего, как это уже отмечалось в этюде 1.5, обычно считается, что противник знает сам шифр и имеет возможности для его предварительного изучения. Противник также знает некоторые характеристики открытых текстов (защищаемой информации), например, общую тематику сообщений, их стиль, некоторые стандарты, форматы и т.д.

Из более специфических приведем еще три примера возможностей противника:

- противник может перехватывать все шифрованные сообщения, но не имеет соответствующих им открытых текстов;
- противник может перехватывать все шифрованные сообщения и добывать соответствующие им открытые тексты;
- противник имеет доступ к шифру (но не к ключам!) и поэтому может зашифровывать и дешифровывать любую информацию.

Рекомендуем самостоятельно придумать еще несколько возможностей противника. Подскажем, например, использование так называемого «*вероятного слова*» в открытом тексте: противнику из каких-либо соображений известно, что в открытом тексте встречается конкретное слово. Иногда такая информация облегчает процесс вскрытия шифра.

На протяжении многих веков среди специалистов не утихали споры о стойкости шифров и о возможности построения абсолютно стойкого шифра. Приведем два характерных высказывания на этот счет.

Английский математик Чарльз Беббидж (XIX в): «Всякий человек, даже если он не знаком с техникой вскрытия шифров, твердо считает, что сможет изобрести абсолютно стойкий шифр, и чем более умен и образован этот человек, тем более твердо это убеждение. Я сам разделял эту уверенность в течение многих лет».

«Отец кибернетики» Норберт Винер (XX в): «Любой шифр может быть вскрыт, если только в этом есть настоятельная необходимость и информация, которую предполагается получить, стоит затраченных средств, усилий и времени...»

Мы вернемся к этому вопросу в этюде 2.5 после рассказа о работах Клода Шеннона.

1.7. Криптография и криптология

Криптпология — наука, состоящая из двух ветвей: криптографии и криптоанализа.

Криптография — наука о способах преобразования (шифрования) информации с целью ее защиты от незаконных пользователей.

Криптоанализ — наука (и практика ее применения) о методах и способах вскрытия шифров.

В последнее время наряду со словом «криптография» часто встречается и слово «криптология», но соотношение между ними не всегда понимается правильно. Сейчас происходит окончательное формирование этих научных дисциплин, уточняются их предмет и задачи.

Соотношение криптографии и криптоанализа очевидно: криптография — защита, т.е. разработка шифров, а криптоанализ — нападение, т.е. атака на шифры. Однако эти две дисциплины связаны друг с другом, и не бывает хороших криптографов, не владеющих методами криптоанализа. Дело в том, что стойкость разработанного шифра можно доказать только с помощью проведения различных атак на шифр, становясь мысленно в положение противника (см. этюды 1.6, 2.6).

1.8. Почему нужно много разных шифрмашин

Потому что не существует единого, подходящего для всех случаев способа шифрования информации. Выбор криптографической системы зависит от особенностей информации, ее ценности и возможностей владельцев по защите своей информации.

Прежде всего подчеркнем большое разнообразие видов защищаемой информации: документальная, телефонная, телевизионная, компьютерная и т.д. Каждый вид информации имеет свои специфические особенности, и эти особенности сильно влияют на выбор методов шифрования информации. Большое значение имеют объемы и требуемая скорость передачи шифрованной информации. Выбор вида шифра, его параметров и его стойкости существенно зависит от характера защищаемых секретов или тайны. Некоторые тайны (например, государственные, военные и др.) должны сохраняться десятилетиями, а некоторые (например, биржевые) — уже через несколько часов можно разгласить. Необходимо учитывать также и возможности того противника, от которого защищается данная информация. Одно дело — противостоять одиночке или даже банде уголовников, а другое дело мощной государственной структуре.

Из-за такого разнообразия требований приходится разрабатывать различные шифры, которые реализуются в сотнях типов шифрующих машин и устройств. Вместе с тем в наиболее развитых в криптографическом отношении странах существуют стандартные шифры: например, DES — в США и СКЗД — в России.

1.9. Зависимость криптографии от уровня технологий

Результаты криптографии реализуются в виде шифрующих устройств, встроенных в современные сети связи. Поэтому криптографы ограничены в выборе средств тем уровнем техники и технологии, который достигнут на данный момент. Такая зависимость отражается и на выборе используемого в криптографии математического аппарата.

Условно можно выделить три принципиально разные этапа в развитии математического аппарата криптографии.

До 40-х годов XX века были только электромеханические шифрмашины, поэтому и спектр математических преобразовании был ограничен: применялись в основном методы комбинаторного анализа и теории вероятностей.

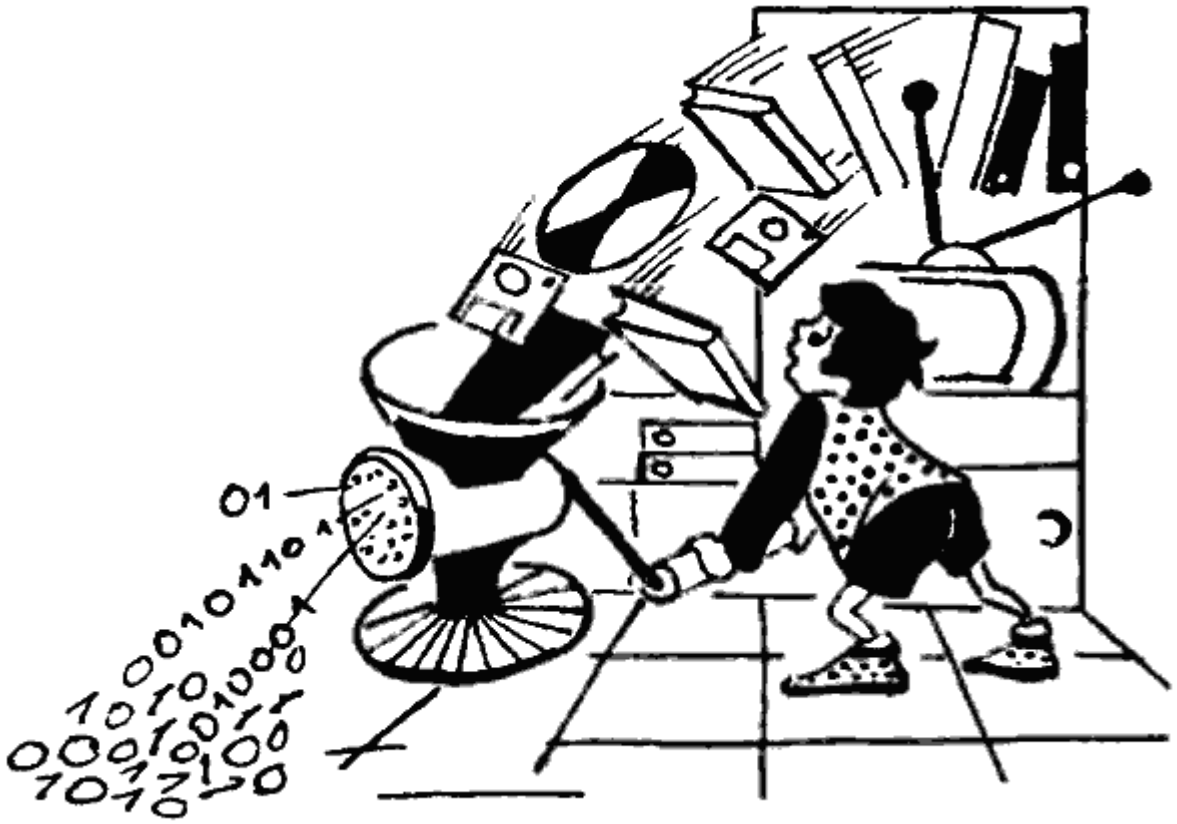
После появления электронной техники, а тем более компьютеров, сильно изменился и математический аппарат криптографии. Получили развитие прикладные идеи и методы теории информации, алгебры, теории конечных автоматов.

Работы Диффи и Хеллмэна (70-е годы) послужили толчком для бурного развития новых направлений математики: теории односторонних функций, доказательств с нулевым разглашением. Сейчас прогресс именно в этих направлениях определяет практические возможности криптографии.

Глава 2 Математические основы криптографии

Большое влияние на развитие криптографии оказали появившиеся в середине нашего века работы американского математика Клода Шеннона. В этих работах были заложены основы теории информации, а также был разработан математический аппарат для исследований во многих областях науки, связанных с информацией. В данной главе мы кратко ознакомим вас с основополагающими математическими понятиями и идеями, без знания которых успешная работа в области криптографии невозможна.

2.1. Приведение любой информации к двоичному виду



Для того, чтобы доказывать математические теоремы, нужно четко определить объекты, с которыми мы имеем дело. При шифровании текста необходимо, в первую очередь, знать, какие символы могут в нем встречаться, или, проще говоря, знать *алфавит*. Но алфавитов существует великое множество. Передаваемая информация может состоять и просто из наборов цифр, скажем, номера счетов в банке и деланные по ним выплаты. Поэтому естественно работать с некоторым обобщенным алфавитом — тогда одну и ту же теорему не нужно будет отдельно доказывать, например, для текстов на русском и на английском языке.

В теоретической криптографии принято работать с универсальным алфавитом, состоящим из всех *двоичных слов* некоторой длины. Двоичное слово длины n — это набор из n нулей и единиц. Соответствующий алфавит состоит из 2^nсимволов. Выбор такого алфавита объясняется многими соображениями.

При использовании компьютеров удобно представлять информацию в виде последовательностей нулей и единиц. Это, в частности, обусловлено применяемыми

техническими средствами: в компьютере используются элементы, которые могут находиться в одном из двух состояний. Одно из них обозначается «0», а другое — «1».

С другой стороны, слова в любом алфавите можно легко перевести в двоичные слова. Пусть мы имеем дело с текстами на русском языке и пусть буквы «е» и «ё», а также «и» и «й» не различаются, а пробел между словами считается отдельной буквой (обозначение: _). Тогда наш алфавит состоит из тридцати двух символов. Рассмотрим теперь *телеграфный код* — старое техническое применение *двоичной системы счисления*. Он состоит тоже из 32 символов — двоичных слов длины 5. (Подробно о двоичной и других системах счисления можно прочитать в брошюре С.В. Фомина «Системы счисления» из серии «Популярные лекции по математике».) Сопоставим каждой букве двоичное слово длины 5 следующим образом:

_ → 00000, *А* → 00001, *Б* → 00010, *В* → 00011, *Г* → 00100, *Д* → 00101, ... , *Ю* → 11110, *Я* → 11111.

Заменив в тексте каждую букву на соответствующее двоичное слово, получим двоичный вид нашей информации — некоторую последовательность нулей и единиц (или, как принято говорить, *битов*). Подобным образом можно поступить и с любым другим алфавитом.

На практике создаются специальные устройства, которые позволяют автоматически переводить вводимую человеком текстовую информацию в двоичную.

Более того, в настоящее время практически любая информация — речь, телевизионные сигналы, музыка и др. — может храниться и пересылаться в двоичном виде. Для работы с такой информацией используют специальные устройства: например, АЦП и ЦАП (аналого-цифровой и цифро-аналоговый преобразователи), устройства для цифровой записи и воспроизведения музыки.

Таким образом, двоичные слова и двоичные последовательности — типовые объекты в криптографических исследованиях.

Подумайте сами:

1. Докажите, что каждое натуральное число n единственным образом представляется в виде $n=2^k+a_{k-1}2^{k-1}+...+a_1 2+a_0$, где k — некоторе целое неотрицательное число, а каждое из чисел a_{k-1}, ..., a_0 — либо 0, либо 1.

2.2. Случайность и закономерность в двоичных последовательностях

Понятие последовательности известно еще со школьных лет. Однако последовательности, которые там изучались, были *детерминированными* — они однозначно восстанавливались по их нескольким элементам. Так, арифметическая и геометрическая прогрессии восстанавливаются по любым двум своим подряд идущим членам. Значения многочлена в целых точках также образуют детерминированную последовательность: она определяется любыми $n+1$ своими членами, где n— степень данного многочлена (докажите это!).

Но существуют и другие последовательности, так называемые *случайные*. Для них, в отличие от детерминированных, вообще говоря, нельзя определить очередной член последовательности, зная предыдущие. Опишем простейший способ получения двоичной случайной последовательности.

Пусть мы подбрасываем «правильную» монету. В зависимости от того, как она падает, полагаем очередной член последовательности равным 0 (орел) или 1 (решка). Как показывает опыт, обычно нельзя угадать, как монета упадет в очередной раз. Однако, если подбрасывать монету достаточно долго, примерно в половине случаев выпадет орел, а в половине — решка. Говорят, что монета падает случайным образом, причем в каждом подбрасывании с одинаковой*вероятностью* ½ выпадает орел (0) или решка (1).

Однако бывают ситуации («кривая монета»), когда орел и решка выпадают с разной вероятностью — p и q соответственно ($p \neq q$). Отметим, что $p+q=1$! В случайной двоичной последовательности, полученной на основе подбрасывания «кривой монеты», p можно считать частотой появления нуля, а q — частотой появления единицы.

Для тех кто изучал пределы, уточним: если обозначить через $S_0(k)$ число нулей, а через $S_1(k)$ — число единиц среди k первых членов нашей последовательности, то тогда предел отношения $S_0(k)/k$ равен p и предел отношения $S_1(k)/k$ равенq при k стремящемся к бесконечности.

$$\lim_{k\to\infty} \frac{S_0(k)}{k} = p, \quad \lim_{k\to\infty} \frac{S_1(k)}{k} = q.$$

Контрольный вопрос. Пусть мы случайным образом подбрасываем монету, причём $p=q=½$ и первые сто членов соответствующей последовательности равны 1 (100 раз подряд выпала решка). Чему равно вероятность того, что 101-ым членом этой последовательности снова будет 1?

Правильный ответ на этот вопрос: ½. Так как $q=½$, а случайность нашей последовательности как раз и означает, что каждый очередной её член равен 1 с вероятностью q *независимо* от того, какими были предыдущие её члены.

Обычно последовательности, с которыми на практике приходится иметь дело, вообще говоря, не строго случайные (неслучайные). Изучение случайных и неслучайных двоичных последовательностей имеет важное значение для криптографии. Например, выявление закономерностей в шифрованных сообщениях очень полезно при вскрытии шифра (см. этюд 2.7). В этюде 2.5 вы также узнаете, что для построения абсолютно стойкого шифра необходимо уметь получать совершенно случайный ключ.

Задачам различения случайной и неслучайной последовательностей, а также выявления закономерностей в неслучайных последовательностях посвящено много исследований в различных областях математики. Так, например, один из основных разделов математической статистики — это *проверка статистических гипотез*, в котором, в частности, разрабатываются методы различения гипотез о природе и характеристиках наблюдаемых последовательностей. Другой пример — это активно изучаемый в современной теоретической криптографии гипотетический объект — *псевдослучайный генератор*. При изучении этого объекта используются многочисленные результаты теории сложности алгоритмов и вычислений. Говоря неформально, псевдослучайный генератор вырабатывает такие последовательности, которые трудно отличить от случайных и из которых трудно извлечь закономерности. Строгие определения необходимых понятий выходят за рамки нашей книги.

Близким по духу, но более простым и хорошо известным, особенно для программистов, является такой объект, как *датчик случайных чисел*. Это — некоторое устройство или программа, которая вырабатывает *псевдослучайные последовательности*. Псевдослучайные последовательности в некоторых ситуациях считают неотличимыми от случайных, причем для разных ситуаций и задач подбирают подходящие датчики. Чем более сильные требования накладываются на случайность вырабатываемых последовательностей, тем более сложным является соответствующий датчик случайных чисел. Многие шифрмашины можно считать датчиками случайных чисел, удовлетворяющими очень высоким требованиям на качество вырабатываемых последовательностей.

Опишем, например, один простейший датчик, предложенный в 1949 году Д.Х. Лемером и в дальнейшем получивший название *линейного конгруэнтного метода*. Для заданного начального числа a_0 он вырабатывает бесконечную последовательность натуральных чисел $\{a_k\}$ по следующему *рекуррентному закону*:

$a_k = d + a_{k-1} \cdot \ell \pmod{N}$.

Здесь параметры датчика d, ℓ, N — некоторые целые числа. Запись $a = b \pmod{N}$, вообще говоря, означает, что $a - b$ делится на число N; в данном случае в качестве a_k берется остаток от деления $d + a_{k-1} \cdot \ell$ на N.

Поскольку все члены последовательности $\{a_k\}$ — неотрицательные целые числа, не превосходящие $N-1$, то среди них найдутся два одинаковых, скажем a_i и a_{i+t}. Тогда, как легко видеть, $a_i = a_{i+t}$ для $k \geq i$, т.е. последовательность —*периодическая* с длиной *периода* t. Конечно, периодичность не вполне согласуется с нашими представлениями о случайности, но, оказывается, можно подбирать такие параметры датчика, чтобы период был достаточно большим и у последовательности были многие признаки случайности.

Следует отметить, что «хорошей во всех отношениях случайной последовательности» практически не существует: насколько «хорошей» является случайная последовательность, зависит от ее назначения.

Подумайте сами:

1. Докажите следующее утверждение: вероятность того, что при k подбрасываниях кривой монеты ℓ раз выпадет орёл, равняется:

$$\frac{k!}{(k-l)!l!} p^l q^{k-l}$$

2. Придумайте такие числа d, ℓ и N, чтобы N было не слишком маленьким и длина периода последовательности, полученной линейным конгруэнтным методом, была близка к N.

3. Придумайте какой-нибудь свой датчик случайных чисел.

2.3. Что такое алгоритм и его сложность

Под *алгоритмом*, если говорить неформально, можно понимать четко описанную последовательность действий, приводящую к определенному результату.

Понятие алгоритма очень долго оставалось интуитивным понятием. Только в 30-е годы XX века в работах выдающихся математиков Д. Гильберта, А. Черча, С. Клини, Э. Поста и А. Тьюринга были предложены формальные определения алгоритма на основе понятия *рекурсивной функции* и на основе *описания алгоритмического процесса*. Тем самым формировалась теория алгоритмов — новое направление в математике, которое стало впоследствии теоретической основой развития вычислительной техники. В настоящее время теория алгоритмов бурно развивается, многие ее понятия проясняются и уточняются (*доказуемость*, *разрешимость*, *эффективность* и др.).

С нематематическими алгоритмами мы постоянно встречаемся в жизни (таковыми можно считать, например, рецепт приготовления борща или инструкцию о проведении экзамена в школе). Простейшим примером математического алгоритма может служить хорошо известный алгоритм Евклида, при помощи которого можно найти наибольший общий делитель двух чисел. А такой вид деятельности, как программирование — это постоянная работа с алгоритмами.

Очень важным понятием в математике (интуитивно ясным, но не очень просто формализуемым) является *сложность алгоритма*. Приведем простой пример. Пусть требуется угадать задуманное число, про которое известно, что оно натуральное и не превосходит 1000. Разрешается задавать вопросы, на которые можно ответить «да» или «нет». Одним из способов (алгоритмов) угадывания может быть такой: последовательно перебираются все числа от 1 до 1000 до тех пор, пока нужное число не будет найдено. В худшем случае для этого потребуется 999 вопросов. Однако можно предложить и другой алгоритм, позволяющий угадать число за 10 вопросов: сначала выясняется, больше ли угаданное число 500 или нет, если да, то больше 750 или нет и т.д. С каждым шагом число возможных кандидатов уменьшается в два раза. Здесь сложностью алгоритма можно считать число вопросов. Тогда первый алгоритм в 100 раз «сложнее» второго.

Если алгоритм проводит серии вычислений, сложностью алгоритма можно считать число совершаемых операций. При этом, если в алгоритме встречаются только умножение и сложение, под сложностью часто понимается только число умножений, поскольку эта операция требует существенно большего времени. На практике необходимо также учитывать стоимость операций, выполняемых алгоритмом, и т.п.

В математической теории сложности вычислений рассматриваются алгоритмы решения не конкретных задач, а так называемых *массовых задач*. Массовую задачу удобно представлять себе в виде бесконечной серии индивидуальных задач. Индивидуальная задача характеризуется некоторым *размером*, т.е. объемом входных данных, требуемых для описания этой задачи. Если размер индивидуальной задачи — некоторое натуральное число n, тогда сложность алгоритма решения массовой задачи становится функцией от n. Приведем два примера.

Рассмотрим алгоритм простого перебора всех двоичных ключей длины n. Ясно, что таких ключей — 2^n, и поэтому в данном алгоритме 2^n шагов, т.е. его сложность равна 2^n. Это — простейший пример *экспоненциального* алгоритма (его сложность выражается через n экспонентой). Большинство экспоненциальных алгоритмов — это просто варианты полного перебора.

Рассмотрим теперь алгоритм умножения столбиком двух n-значных чисел. Он состоит из n^2 умножений однозначных чисел, т.е. его сложность, измеренная количеством таких умножений, равна n^2. Это — простейший пример*полиномиального* алгоритма (его сложность выражается через n полиномом).

Достаточно очевидно, что для решения одной и той же математической задачи могут быть предложены различные алгоритмы. Поэтому под *сложностью задачи* понимают минимальную сложность алгоритмов ее решения. Возвращаясь теперь к этюду 1.6, можно сказать в новых терминах, что стойкость шифра — это сложность задачи его вскрытия.

В математике есть много задач, для решения которых пока не удалось построить полиномиальный алгоритм. К ним относится, например, задача коммивояжера: есть n городов, соединенных сетью дорог, и для каждой дороги указана стоимость проезда по ней; требуется указать такой маршрут, проходящий через все города, чтобы стоимость проезда по этому маршруту была минимальной.

Подумайте сами:

1. Можете ли вы предложить алгоритм умножения двух n-значных чисел, требующий меньшего числа умножений однозначных чисел, чем при умножении столбиком?

2.4. Шифры замены и перестановки

В своей работе «Математическая теория секретной связи» Клод Шеннон обобщил накопленный до него опыт разработки шифров. Оказалось, что даже в сложных шифрах в качестве типичных компонентов можно выделить *шифры замены*,*шифры перестановки* или их сочетания. Эти шифры можно считать как бы базовыми.

Шифр замены является простейшим, наиболее популярным шифром. Типичными примерами являются шифр Цезаря, «цифирная азбука» Петра Великого и «пляшущие человечки» А. Конан-Дойля. Как видно из самого названия, шифр замены осуществляет преобразование замены букв или других «частей» открытого текста на аналогичные «части» шифрованного текста. Понятно, что, увеличив алфавиты, т.е. объявив «части» буквами, можно любой шифр замены свести к замене букв. Теперь уже легко дать математическое описание шифра замены. Пусть X и Y — два алфавита открытого и соответственно шифрованного текстов, состоящие из одинакового числа символов. Пусть также $g : X \to Y$ — взаимно-однозначное отображение X в Y. Это значит, что каждой букве x алфавита X сопоставляется однозначно определенная буква y алфавита Y, которую мы обозначаем символом $g(x)$, причем разным буквам сопоставляются разные буквы. Тогда шифр замены действует так: открытый текст $x_1x_2...x_n$ преобразуется в шифрованный текст $g(x_1)g(x_2)...g(x_n)$. К вопросу о вскрытии шифра замены мы вернемся в этюде 2.8.

Шифр перестановки, как видно из названия, осуществляет преобразование перестановки букв в открытом тексте. Типичным и древнейшим примером шифра перестановки является шифр «Сциталь». Обычно открытый текст разбивается на отрезки равной длины, и каждый отрезок шифруется (т.е. в нем переставляются буквы) независимо. Пусть, например, длина отрезков равна n и σ — взаимно-однозначное отображение множества $\{1,2, ..., n\}$ в себя. Тогда шифр перестановки действует так: отрезок открытого текста $x_1...x_n$ преобразуется в отрезок шифрованного текста $x_{\sigma(1)}...x_{\sigma(n)}$.

Важной проблемой при практическом использовании шифров замены и перестановки является проблема удобного запоминания отображений g и σ. Ясно, что легко запоминать некоторые отображения: например, отображения «небольших» размеров, отображения, реализуемые каким-нибудь предметом (сциталь в шифре «Сциталь» и т.п.). Если же отображение «большого» размера, то процесс запоминания сильно усложняется. Например, широко известны*биграммные шифры*. В них преобразовывались биграммы (пары букв). Поскольку количество биграмм превышает 1000, то на практике биграммные шифры выглядят как специальные книжки.

Для облегчения запоминания отображений g и σ изобретались различные хитроумные способы. Правда, «расплатой» за это было упрощение используемых отображений и тем самым уменьшение стойкости шифров.

Популярным способом запоминания отображения σ, т.е. шифра перестановки, является следующий. Пусть, например, n=20. Берем прямоугольную таблицу размера 4x5, вписываем в нее открытый текст «по строкам», а шифрованный текст считываем «по столбцам». Возможны и более хитрые способы вписывания и считывания.

Подумайте сами:

1. Выпишите отображение g для шифра Цезаря.
2. Выпишите отображение σ для описанного шифра перестановки — прямоугольника 4×5.

2.5. Возможен ли абсолютно стойкий шифр

Да, и единственным таким шифром является какая-нибудь форма так называемой *ленты однократного использования*, в которой открытый текст «объединяется» с полностью случайным ключом такой же длины, Этот результат был доказан К. Шенноном с помощью разработанного им теоретико-информационного метода исследования шифров. Мы не будем здесь останавливаться на этом подробно, но заинтересованному читателю рекомендуем изучить работу К. Шеннона, она указана в списке литературы.

Обсудим особенности строения абсолютно стойкого шифра и возможности его практического использования. Типичным и наиболее простым примером реализации абсолютно стойкого шифра является *шифр Вернама*, который осуществляет побитовое сложение n-битового открытого текста и n-битового ключа: $y_i=x_1 \oplus k_i$, i=1, ..., n.

Здесь x_1, ..., x_n — открытый текст, k_1, ..., k_n — ключ, y_1, ..., y_n — шифрованный текст; символы складываются по таким правилам: $0 \oplus 0=0$, $0 \oplus 1=1 \oplus 0=1$, $1 \oplus 1=0$.

Подчеркнем теперь, что для абсолютной стойкости существенным является каждое из следующих требований к ленте однократного использования:

1) полная случайность (равновероятность) ключа (это, в частности, означает, что ключ нельзя вырабатывать с помощью какого-либо детерминированного устройства);

2) равенство длины ключа и длины открытого текста;

3) однократность использования ключа.

В случае нарушения хотя бы одного из этих условий шифр перестает быть абсолютно стойким и появляются принципиальные возможности для его вскрытия (хотя они могут быть трудно реализуемыми).

Но, оказывается, именно эти условия и делают абсолютно стойкий шифр очень дорогим и непрактичным. Прежде чем пользоваться таким шифром, мы должны обеспечить всех абонентов достаточным запасом случайных ключей и исключить возможность их повторного применения. А это сделать необычайно трудно и дорого. Как отмечал Д. Кан: «Проблема создания, регистрации, распространения и отмены ключей может показаться не слишком сложной тому, кто не имеет опыта передачи сообщений по каналам военной связи, но в военное время объем передаваемых сообщений ставит в тупик даже профессиональных связистов. За сутки могут быть зашифрованы сотни тысяч слов. Создание миллионов ключевых знаков потребовало бы огромных финансовых издержек и было бы сопряжено с большими затратами времени. Так как каждый текст должен иметь свой собственный, единственный и неповторимый ключ, применение идеальной системы потребовало бы передачи, по крайней мере, такого количества знаков, которое эквивалентно всему объему передаваемой военной информации».

В силу указанных причин, абсолютно стойкие шифры применяются только в сетях связи с небольшим объемом передаваемой информации, обычно это сети для передачи особо важной государственной информации.

2.6. Стойкость теоретическая и практическая

Теперь мы уже понимаем, что чаще всего для защиты своей информации законные пользователи вынуждены применять неабсолютно стойкие шифры. Такие шифры, по крайней мере теоретически, могут быть вскрыты. Вопрос только в том, хватит ли у противника сил, средств и времени для разработки и реализации соответствующих алгоритмов.

Обычно эту мысль выражают так: противник с неограниченными ресурсами может вскрыть любой неабсолютно стойкий шифр.

Как же должен действовать в этой ситуации законный пользователь, выбирая для себя шифр? Лучше всего, конечно, было бы доказать, что никакой противник не может вскрыть выбранный шифр, скажем, за 10 лет и тем самым получить теоретическую оценку стойкости. К сожалению, математическая теория еще не дает нужных теорем — они относятся к нерешенной *проблеме нижних оценок сложности задач*.

Поэтому у пользователя остается единственный путь — получение практических оценок стойкости. Этот путь состоит из следующих этапов:

— понять и четко сформулировать, от какого противника мы собираемся защищать информацию; необходимо уяснить, что именно противник знает или сможет узнать о системе шифра, какие силы и средства он сможет применить для его вскрытия;

— мысленно стать в положение противника и пытаться с его позиций атаковать шифр, т.е. разрабатывать различные алгоритмы вскрытия шифра; при этом необходимо в максимальной мере обеспечить моделирование сил, средств и возможностей противника;

— наилучший из разработанных алгоритмов использовать для практической оценки стойкости шифра.

Здесь полезно для иллюстрации упомянуть о двух простейших методах вскрытия шифра: случайного угадывания ключа (он срабатывает с маленькой вероятностью, зато имеет маленькую сложность) и перебора всех подряд ключей вплоть до нахождения истинного (он срабатывает всегда, зато имеет очень большую сложность).

2.7. Всегда ли нужна атака на ключ

Нет, для некоторых шифров можно сразу, даже не зная ключа, восстанавливать открытый текст по шифрованному.

Эту мысль удобнее всего проиллюстрировать на примере шифра замены, для которого уже давно разработаны методы вскрытия.

Напомним, что шифр замены математически описывается с помощью некоторой подстановки g (см. этюд 2.4). Такой шифр преобразует открытый текст в шифрованный по следующему правилу: каждая буква x заменяется на букву $g(x)$. Вскрытие шифра основано на двух следующих закономерностях:

1) в осмысленных текстах любого естественного языка различные буквы встречаются с разной частотой, а действие подстановки g «переносит» эту закономерность на шифрованный текст;

2) любой естественный язык обладает так называемой *избыточностью*, что позволяет с большой вероятностью «угадывать» смысл сообщения, даже если часть букв в сообщении неизвестна.

Приведем для примера относительные частоты букв алфавита русского языка.

N	Буква	Относит. частота
1	а	0,062
2	б	0,014
3	в	0,038
4	г	0,013
5	д	0,025
6	е, ё	0,072
7	ж	0,007
8	з	0,016
9	и	0,062
10	й	0,010
11	к	0,028

12	л	0,035
13	м	0,026
14	н	0,053
15	о	0,090
16	п	0,023
17	р	0,040
18	с	0,045
19	т	0,053
20	у	0,021
21	ф	0,002
22	х	0,009
23	ц	0,004
24	ч	0,012
25	ш	0,006
26	щ	0,003
27	ы	0,016
28	ъ, ь	0,014
29	э	0,003
30	ю	0,006
31	я	0,018
32	пробел	0,175

Подобные таблицы используются для вскрытия шифра простой замены следующим образом. Составляем таблицу частот встречаемости букв в шифртексте. Считаем, что при замене наиболее частые буквы переходят в наиболее частые. Последовательно перебирая различные варианты, пытаемся либо прийти к противоречию с законами русского языка, либо получить читаемые куски сообщения. Далее по возможности продляем читаемые куски либо по смыслу, либо по законам русского языка.

Подробный разбор даже одного примера может занять слишком много места. Любознательным читателям рекомендуем проделать это самостоятельно для какого-нибудь своего шифра замены. Можно также прочитать подробное описание трех примеров:

— в рассказе Э. По «Золотой жук»;

— в рассказе А. Конан-Дойля «Пляшущие человечки»;

— в книге М.Н. Аршинова и Л.Е. Садовского «Коды и математика».

2.8. Криптография, комбинаторные алгоритмы и вычислительная техника

Зададимся теперь вопросом: от прогресса в каких областях науки зависят оценки практической стойкости шифров? Внимательный читатель сам из предыдущего изложения ответит на этот вопрос: в первую очередь это — теория сложности алгоритмов и вычислений, а также сложность реализации алгоритмов на вычислительной технике. В последние годы эти области бурно развиваются, в них получены интересные результаты, которые, в частности, влияют на оценки практической стойкости шифров. Многие полезные результаты носят характер «ухищрений» для ускорения алгоритмов и поэтому быстро входят в массовую практику программистов. Особенно это относится к области *комбинаторных алгоритмов*, выросшей из хорошо известных типичных задач быстрого *поиска* и *сортировки данных*. Систематическое подробное изложение этих вопросов содержится в популярном трехтомнике Д. Кнута «Искусство программирования для ЭВМ».

Отметим, что к области комбинаторных алгоритмов относятся также алгоритмы для хорошо известных игр-головоломок типа «кубика Рубика».

Алгоритмы вскрытия шифров, как правило, используют большое количество различных приемов сокращения перебора ключей (или других элементов шифра), а также поиска, сравнения и отбраковки данных. Поэтому в оценки стойкости шифров входят различные оценки из теории комбинаторных алгоритмов.

Подумайте сами:

1. Докажите, что наименьший элемент среди чисел $\{x_1, \ldots, x_n\}$ нельзя найти меньше, чем за $n-1$ сравнение.

2. Предложите алгоритм расположения чисел $\{x_1, \ldots, x_n\}$ в порядке возрастания, использующий меньше, чем $n(n-1)/2$ сравнений (т.е. более эффективный, чем тривиальный алгоритм последовательного сравнения каждого числа с каждым).

3. На полке в беспорядке стоят n томов собрания сочинений. Хозяин, увидев два тома, стоящие в неправильном порядке, меняет их местами. Докажите, что не позднее, чем через n^2 таких перестановок, тома будут расставлены по порядку.

4. На сортировочной станции имеется несколько поездов. Разрешается либо расцепить поезд, состоящий из нескольких вагонов, на два поезда, либо удалить поезд, если в нём всего один вагон. Докажите, что, выполняя эти действия в произвольном порядке, мы рано или поздно удалим все вагоны.

5. Задумано и введено в компьютер n натуральных чисел $a_1, \ldots, a_n$. За один шаг разрешается ввести в компьютер любые n других натуральных чисел $b_1, \ldots, b_n$. После этого компьютер вычисляет сумму $a_1b_1 + \ldots + a_nb_n$ и выводит результат на экран. Ясно, что этот результат содержит

некоторую информацию о задуманных числах. За какое минимальное число шагов всегда можно угадать задуманные числа?

Глава 3 Новые направления

В 1983 году в книжке «Коды и математика» М.Н. Аршинова и Л.Е. Садовского (библиотечка «Квант») было написано: «Приемов тайнописи — великое множество, и, скорее всего, это та область, где уже нет нужды придумывать что-нибудь существенно новое». Однако это было очередное большое заблуждение относительно криптографии. Еще в 1976 году была опубликована работа молодых американских математиков У. Диффи и М.Э. Хеллмэна «Новые направления в криптографии», которая не только существенно изменила криптографию, но и привела к появлению и бурному развитию новых направлений в математике. В настоящей главе мы опишем основные понятия «новой криптографии».

3.1. Односторонняя функция

Односторонней называется функция F:$X \to Y$, обладающая двумя свойствами:

а) существует полиномиальный алгоритм вычисления значений $F(x)$;

б) не существует полиномиального алгоритма *инвертирования* функции F, т.е. решения уравнения $F(x)=y$ относительно x.

Отметим, что односторонняя функция существенно отличается от функций, привычных со школьной скамьи, из-за ограничений на сложность ее вычисления и инвертирования. Это новое понятие в математике введено в 1975 году Диффи и Хеллмэном. Но за истекшие 19 лет так и не удалось построить ни одного примера односторонней функции. Тем не менее, активное изучение свойств этого, пока гипотетического, математического объекта позволило установить его связь с другими более изученными объектами. При этом удалось доказать, что проблема существования односторонней функции эквивалентна одной из хорошо известных нерешенных проблем — «совпадают ли *классы сложностей P и NP*»? Строгое определение классов *P* и *NP* выходит далеко за рамки настоящей книги. Подготовленным читателям рекомендуем фундаментальную монографию М. Гэри и Д. Джонсона «Вычислительные машины и труднорешаемые задачи».

Говоря неформально, класс *P* состоит из задач с полиномиальной сложностью. Более строго, класс *P* — это класс *языков*, распознаваемых за полиномиальное время на *детерминированной машине Тьюринга*. Если такую машину Тьюринга дополнить гипотетической способностью «угадывания», получается более сильная модель — *недетерминированная машина Тьюринга*. Класс *NP* — это класс языков, распознаваемых за полиномиальное время на недетерминированной машине Тьюринга. Проблема совпадения классов *P* и *NP* — это проблема соотношения возможностей двух моделей вычислений: детерминированная и недетерминированная машина Тьюринга.

Другим понятием, более близким к традиционной криптографии, в которой есть секретный ключ, является понятие *односторонней функции с секретом*. Иногда еще употребляются термины *функция с ловушкой*, *функция опускной двери*(английское название: one-way trap-door function).

Односторонней функцией с секретом K называется функция F_K: $X \to Y$, зависящая от параметра *K* и обладающая тремя свойствами:

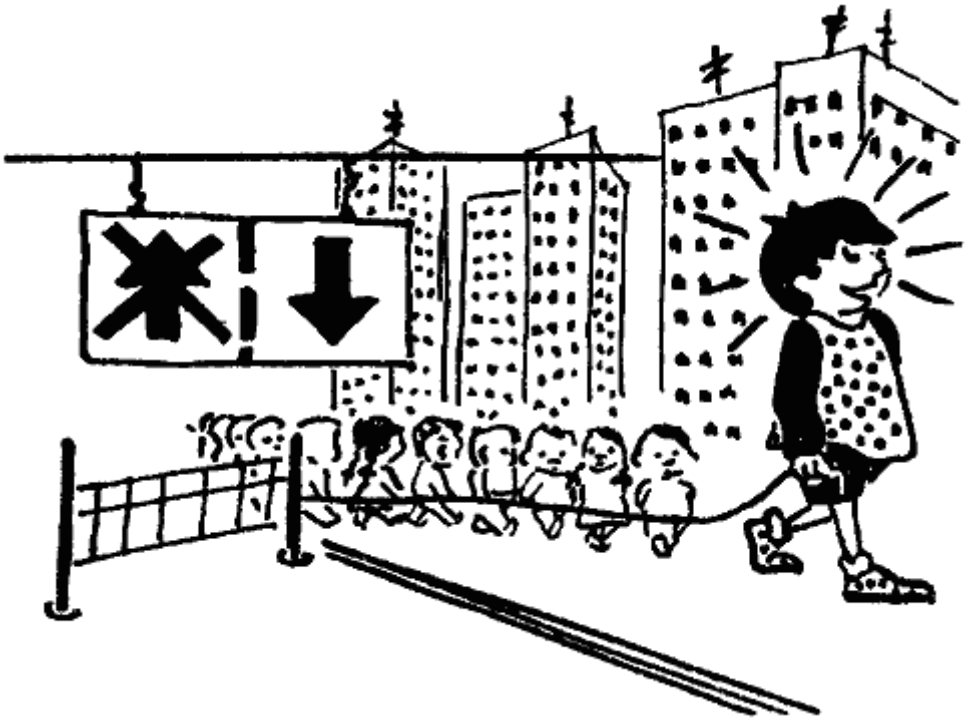
а) при любом *K* существует полиномиальный алгоритм вычисления значений $F_K(x)$;

б) при неизвестном *K* не существует полиномиального алгоритма инвертирования F_K;

в) при известном *K* существует полиномиальный алгоритм инвертирования F_K.

Про существование односторонних функций с секретом можно сказать то же самое, что было сказано ранее про односторонние функции. Для практических целей криптографии было построено несколько функций, которые могут оказаться односторонними. Это означает, что для них свойство б) пока строго не доказано, но известно, что задача инвертирования эквивалентна некоторой давно изучаемой трудной математической задаче. Примеры таких функций приводятся

в этюдах 3.5, 3.6, 3.7. Стоит отметить, что для некоторых кандидатов на звание односторонней функции были найдены полиномиальные алгоритмы инвертирования и тем самым доказано, что эти функции не являются односторонними.



3.2. Что даёт односторонняя функция для криптографии

Применение односторонней функции в криптографии позволяет:

1) организовать обмен шифрованными сообщениями с использованием только открытых каналов связи, т.е. отказаться от секретных каналов связи для предварительного обмена ключами;

2) включить в задачу вскрытия шифра трудную математическую задачу и тем самым повысить обоснованность стойкости шифра;

3) решать новые криптографические задачи, отличные от шифрования (*цифровая подпись* и др.).

Прежде чем разбирать конкретные примеры, опишем идею Диффи и Хеллмэна в общем виде.

Пользователь A, который хочет получать шифрованные сообщения, должен сначала выбрать какую-нибудь одностороннюю функцию F_K с секретом K. Он сообщает всем заинтересованным (например, публикует) описание функции F_K в качестве своего алгоритма шифрования. Но при этом значение секрета K он никому не сообщает и держит в секрете. Если теперь пользователь B хочет послать A защищаемую информацию $x \in X$, то он вычисляет $F_K(x)$ и посылает по открытому каналу к A. Поскольку A для своего секрета K умеет инвертировать F_K, то он вычисляет x по полученному $F_K(x)$. Никто другой не знает K и поэтому в силу свойства б) односторонней функции с секретом не сможет за полиномиальное время по известному шифрованному сообщению $F_K(x)$ вычислить защищаемую информацию x.

Таким образом, построена система передачи защищаемой информации, причем выполнены два новых свойства:

1) для передачи сообщений не требуется предварительный обмен ключами по секретному каналу связи;

2) стойкость шифра зависит от сложности решения трудной математической задачи инвертирования односторонней функции с секретом.

Описанную систему называют *криптосистемой с открытым ключом*, поскольку алгоритм шифрования F_K является общедоступным или открытым. В последнее время такие криптосистемы

еще называют *асимметричными*, поскольку в них есть асимметрия в алгоритмах: алгоритмы шифрования и дешифрования различны. В отличие от таких систем традиционные шифры называют *симметричными*: в них ключ для шифрования и дешифрования один и тот же, и именно поэтому его нужно хранить в секрете. Для асимметричных систем алгоритм шифрования общеизвестен, но восстановить по нему алгоритм дешифрования за полиномиальное время невозможно.

Описанную выше идею Диффи и Хеллмэн предложили использовать также для цифровой подписи сообщений, которую невозможно подделать за полиномиальное время. Пусть пользователю *A* необходимо подписать сообщение *x*. Он, зная секрет *K*, находит такое *y*, что $F_K(y) = x$, и посылает *y* пользователю *B* в качестве своей цифровой подписи. Пользователь *B* хранит *y* в качестве доказательства того, что *A* подписал сообщение *x*. Это доказательство неопровержимо, поскольку никто другой в силу свойства б) односторонней функции с секретом не сможет за полиномиальное время по известному сообщению $x=F_K(y)$ подделать цифровую подпись *y*.

В дальнейшем на основе аналогичных рассуждений был предложен еще целый ряд так называемых *криптографических протоколов*. Эти протоколы позволили решить много новых задач взаимодействия удаленных пользователей по техническим каналам связи в условиях различных угроз (подробнее об этом см. этюд 3.8).

3.3. Числа и поля

Занимаясь математикой, вы постоянно пользуетесь очевидными свойствами действительных чисел, даже не замечая этого, например: сумма чисел не зависит от порядка слагаемых.

Приведем основные свойства операций сложения и умножения на множестве действительных чисел *F*.

1) Для каждых трех элементов $a, b, c \in F$ $a+(b+c)=(a+b)+c$.

2) В множестве *F* есть элемент 0 такой, что для каждого $a \in F$ $a+0=0+a=a$.

3) Для каждого элемента $a \in F$ существует такой элемент $x \in F$, что $a+x=x+a=0$ (такой элемент называется противоположным к данному).

4) Для каждых двух элементов $a, b \in F$ $a+b=b+a$.

5) Для каждых трех элементов $a, b, c \in F$ $a\cdot(b\cdot c)=(a\cdot b)\cdot c$.

6) В множестве F есть элемент 1 (не равный 0) такой, что для каждого $a \in F$ $a \cdot 1 = 1 \cdot a = a$.

7) Для каждого элемента $a \in F$, $a \neq 0$ существует такой элемент $x \in F$, что $a \cdot x = x \cdot a = 1$ (такой элемент называется обратным к данному).

8) Для каждых двух элементов a, $b \in F$ $a \cdot b = b \cdot a$.

9) Для каждых трех элементов a, b, $c \in F$ $a \cdot (b + c) = a \cdot b + a \cdot c$.

Свойства 1) – 4) — это свойства операции сложения, свойства 5) – 8) — свойства операции умножения, а свойство 9) устанавливает связь между этими двумя операциями.

Оказывается, в математике существует много других множеств с парами операций на них, обладающих теми же самыми свойствами. Для таких множеств есть даже специальное название: *поле*.

Полем называется множество F с двумя отображениями («операциями»), каждое из которых сопоставляет любой паре элементов из F однозначно определенный третий элемент из F, и эти отображения (условно обозначаемые «+» и «·») удовлетворяют девяти аксиомам (свойствам), приведенным выше.

Особенно важными для криптографии являются *конечные поля*. Сконструируем одно из таких полей.

Пусть p — простое число. Рассмотрим множество чисел $\{0, 1, 2, ..., p-1\}$ с операциями сложения и умножения по модулю p (суммой двух чисел считаем остаток от деления на p обычной суммы, произведением — остаток от деления на p обычного произведения). Легко проверить, что свойства 1) – 4) выполнены: для свойств 1) и 4) это очевидно, элемент 0 в свойстве 2) — это обычный нуль, противоположный к элементу a в свойстве 3) — это элемент $p - a$. Так же легко проверяются свойства 5), 6), 8) и 9). Свойство 7) надо доказывать. Предлагаем вам доказать это самостоятельно, поясним только идею: для каждого $a \in \{0, 1, 2, ..., p-1\}$ требуется найти такие x и y, что $ax = 1 + py$, т.е. $ax - py = 1$, а такие x и y всегда можно найти с помощью алгоритма Евклида.

Конечное поле — очень интересный математический объект. Оказывается, например, что число элементов в конечном поле может быть только степенью простого числа, и наоборот, для любого простого числа p и для любого натурального числа n существует и в некотором смысле единственное поле из p^n элементов. Для него введено даже специальное обозначение: $GF(p^n)$.

Поясним более подробно, в каком смысле поле из p^n элементов единственно. В математике принято не различать многие объекты, изучаемые свойства которых совпадают. Например, для того, чтобы складывать и умножать, вовсе не обязательно учить отдельно таблицы сложения и умножения для яблок, и отдельно — для стульев. Достаточно уметь складывать числа. Число в данной ситуации можно представлять как количество единиц некоторого обобщенного продукта, неважно какого. В теории полей два поля F и G считаются «одинаковыми» или *изоморфными*, если можно построить такое взаимно-однозначное отображение $s: F \to G$, чтобы для любых $x_1, x_2 \in F$ выполнялись условия $s(x_1 + x_2) = s(x_1) + s(x_2)$, $s(x_1 x_2) = s(x_1) s(x_2)$. Фактически это означает, что можно взаимно-однозначно сопоставить всем элементам одного поля элементы другого так, что таблицы умножения и сложения в этих полях будут «одинаковыми». Легко, например, доказать, что при изоморфизме нуль переходит в нуль, единица — в единицу.

Яркий пример использования полей в криптографии вы найдете в этюде 3.5, описывающем криптосистему RSA. Для ее полного понимания рекомендуем решить (или прочитать в любой книге по теории чисел, например, в книге И.М. Виноградова «Основы теории чисел» или в книге О. Оре «Приглашение в теорию чисел») приведенные ниже задачи.

Подумайте сами:

1. Функцией Эйлера (обозначение $\varphi(n)$) называется количество неотрицательных целых чисел, меньших n и взаимно простых с n. Пусть $n = p_1^{\alpha_1} \cdot ... \cdot p_k^{\alpha_k}$, где p_1, ..., p_k — различные простые числа, а α_1, ..., α_k — натуральные. Доказать, что

$$\varphi(n) = n \cdot (1 - \frac{1}{p_1}) \cdot \ldots \cdot (1 - \frac{1}{p_k})$$

2. (Малая теорема Ферма). Пусть *p* — простое число, *a* — число взаимно простое с *p*. Докажите, что тогда

$$a^{p-1} = 1 \,(\mathrm{mod} p)$$

3. (Теорема Эйлера). Пусть *a* и *n* — взаимно простые числа. Докажите, что тогда

$$a^{\varphi(n)} = 1 \,(\mathrm{mod} n)$$

3.4. Проблемы факторизации чисел и дискретного логарифмирования

Еще в младших классах школы все решают задачи по разложению чисел на простые множители. Делается это просто делением данного числа на последовательные простые числа. Если число большое, то этот алгоритм будет работать долго (даже на компьютере). Если же число очень большое (скажем, состоит из 200 знаков), самому современному компьютеру могут понадобиться годы работы. И, как это ни странно, до сих пор математики не придумали никакого другого алгоритма, работающего существенно быстрее. Проблема построения такого алгоритма называется проблемой факторизации чисел. С другой стороны, существуют быстрые алгоритмы, позволяющие с большой вероятностью определять, является ли данное число простым или нет (но никакого разложения числа на простые множители эти алгоритмы не находят).

Криптографические приложения проблемы факторизации чисел и, особенно, заинтересованность пользователей банковских систем цифровой подписи привели к резкому увеличению исследований, связанных с разложением чисел на множители. В последние годы благодаря применению тонких методов теории чисел и алгебраической геометрии было разработано несколько эффективных алгоритмов факторизации. Наилучший из таких алгоритмов еще не является полиномиальным, но уже и не экспоненциальный, он относится к классу так

называемых *субэкспоненциальных* алгоритмов (говоря строго, его сложность превосходит любой полином от n, но меньше, чем 2^N, где $N=n^\varepsilon$ для любого $\varepsilon>0$).

Среди последних достижений в этой области можно упомянуть об успехе Ленстры и Монасси, разложивших в июне 1990 года 155-разрядное число на три простых. Для этого они использовали 1000 объединенных ЭВМ и шесть недель их машинного времени. Вычисления проводились с помощью алгоритма английского математика Дж. Полларда. Ленстра и Монасси считают, что в настоящее время (1991 г.) можно в течение года разложить новые классы целых чисел длиной до 155 разрядов, затратив на это $200 млн.

Еще одна большая проблема — дискретное логарифмирование в конечных полях. Пусть, например, нам даны элементы a и b из конечного поля F, причем известно, что $a=b^x$ при некотором натуральном x. Задача дискретного логарифмирования состоит в том, чтобы определить это x. Можно, разумеется, просто перебирать последовательно все натуральные числа, проверяя, выполнено ли указанное равенство, но это будет экспоненциальный алгоритм. Пока наилучший из разработанных математиками алгоритмов дискретного логарифмирования является субэкспоненциальным.

В настоящее время эти описанные трудные математические проблемы имеют многочисленные криптографические приложения (см. этюды 3.5, 3.6, 3.7).

3.5. Криптосистема RSA

В этюде 3.2 описано, как Диффи и Хеллмэн с помощью односторонней функции с секретом построили криптосистему с открытым ключом. Правда, они не предложили функций, удобных для реализации.

Однако уже в начале 1977 г. американские специалисты по компьютерным наукам Р. Ривест, А. Шамир и Л. Адлеман придумали одну такую функцию. Система на основе этой функции оказалась очень практичной и получила широкое распространение под названием «система RSA» по первым английским буквам фамилий авторов.

Опишем систему RSA. При этом мы будем использовать без подробных пояснений обозначения и результаты этюдов 3.2 и 3.3. Пусть $n=pq$, где p и q — большие простые числа, а e — некоторое число, взаимно простое с $\varphi(n)$. Найдем число d из уравнения: $d\cdot e=1(\mathrm{mod}\varphi(n))$.

Числа p, q и d будем считать секретными и обозначим секрет $K=\{p, q, d\}$. Числа n и e будем считать общедоступными. Множества открытых сообщений X и шифрованных сообщений Y будем считать равными: $X = Y = \{1, 2, \ldots , n-1\}$.

Функцию $F_K : X \to Y$ определим равенством: $F_K(x) = x^e(\mathrm{mod}n)$.

Свойство а) односторонней функции с секретом выполнено для F_K очевидным образом. Проверим свойство в). Для этого просто укажем, как при известном K инвертировать функцию F_K:

решением уравнения $F_K(x) = y$ будет $x = y^d (\mathrm{mod} n)$. Подробное доказательство этого факта оставляем читателю, приведем лишь необходимые выкладки без комментариев:

$d \cdot e = \varphi(n) \cdot m + 1$

$(x^e)^d (\mathrm{mod} n) = x^{\varphi(n) \cdot m + 1} (\mathrm{mod} n) = (x^{\varphi(n)})^m \cdot x (\mathrm{mod} n) = (1)^m \cdot x (\mathrm{mod} n) = x$.

Свойство б) для функции F_K строго не доказано. Пока общепризнано, что для инвертирования F_K необходимо разложить n на множители, а, как указывалось в этюде 3.4, задача факторизации целых чисел относится к трудным математическим задачам.

Таким образом, описанную функцию F_K можно считать кандидатом на звание односторонней функции с секретом. Система RSA строится с помощью этой функции так, как рассказано в этюде 3.2.

В газете «Известия» за 29 апреля 1994 г. под заголовком «Сверхсекретный шифр разгадан за 17 лет» появилось следующее сообщение: «Когда в 1977 году математики Рональд Ривест, Ади Шамир и Леонард Адлеман зашифровали фразу из нескольких слов, используя комбинацию из 129 цифр, они утверждали, что на разгадку понадобятся триллионы лет. Однако ключ к самому сложному в мире шифру «РСА-129» (RSA) был найден за 17 лет... Разгадка шифра за такой относительно короткий срок имеет огромное значение для государственных организаций и предпринимателей, которые пользуются аналогичными длинными цифровыми кодами для защиты секретных сведений в своих компьютерных базах данных...» Пока это сообщение не подтверждено научными публикациями, ясно лишь, что речь идет о том, что удалось разложить на множители то 129-значное число, которое было использовано в 1977 году. Но уже давно в реальных системах RSA используются более длинные числа.

Подумайте сами:

1. Разберите примеры работы системы RSA, приведённые на стр. 241–243 книги М. Гарднера «От мозаик Пенроуза к надёжным шрифтам».

3.6. Открытое распределение ключей

Кроме принципа построения криптосистемы с открытым ключом, Диффи и Хеллмэн в той же работе предложили еще одну новую идею — *открытое распределение ключей*. Они задались вопросом: можно ли организовать такую процедуру взаимодействия абонентов *A* и *B* по открытым каналам связи, чтобы решить следующие задачи:

1) вначале у *A* и *B* нет никакой общей секретной информации, но в конце процедуры такая общая секретная информация (общий ключ) у *A* и *B* появляется, т.е. вырабатывается;

2) противник, который перехватывает все передачи информации и знает, что хотят получить *A* и *B*, тем не менее не может восстановить выработанный общий ключ *A* и *B*.

Диффи и Хеллмэн предложили решать эти задачи с помощью функции $F(x) = a^x (\mathrm{mod} p)$, где p — большое простое число, x — произвольное натуральное число, a — некоторый *примитивный элемент* поля $GF(p)$.

Примитивным называется такой элемент a из $GF(p)$, что каждый элемент поля, за исключением нуля, может быть представлен в виде степени a. Можно доказать, хотя это и не просто, что примитивный элемент всегда существует.

Общепризнано, что инвертирование функции $a^x (\mathrm{mod} p)$, т.е. дискретное логарифмирование, является трудной математической задачей (см. этюд 3.4).

Сама процедура или, как принято говорить, *протокол выработки общего ключа* описывается следующим образом.

Числа p и a считаются общедоступными.

Абоненты *A* и *B* независимо друг от друга случайно выбирают по одному натуральному числу — скажем $x(A)$ и $x(B)$. Эти элементы они держат в секрете. Далее каждый из них вычисляет новый элемент:

$y(A)=a^{x(A)} (\mathrm{mod} p)$, $y(B)=a^{x(B)} (\mathrm{mod} p)$.

Потом они обмениваются этими элементами по каналу связи. Теперь абонент *A*, получив $y(B)$ и зная свой секретный элемент $x(A)$, вычисляет новый элемент

$y(B)^{x(A)} (\mathrm{mod} p)=(a^{x(B)})^{x(A)} (\mathrm{mod} p)$.

Аналогично поступает абонент *B*:

$y(A)^{x(B)} (\mathrm{mod} p)=(a^{x(A)})^{x(B)} (\mathrm{mod} p)$.

Из свойств поля следует, что тем самым у *A* и *B* появился общий элемент поля, равный $a^{x(A)x(B)}$. Этот элемент и объявляется общим ключом *A* и *B*.

Из описания протокола видно, что противник знает p, a, $a^{x(A)}$, $a^{x(B)}$, не знает $x(A)$ и $x(B)$ и хочет узнать $a^{x(A)x(B)}$. В настоящее время нет алгоритмов действий противника, более эффективных, чем дискретное логарифмирование, а это — трудная математическая задача. (Рекомендуем самостоятельно найти за противника общий ключ, используя алгоритм дискретного логарифмирования и не принимая во внимание вопросы его сложности.)

3.7. Цифровая подпись

Идея *цифровой подписи* (иногда ее еще называют *электронной подписью*) была предложена Диффи и Хеллмэном. Суть идеи — в использовании односторонней функции с секретом F_K (см. этюд 3.2). В настоящее время эта идея реализована в большом количестве систем передачи данных, особенно банковских. Сообщение, подписанное цифровой подписью, можно представлять себе как пару (x, y), где x — сообщение (платежное поручение в примере с банком и т.п.), F_K: $X \to Y$ — односторонняя функция, известная всем взаимодействующим абонентам, y — решение уравнения $F_K(y)=x$. Из определения функции F_K (см. этюд 3.2) очевидны следующие достоинства цифровой подписи:

1) подписать сообщение x, т.е. решить уравнение $F_K(y)=x$, может только абонент — обладатель данного секрета K; другими словами, подделать подпись невозможно;

2) проверить подлинность подписи может любой абонент, знающий открытый ключ, т.е. саму функцию F_K;

3) при возникновении споров отказаться от подписи невозможно в силу ее неподделываемости;

4) подписанные сообщения (x, y) можно, не опасаясь ущерба, пересылать по любым каналам связи.

Именно перечисленные достоинства и обусловили широкое распространение систем цифровой подписи. Опишем, как практически выглядит использование цифровой подписи, на простейшем примере: работа банка с платежными поручениями своих клиентов. Все абоненты этой сети знают одностороннюю функцию F_K, и каждый клиент имеет свой собственный, никому не известный секрет K. Клиент подписывает платежное поручение x с помощью функции F_K со своим секретом K и посылает подписанное платежное поручение в банк. Банк, получив сообщение от клиента и зная открытый ключ, проверяет подлинность подписи клиента и только после этого выполняет его платежное поручение. В силу отмеченных выше достоинств цифровой подписи и банк, и клиент уверены, что их интересы не пострадают.

Широкое развитие систем электронных платежей, электронной почты и других систем передачи данных потребовало большого разнообразия цифровых подписей. Это привело к развитию теории протоколов цифровой подписи, которая в настоящее время составляет большой

раздел теоретической криптографии. В рамках этой теории систематизированы различные виды атак противника на систему цифровой подписи, различные виды успехов, которые противник может достигнуть, различные виды стойкости схем цифровой подписи. Удалось также доказать в некотором смысле эквивалентность существования двух гипотетических объектов: односторонней функции и стойкой схемы цифровой подписи.

Подумайте сами:

1. Пользуясь общей схемой из этюда 3.2, опишите схему цифровой подписи RSA.

3.8. Что такое криптографический протокол

Под *криптографическим протоколом* понимают такую процедуру взаимодействия абонентов, в результате которой абоненты (не противники!) достигают своей цели, а противник — не достигает.

Успехи, достигнутые в разработке схем цифровой подписи и открытого распределения ключей, позволили применить эти идеи также и к другим задачам взаимодействия удаленных абонентов. Так возникло большое новое направление теоретической криптографии — криптографические протоколы. В настоящее время здесь еще нет устоявшихся определений и общепринятой терминологии, однако мы считаем необходимым дать читателю неформальное представление об этой новой интересной области.

Объектом изучения теории криптографических протоколов являются удаленные абоненты, взаимодействующие по открытым каналам связи. Целью взаимодействия абонентов является решение какой-то задачи. Имеется также противник, который преследует собственные цели. При этом противник в разных задачах может иметь разные возможности: например, может взаимодействовать с абонентами от имени других абонентов или вмешиваться в обмены информацией между абонентами и т.д. Противником может даже оказаться один из абонентов или несколько абонентов, вступивших в сговор.

Полезно самостоятельно продумать введенные понятия на примерах изученных ранее протоколов открытого распределения ключей и цифровой подписи.

Приведем еще несколько примеров задач, решаемых удаленными абонентами.

1. Взаимодействуют два не доверяющих друг другу абонента. Они хотят подписать контракт. Это надо сделать так, чтобы не допустить следующую ситуацию: один из абонентов получил подпись другого, а сам не подписался.

Протокол решения этой задачи принято называть *протоколом подписания контракта*.

2. Взаимодействуют два не доверяющих друг другу абонента. Они хотят бросить жребий с помощью монеты. Это надо сделать так, чтобы абонент, подбрасывающий монету, не мог изменить результат подбрасывания после получения догадки от абонента, угадывающего этот результат.

Протокол решения этой задачи принято называть *протоколом подбрасывания монеты*.

Опишем один из простейших протоколов подбрасывания монеты по телефону (так называемая схема Блюма-Микали). Для его реализации у абонентов A и B должна быть односторонняя функция f: $X \to Y$, удовлетворяющая следующим условиям:

1) X — конечное множество целых чисел, которое содержит одинаковое количество четных и нечетных чисел;

2) любые числа $x_1, x_2 \in X$, имеющие один образ $f(x_1)=f(x_2)$, имеют одну четность;

3) по заданному образу $f(x)$ «трудно» вычислить четность неизвестного аргумента x.

Роль подбрасывания монеты играет случайный и равновероятный выбор элемента $x \in X$, а роль орла и решки — четность и нечетность x соответственно. Пусть A — абонент, подбрасывающий монету, а B — абонент, угадывающий результат. Протокол состоит из следующих шагов:

1) A выбирает x («подбрасывает монету»), зашифровывает x, т.е. вычисляет $y=f(x)$, и посылает y абоненту B;

2) B получает y, пытается угадать четность x и посылает свою догадку абоненту A;

3) A получает догадку от B и сообщает B, угадал ли он, посылая ему выбранное число x;

4) B проверяет, не обманывает ли A, вычисляя значение $f(x)$ и сравнивая его с полученным на втором шаге значением y.

3. Взаимодействуют два абонента A и B (типичный при мер: A — клиент банка, B — банк). Абонент A хочет доказать абоненту B, что он именно A, а не противник.

Протокол решения этой задачи принято называть *протоколом идентификации абонента*.

4. Взаимодействуют несколько удаленных абонентов, получивших приказы из одного центра. Часть абонентов, включая центр, могут быть противниками. Необходимо выработать единую стратегию действий, выигрышную для абонентов.

Эту задачу принято называть задачей о византийских генералах, а протокол ее решения — *протоколом византийского соглашения*.

Опишем пример, которому эта задача обязана своим названием. Византия. Ночь перед великой битвой. Византийская армия состоит из n легионов, каждый из которых подчиняется своему генералу. Кроме того, у армии есть главнокомандующий, который руководит генералами. Однако империя находится в упадке и до одной трети генералов, включая главнокомандующего, могут быть предателями. В течение ночи каждый из генералов получает от главнокомандующего приказ о действиях на утро, причем возможны два варианта приказа: «атаковать» или «отступать». Если все честные генералы атакуют, то они побеждают. Если все они отступают, то им удается сохранить армию. Но если часть из них атакует, а часть отступает, то они терпят

поражение. Если главнокомандующий окажется предателем, то он может дать разным генералам разные приказы, поэтому приказы главнокомандующего не стоит выполнять беспрекословно. Если каждый генерал будет действовать независимо от остальных, результаты могут оказаться плачевными. Очевидно, что генералы нуждаются в обмене информацией друг с другом (относительно полученных приказов) с тем, чтобы прийти к соглашению.

Осмысление различных протоколов и методов их построения привело в 1985–1986 гг. к появлению двух плодотворных математических моделей — *интерактивной системы доказательства* и *доказательства с нулевым разглашением*.

Математические исследования этих новых объектов позволили доказать несколько утверждений, весьма полезных при разработке криптографических протоколов.

Под интерактивной системой доказательства (*P*, *V*, *S*) понимают протокол взаимодействия двух абонентов: *P* (доказывающий) и *V* (проверяющий). Абонент *P* хочет доказать *V*, что утверждение *S* истинно. При этом абонент *V* самостоятельно, без помощи *P*, не может доказать утверждение *S* (поэтому *V* и называется проверяющим). Абонент *P* может быть и противником, который хочет доказать *V*, что утверждение *S* истинно, хотя оно ложно. Протокол может состоять из многих *раундов* обмена сообщениями между *P* и *V* и должен удовлетворять двум условиям:

1) *полнота* — если *S* действительно истинно, то абонент *P* почти наверняка убедит абонента *V* признать это;

2) *корректность* — если *S* ложно, то абонент *P* вряд ли убедит абонента *V*, что *S* истинно.

Здесь словами «почти наверняка» и «вряд ли» мы заменили точные математические формулировки, использующие понятие вероятности.

Подчеркнем, что в определении системы (*P*, *V*, *S*) не допускалось, что *V* может быть противником. А если *V* оказался противником, который хочет «выведать» у *P* какую-нибудь новую полезную для себя информацию об утверждении *S*? В этом случае *P*, естественно, может не хотеть, чтобы это случилось в результате работы протокола (*P*, *V*, *S*). Протокол (*P*, *V*, *S*), решающий такую задачу, называется доказательством с нулевым разглашением и должен удовлетворять, кроме условий 1 и 2, еще и следующему условию:

3) *нулевое разглашение* (или стойкость) — в результате работы протокола (*P*, *V*, *S*) абонент *V* не увеличит свои знания об утверждении *S* или, другими словами, не сможет извлечь никакой информации о том, почему *S* истинно.

Самое удивительное, что в 1991 году для широкого класса математических проблем (включающего так называемые *NP-полные задачи*) удалось доказать существование доказательств с нулевым разглашением. Впрочем, это доказано только в предположении, что существует односторонняя функция.

Приведем одно практическое применение теории доказательств с нулевым разглашением — *«интеллектуальные карточки»* (неподделываемые удостоверения личности, кредитные карточки и т.п.). В них вмонтирован микропроцессор, реализующий действия абонента *P* в протоколе, претендующем быть протоколом доказательства с нулевым разглашением (*P*, *V*, *S*). Здесь абонент *P* — владелец карточки, а абонент *V* — например, компьютер в банке или в проходной секретного учреждения. Подумайте, почему в таком случае можно обеспечить неподделываемость удостоверений личности и кредитных карточек.

Заключение

Вы прочли первую книгу по криптографии.

Если вам хочется подробней узнать историю криптографии, события и легенды, связанные с ней, то рекомендуем попытаться найти и прочесть упомянутые в этюде 1.4 книги Д. Кана и Т.А. Соболевой, а также любые номера журнала «Cryptology».

Если вы увлекаетесь программированием и вам захотелось самому реализовать какие-нибудь криптографические алгоритмы, то прежде всего полезно овладеть упомянутой в этюде 2.8

книгой Д. Кнута. Затем можно обратиться к одной из многочисленных книг для программистов по вопросам защиты информации в ЭВМ.

Если вас интересуют математические вопросы криптографии, то в первую очередь необходимо углубиться в те разделы математики, которые упомянуты в этюдах 2.1, 2.2, 2.3, 3.3, 3.4 и 3.8. Систематическое образование в этой области можно получить в любом из вузов, указанных во введении.

Что еще можно почитать о криптографии

1. Т.А. Соболева. Тайнопись в истории России. (История криптографической службы России XVIII — начала XX в.). М., 1994.

2. К. Шеннон. Работы по теории информации и кибернетике. М., ИЛ, 1963.

3. У. Диффи, М.Э. Хеллмэн. Защищенность и имитостойкость. Введение в криптографию. ТИИЭР, том 67, N 3, 1979.

4. Г. Фролов. Тайны тайнописи. М., 1992.

5. М. Гарднер. От мозаик Пенроуза к надежным шифрам. М., Мир, 1993.

6. А.Н. Лебедев. Криптография с «открытым ключом» и возможности ее практического применения. «Защита информации», выпуск 2, 1992.

ПРИЛОЖЕНИЕ Избранные задачи олимпиад по криптографии

Институт криптографии, связи и информатики (ИКСИ) входит в состав Академии Федеральной службы контрразведки Российской Федерации. ИКСИ имеет в своем составе два факультета: информатики и специальной техники. Институт готовит высококвалифицированных специалистов в области защиты информации, криптографии, специальной связи, компьютерной безопасности.

Для школьников при ИКСИ действует вечерняя физико-математическая школа. С 1991 года институт проводит олимпиады по криптографии и математике, избранные задачи которых публикуются в данном приложении.

Задачи

1. Ключом шифра, называемого «решетка», является трафарет, сделанный из квадратного листа клетчатой бумаги размером $n \times n$ (n — четно). Некоторые из клеток вырезаются с тем, чтобы в получившиеся отверстия на чистый лист бумаги того же размера можно было вписывать буквы текста, подлежащего зашифрованию. Одна из сторон трафарета является помеченной. Кроме того, трафарет должен обладать одним важным свойством: при наложении его на чистый лист бумаги четырьмя возможными способами (помеченной стороной вверх, вправо, вниз, влево) его вырезы полностью покрывают всю площадь квадрата, причем каждая клетка оказывается под вырезом ровно один раз.

Буквы сообщения, имеющего длину n^2, последовательно вписываются в вырезы трафарета при каждом из четырех его указанных положений. После снятия трафарета на листе бумаги оказывается зашифрованное сообщение.

Найдите число различных ключей для произвольного четного числа n.

2. В адрес олимпиады пришла шифртелеграмма

ЦДОЗИФКДЦЮ.

Прочитайте зашифрованное сообщение, если известно, что использовался шифр, по которому к двузначному порядковому номеру буквы в алфавите (от 01 до 33) прибавлялось значение многочлена

$f(x) = x^6 + 3x^5 + x^4 + x^3 + 4x^2 + 4x + 5$,

вычисленное либо при $x = x_1$, либо при $x = x_2$ (в случайном порядке), где x_1, x_2 — корни трехчлена $x^2 + 3x + 1$, а затем полученное число заменялось соответствующей ему буквой.

3. Одна фирма предложила устройство для автоматической проверки пароля. Паролем может быть любой непустой упорядоченный набор букв в алфавите $\{a, b, c\}$. Будем обозначать такие наборы большими латинскими буквами. Устройство перерабатывает введенный в него набор P в набор $Q = \varphi(P)$. Отображение φ держится в секрете, однако про него известно, что оно определено не для каждого набора букв P и обладает следующими свойствами. Для любого набора букв P

1) $\varphi(aP) = P$;

2) $\varphi(bP) = \varphi(P)a\varphi(P)$;

3) набор $\varphi(cP)$ получается из набора $\varphi(P)$ выписыванием букв в обратном порядке.

Устройство признает предъявленный пароль верным, если $\varphi(P) = P$. Например, трехбуквенный набор bab является паролем, так как $\varphi(bab) = \varphi(ab)a\varphi(ab) = bab$. Подберите пароль, состоящий более, чем из трех букв.

4. Коммерсант для передачи цифровой информации с целью контроля передачи разбивает строчку передаваемых цифр на пятерки и после каждых двух пятерок приписывает две последние цифры от суммы чисел, изображенных этими пятерками. Затем процесс шифрования осуществляется путем прибавления к шифруемым цифрам членов арифметической прогрессии с последующей заменой сумм цифр остатками от деления на 10. Прочитайте зашифрованное сообщение:

4 2 3 4 6 1 4 0 5 3 1 3.

5. Рассмотрим модель шифра для цифрового текста, в котором каждая цифра заменяется остатком от деления значения многочлена

$f(x) = b(x^3 + 7x^7 + 3x + a)$

на число 10, где a,b — фиксированные натуральные числа. Выяснить, при каких значениях a и b возможно однозначное расшифрование.

6. Фирма предложила на рынок кодовый замок. При установке владелец замка сопоставляет каждой из 26 латинских букв, расположенных на клавиатуре, произвольное натуральное число (известное лишь обладателю замка). После выбора произвольной комбинации попарно различных букв, происходит суммирование числовых значений набранных букв и замок открывается, если сумма делится на 26. Докажите, что для любых числовых значений букв существует комбинация, открывающая замок.

7. Рассматривается шифр, в котором буквы русского 30-буквенного алфавита Ω занумерованы по следующей таблице:

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

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

Для зашифрования сообщения $\tau = t_1t_2...t_n$ выбирается некоторая последовательность $\kappa = \gamma_1\gamma_2...\gamma_n$ (ключ), состоящая из букв алфавита Ω. Зашифрование состоит в позначном сложении соответствующих букв из τ и κ с последующей заменой суммы буквой алфавита Ω, номер которой равен остатку от деления этой суммы на число 30.

Известно, что два сообщения τ_1 и τ_2 зашифрованы с помощью одного ключа (κ) и что каждое из них содержит слово «корабли». Восстановить τ_1 и τ_2 по текстам данных криптограмм:

σ_1=ЮПТЦАРГШАЛЖЖЕВЦЩЫРВУУ

σ_2=ЮПЯТБНЩМСДТЛЖГПСГХСЦЦ

8. Перехвачена «шифровка»: РБЬНПТСИТСРРЕЗОХ

Относительно шифра известно следующее:

— используется шифр предыдущей задачи;

— в качестве ключа используется произвольная последовательность, составленная из букв: А,Б,В.

Прочтите зашифрованное сообщение.

9. Шифр простой замены в алфавите $A = \{a_1, a_2, ..., a_n\}$, состоящем из n различных букв, заключается в замене каждой буквы шифруемого текста буквой того же алфавита, причем разные буквы заменяются разными. Ключом шифра простой замены называется таблица, в которой указано, какой буквой надо заменить каждую букву алфавита A. Если слово СРОЧНО зашифровать простой заменой с помощью ключа:

АБВГДЕЖЗИКЛМНОПРСТУФХЦЧШЩЪЫЭЮЯ

ЧЯЮЭЫЫЦШЦХФУБДТЗВРПМЛКАИОЖЕСГН,

то получится слово ВЗДАБД. Зашифровав полученное слово с помощью того же ключа еще раз, получим новое слово ЮШЫЧЯЫ. Сколько всего различных слов можно получить, если указанный процесс шифрования продолжить неограниченно?

10. Сообщение, зашифрованное в пункте **А** шифром простой замены в алфавите из букв русского языка и знака пробела (_) между словами, передается в пункт **Б** отрезками по 12 символов. При передаче очередного отрезка сначала передаются все его знаки, стоящие на четных местах в порядке возрастания их номеров, начиная со второго, а затем — все знаки, стоящие на нечетных местах, также в порядке возрастания их номеров, начиная с первого. В пункте **Б**полученное шифрованное сообщение дополнительно шифруется с помощью некоторого другого шифра простой замены в том же алфавите, а затем таким же образом, как и из пункта **А**, передается в пункт **В**. По перехваченным в пункте **В**отрезкам:

СО_ГЖТПНБЛЖО

РСТКДКСПХЕУБ

_Е_ПФПУБ_ЮОБ

СП_ЕОКЖУУЛЖЛ

СМЦХБЭКГОЩПЫ

УЛКЛ_ИКНТЛЖГ,

восстановите исходное сообщение зная, что в одном из передаваемых отрезков зашифровано слово КРИПТОГРАФИЯ.

11. Дана последовательность $C_1, C_2, C_3, \ldots, C_n, \ldots$, в которой C_n есть последняя цифра числа n^n. Доказать, что эта последовательность периодическая и ее период равен 20.

12. Знаки алфавита, состоящего из букв русского языка и символа пробела между словами (_), заменим парами цифр согласно таблице:

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

01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

Для зашифрования сообщения длины m, записанного в этом алфавите, сначала преобразуем буквенный текст в цифровой $T = t_1, t_2, \ldots, t_{2m}$,а затем, выбрав отрезок $K = C_{n+1}, C_{n+2}, \ldots, C_{n+2m}$ последовательности из задачи 11, осуществим последовательное поразрядное сложение цифр текста T с цифрами отрезка K, причем в качестве очередного знака шифрованного текста берется цифра единиц соответствующей суммы (младший разряд).

Прочитайте зашифрованное сообщение:

2 3 3 9 8 6 7 2 1 6 4 5 8 1 6 0 6 7 0 6 1 7 3 1 5 5 8 8.

* * *

1

David Kahn, Codebreakers. The story of Secret Writing. New-York, Macmillan, 1967.
(обратно)

2

Т.А. Соболева. Тайнопись в истории России. (История криптографической службы России XVIII — начала XX в.). М., 1994.
(обратно)

3

У. Диффи, М. Э. Хеллмэн. Защищенность и имитостойкость. Введение в криптографию. ТИИЭР, т. 67, N 3, 1979.
(обратно)

Оглавление