Лекция 01. Введение в криптографический анализ методов открытой криптографии

Цель лекции: приступить к изучению основ криптографического анализа методов открытой криптографии. Лекция является обязательной для понимания следующих тем курса.

План лекции:
Введение
1 Классическая система секретной связи
Заключение
Контрольные вопросы

Ключевые слова: [выбрать самостоятельно].

Содержание лекции:

Введение

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

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

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

1 Классическая система секретной связи

В криптографии обычно предполагается, что у лица, передающего сообщения и (или) их принимающего, есть некоторый противник E, который может быть конкурентом в бизнесе, членом преступной группировки, представителем иностранной разведки или даже чрезмерно ревнивой женой, и

этот противник может перехватывать сообщения, передаваемые по открытому каналу, и анализировать их. Часто удобно рассматривать противника как некую особу по имени Ева, которая имеет в своем распоряжении мощную вычислительную технику и владеет методами криптоанализа. Естественно, Алиса и Боб хотят, чтобы их сообщения были непонятны Еве, и используют для этого специальные шифры.

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

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

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

$$c = (m + k) \bmod 32, \qquad (1.1)$$

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

Чтобы дешифровать зашифрованный текст, нужно применить «обратный» алгоритм

$$m = (c - k) \bmod 32. \qquad (1.2)$$

Можно представить себе ситуацию, когда источник и получатель сообщений договорились использовать шифр (1.1), но для того, чтобы усложнить задачу противника, решили иногда менять ключ шифра. Для этого Алиса каким-либо образом генерирует число k, передает его Бобу по закрытому каналу связи, и после этого они обмениваются сообщениями, зашифрованными с помощью этого ключа k. Замену ключа можно проводить, например, перед каждым сеансом связи или после передачи фиксированного числа букв (скажем, каждую десятку символов шифровать со своим k) и т.п. В таком случае говорят, что ключ порождается источником ключа. Схема рассмотренной криптосистемы с секретным ключом приведена на рис. 1.

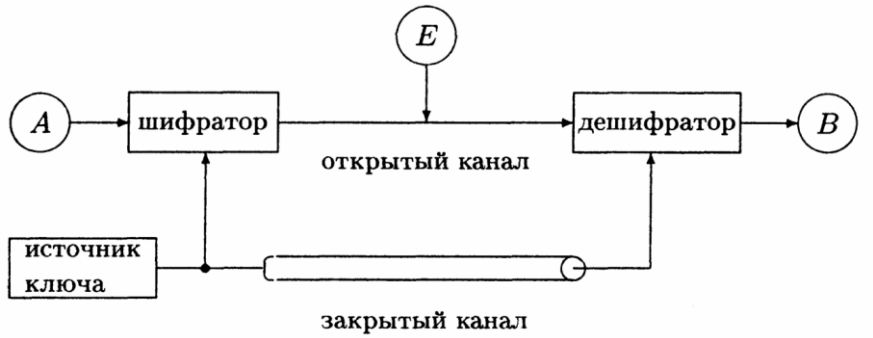

Рис. 1. Классическая система секретной связи

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

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

Наиболее очевидная попытка расшифровки – последовательный перебор всех возможных ключей (это так называемый метод «грубой силы»). Итак, Ева перебирает последовательно все возможные ключи k = 1, 2, ..., подставляя их в алгоритм дешифрования и оценивая получающиеся результаты. Попробуем и мы использовать этот метод. Результаты

дешифрования по (1.2) при разных ключах и шифротексте «ТИУИПИРГ» сведены в табл. 1.1. В большинстве случаев нам достаточно было расшифровать две-три буквы, чтобы отвергнуть соответствующий ключ (из-за отсутствия слова в русском языке, начинающегося с такого фрагмента).

Таблица 1.1: Расшифровка слова «ТИУИПИРГ» путем перебора ключей

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

Из табл. 1.1 мы видим, что был использован ключ k = 3 и зашифровано сообщение ПЕРЕМЕНА. Причем для того, чтобы проверить остальные возможные значения ключа, нам не требовалось дешифровать все восемь букв, а в большинстве случаев после анализа двух-трех букв ключ отвергался (только при k = 8 надо было дешифровать пять букв, зато при k = 22, 23, 24 хватало и одной, так как в русском языке нет слов, начинающихся с Ь, Ъ, Ы).

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

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

$$c = (m + k) \bmod 10. \qquad (1.3)$$

Допустим, она послала Бобу шифротекст «26047». Ева пытается расшифровать его, последовательно перебирая все возможные ключи. Результаты ее попыток сведены в табл. 1.2.

Таблица 1.2: Расшифровка сообщения «26047» путем перебора ключей

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

5	71592	0	26047

Мы видим, что все полученные варианты равнозначны и Ева не может понять, какая именно комбинация истинна. Анализируя шифротекст, она не может найти значения секретного ключа. Конечно, до перехвата сообщения у Евы было 10^5 возможных значений кодовой комбинации, а после – только 10. Однако важно отметить то, что в данном случае всего 10 значений ключа. Поэтому при таком ключе (одна десятичная цифра) Алиса и Боб и не могли рассчитывать на большую секретность.

В первом примере сообщение – текст на русском языке, поэтому оно подчиняется многочисленным правилам, различные буквы и их сочетания имеют различные вероятности и, в частности, многие наборы букв запрещены. (Это свойство называется избыточностью текста). Поэтому-то и удалось легко найти ключ и дешифровать сообщение, т.е. избыточность позволила «взломать» шифр. В противоположность этому, во втором примере все комбинации цифр допустимы. «Язык» кодового замка не содержит избыточности. Поэтому даже простой шифр, примененный к сообщениям этого языка, становится невскрываемым. В классической работе К. Шеннона [1] построена глубокая и изящная теория шифров с секретным ключом и, в частности, предложена “правильная” количественная мера избыточности. Мы кратко коснемся этих вопросов, когда нами будут описываться современные шифры с секретным ключом.

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

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

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

[1] Шеннон К. Работы по теории информации и кибернетике. – М.: ИЛ, 1963. – С. 333-369 (Теория связи в секретных системах).

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

Заключение

Итак, мы познакомились с основными героями криптографии – Алисой, Бобом и Евой и с важными понятиями этой науки – шифром, ключом, атакой, открытым и защищенным каналом. Заметим, что с последним понятием связан один интригующий факт – возможно построение надежных криптосистем без защищенного канала! В таких системах Алиса и Боб вычисляют секретный ключ так, что Ева не может этого сделать. Это открытие было сделано в основополагающих работах Диффи, Хеллмана и Меркля (см., например, [2]) в 1976 году и открыло новую эру в современной криптографии. Большая часть нашего курса будет связана именно с такими системами, называемыми схемами с открытым, или несимметричным ключом.

Контрольные вопросы

Смотри руководство по организации самостоятельной работы.

[2] 20. Diffie W., Heilman M. E. New directions in cryptography // IEEE Transactions on Information Theory. 1976. V. 22. P. 644-654.