Сергей Дориченко

Стойкие шифры

Шифры замены

Обычные шифры из детективных романов часто устроены так: каждая буква сообщения заменяется каким-нибудь определённым значком или другой буквой. Подобные шифры очень ненадёжны, и вот почему. Буквы в текстах на русском языке (да и на любом языке вообще) встречаются неравномерно. Например, буква «О» в русских текстах встречается чаще всех других букв, а буква «Ъ» – реже всего. У каждой буквы есть своя примерная частота появления в тексте (смотри таблицу на поле справа).

Сочетания букв тоже встречаются неравномерно (например, «ьь» вообще не встречается). Конечно, все эти частоты зависят от конкретного текста – скажем, в биологической статье о жужелицах буква «ж» явно будет встречаться чаще, чем обычно. Но приведённая таблица вполне годится как ориентир.

Так вот, описанный способ шифровки не изменяет частот – просто теперь с аналогичной частотой будет появляться не сама буква, а заменяющий её значок. Высчитав частоту появления каждого значка в шифровке и сравнив полученные данные с таблицей частот, мы можем сделать предположения, какой букве какой значок соответствует. Далее пробуем заменять значки один за одним на буквы, проверяя свои догадки, корректируя их и делая новые, и постепенно расшифровываем текст. Если он не слишком короткий, мы с большой вероятностью его полностью расшифруем (хотя это может оказаться не совсем простым делом). Кстати, намного чаще любой буквы встречается пробел, разделяющий слова. Поэтому если пробел используется в шифровке и тоже заменён на какой-то значок, мы разгадаем его в первую очередь.

Совершенный шифр

Опишем теперь шифр, который принципиально не поддается расшифровке без знания ключа. Сопоставим каждой букве русского алфавита свою последовательность из 0 и 1 длины 5 (пятизначный двоичный код), например: А – 00000, Б – 00001, В – 00010 и так далее (или в каком-то другом порядке). Если буквы Е и Ё кодировать одинаково, то последовательностей как раз хватит (их 32, а в алфавите 33 буквы).

Заменим в тексте каждую букву на её двоичный код, получим последовательность из 0 и 1 (двоичный текст). Это пока ещё не шифровка – мы бы легко разгадали, какая буква на какую последовательность заменена (тем же методом, что и в случае замены букв на значки).

Чтобы зашифровать полученный двоичный текст, нам потребуется ещё ключ – случайная последовательность из 0 и 1 такой же длины. Этот ключ должен быть и у отправителя шифрованного сообщения, и у адресата.

Для зашифровки просто складываем две последовательности нулей и единиц – двоичный текст сообщения и ключ: первую цифру с первой, вторую со второй, и так далее. Но складываем по особым правилам:

$$0+0=0,\ 1+0=1,\ 0+1=1,\ 1+1=0$$

(в математике это называется сложением по модулю 2). Полученная последовательность и будет зашифрованным сообщением. Чтобы расшифровать её, надо просто… снова прибавить к ней ключ! Тогда мы как бы прибавим к исходной последовательности ключ два раза. А по нашим правилам, прибавляя две одинаковые цифры мы ничего не меняем, то есть мы вернёмся к исходному двоичному тексту. Схематически процесс шифрования и дешифрования можно описать так:

$$\text{текст} + \text{ключ} = \text{шифровка};$$
$$\text{шифровка} + \text{ключ} = \text{текст} + \text{ключ} + \text{ключ} = \text{текст}.$$

Ясно, что расшифровать сообщение, не зная ключа, невозможно. Нам как бы дана сумма двух чисел, и нельзя восстановить одно из слагаемых, ничего не зная про другое. Имея на руках лишь шифровку, мы знаем только, что исходный текст может быть абсолютно любым текстом соответствующей длины. Ведь по любому такому тексту можно изготовить ключ, который приведёт ровно к той же самой шифровке!

Недостаток описанного способа в том, что каждый текст требует нового ключа такой же длины – если повторять ключи, появляется возможность расшифровки. Например, мы могли бы попробовать вместо длинного ключа использовать ключ всего из пяти символов, скажем 11010. Разбиваем двоичный текст на пятёрки цифр и прибавляем к каждой пятёрке 11010. Фактически, мы просто заменяем каждую пятёрку цифр на какую-то другую фиксированную пятёрку. В этом случае расшифровать исходный текст так же легко, как если бы

мы просто заменили его двоичным кодом, не прибавляя никакого ключа. Использовать длинные ключи, но всё же существенно более короткие, чем текст, тоже опасно – есть метод определения длины ключа, а после того как длина ключа установлена, можно применить частотный анализ.

Поэтому надо заготовить ключ огромной длины заранее и лишь указывать, например, в начале шифровки, какое место ключа используется. При этом очень важно, чтобы ключ был случайной последовательностью из 0 и 1. Например, последовательности 1111111111111111 и 010101010101010 не случайные. Кстати, придумать случайную последовательность не так-то просто. Трудно даже (но возможно) дать чёткое определение, какие последовательности могут считаться случайными.

Немного истории и литературы

Подобный шифр использовал Макс Кристиансен-Клаузен, шифровальщик выдающегося советского разведчика Рихарда Зорге. Наиболее часто употребляемые буквы английского алфавита s, i, o, e, r, a, t, n заменялись цифрами от 0 до 7, а остальные буквы – числами от 80 до 99 (чтобы не возникало путаницы, когда числа записывались подряд). Ключом служили старые выпуски «Статистического ежегодника Германского рейха» с множеством числовых данных. Ключ записывали под текстом и прибавляли, причём если сумма двух цифр превышала 10, то записывалась только её последняя цифра. Например, вместо 7 + 5 писали 2, отбрасывая десяток (в математике это называется сложением по модулю 10). Восстанавливали исходное сообщение, «вычитая» ключ из шифровки. Когда выходило отрицательное число, как скажем при вычитании 5 из 2, было ясно, что надо вычитать из числа на 10 больше, то есть из 12 – вот и получали 7.

Японские тайные службы перехватили много радиограмм Зорге, но ни одной не сумели расшифровать. Более полный рассказ об этом читайте в замечательной книге Юлиуса Мадера «Репортаж о докторе Зорге».

А герой приключенческих романов Юлиана Семёнова «Семнадцать мгновений весны» и «Приказано выжить» разведчик Штирлиц, больше известный нам по знаменитому кинофильму, использовал в качестве ключа художественную книгу Монтеня. При этом осмысленный текст сообщения«складывался» с осмысленным

же (и значит, не случайным!) текстом ключа. Когда германским контрразведчикам стало известно предполагаемое содержание одной из шифровок, в частности – некоторые слова, которые там могли встречаться, – они попробовали их подставить в разные места шифровки и посмотреть, какой получается ключ. Попав в нужное место, они открывали кусочек ключа, в котором угадывались части осмысленных слов. Восстанавливая эти слова, они раскрывали и новый кусочек шифровки, и так постепенно расшифровали её.

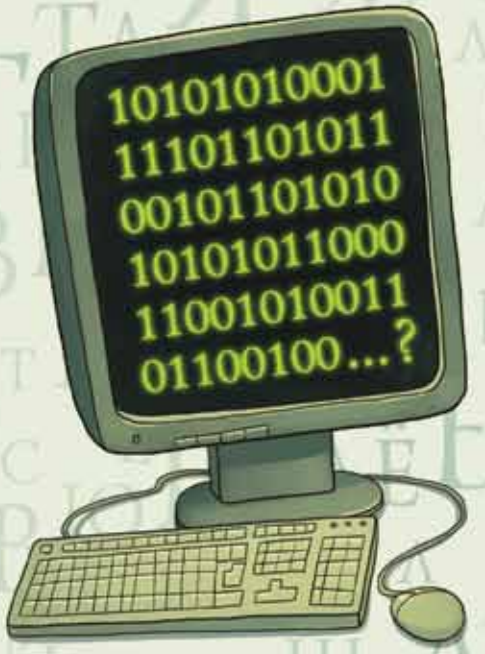
Шифры с открытым ключом

Начиная с 1977 года, стали появляться новые шифры, основанные на глубоких математических идеях, высказанных американскими математиками Диффи и Хеллманом за два года до этого. Представьте себе, что два бизнесмена хотят переписываться друг с другом, надёжно шифруя сообщения, но забыли договориться о ключе. Они находятся в разных странах, всё их общение может прослушиваться конкурентами. Как тут быть? Оказывается, выходы есть. Опишем один из них, но без подробностей, только сам принцип.

Придуман способ шифровки, для которого надо знать лишь произведение pq двух каких-то простых чисел p и q, а сами числа p и q знать не нужно. А вот для расшифровки сообщения обязательно иметь в распоряжении и число p, и число q. «Ну и что тут такого?», – спросите вы. А вот что. Дело в том, что эти простые числа можно взять очень большими. И тут мы сталкиваемся с таким явлением: современные компьютерные мощности огромны, но всё же ограничены. Скажем, компьютер может за разумное время разложить на простые множители 200-значное число, но раскладывание 300-значных чисел ему уже не под силу (любому из известных алгоритмов потребуются многие годы). Всегда есть какая-то подобная граница. А выяснить про число, простое оно или нет, компьютеры могут очень быстро для гораздо более длинных чисел. Так вот, первый из компаньонов может с помощью компьютера найти какие-нибудь два, скажем, 400-значных простых числа p и q, перемножить их и открыто переслать результат второму (а сами числа p и q хранить в тайне). Получив произведение pq, тот зашифрует своё сообщение и отправит обратно первому. И первый его легко расшифрует – он-то знает оба числа p и q. А вот всяким там подслушивателям для расшифровки

придётся сначала разложить на множители произведение pq, в котором 800 знаков – а с этим не справится ни один современный компьютер! Этот метод шифровки называется RSA, по первым буквам фамилий его создателей – Ривеста, Шамира и Адлемана.

Конечно, с развитием компьютерных технологий появляется возможность расшифровывать старые сообщения. Первая шифровка авторов RSA, опубликованная ими в 1977 году как вызов всем дешифровальщикам мира, продержалась 17 лет. Также есть опасность, что будет найден новый, быстрый алгоритм разложения чисел на простые множители. Но есть математическая гипотеза, что все такие алгоритмы работают принципиально не быстрее, чем уже известные.



А у вас получится?

Перед вами текст, который получен из хорошо известного заменой каждой буквы на какую-то другую. Расшифруйте его.

Атокг ацынг цлекытуы цлауенг ьи Чолсв, и уими Чолси уманлоти ки эекпв ниеме вматыфюеме, цатказме утоь чтиьиме, жна ни ацынг ьималчити, андоти атокы д уналакв е, мокыы омв ки чатадо тхс, боцквти:

– Пиё д уимам сото в Укоркаё палатодз, ка ак дцатко садаток е свмион, жна твжбо омв кечсо е щзнг ко марон. Цлежекаё ро дуомв аупатпе ьолпити, жна уесын в коча д уолсйо е д чтиьв. Еш киса вситенг, екижо ак кепачси ко щвсон жотадопам, е Укоркиы палатоди уашликен кис кем удаф дтиунг.

– Ка ко цамаробг те нз Чолсо пип-кещвсг вкежнаренг янв дтиунг?

– Уетгкоо, жом аки оунг, ы ко мачв ох усотинг. Ко десебг лиьдо, пип дотепи ох уети? Ко десебг, жна оё утврин е тфсе е реданкзо? Досг аки щауиы ащабти цатудони! Ко в киу ьикеминг оё уетв! Уети – д ох метам, кодеккам сонупам уолсожпо. Оуте аки уими ко умарон цлакепквнг д жолначе Укоркаё палатодз е еьдтожг еь уолсйи Пиы аупатпе, на мз е цасидка оё ко цамаром! Д сдвш метыш ануфси кижекионуы уис Укоркаё палатодз. Анкоуе нвси содажпв, уцвуне в щатгбача пвуни, цаплзнача плиукзме ычасиме, е, ко мобпиы, даьдлиюиёуы ащлинка!

У янеме утадиме эекпи цасуисети Чолсв ки уцекв атокы, е нан щлауетуы щоринг уа дуош кач.

Художник Виктор Пяткин