КРИПТОАНАЛИЗ КЛАССИЧЕСКИХ ШИФРОВ: ЛАБОРАТОРНЫЙ ПРАКТИКУМ

Автор/создатель: Жданов О.Н., Куденкова И.А.

Курс "Криптографические методы защиты информации" является базовым при подготовке специалистов по защите информации. В настоящем лабораторном практикуме к курсу на примерах классических шифров иллюстрируются некоторые важные приемы и методы криптоанализа. Приведенные в пособии теоретические сведения дополнены подробно изложенными примерами выполнения заданий.

Страницы ← предыдущая следующая →
1 2 3 4 5 6 7 8 9 10

Федеральное агентство по образованию

Сибирский государственный аэрокосмический университет

имени академика М. Ф. Решетнева

О. Н. ЖДАНОВ,

И. А. КУДЕНКОВА

КРИПТОАНАЛИЗ КЛАССИЧЕСКИХ ШИФРОВ

Лабораторный практикум для студентов,

обучающихся по специальностям

«Комплексное обеспечение информационной безопасноти автоматизированных

систем» и

«Информационная безопасность телекоммуникационных систем»

Красноярск 2008

Оглавление

ВВЕДЕНИЕ

Курс «Криптографические методы защиты информации» является базовым при подготовке специалистов по защите информации. На основе знаний криптографии выстраивается система подготовки специалистов. При этом изучение методов защиты неразрывно связано с изучением возможных атак на алгоритмы и на их реализации. Хорошо известно, что для усвоения материала необходима активная самостоятельная работа студентов. Поэтому представляется

целесообразным проведение лабораторных работ по криптоанализу. Работы по анализу таких шифров, как DES, ГОСТ 28147-89, IDEA требуют большого ресурса и для начинающего являются чрезвычайно сложными. В то же время на примерах классических шифров можно проиллюстрировать некоторые важные приемы и методы криптоанализа. Как показывает практика работы, студенты после анализа шифров перестановки, простой замены и Виженера уверенно и достаточно быстро входят в круг идей современной криптографии. Таким образом, настоящее пособие выполняет пропедевтическую функцию. После анализа классических шифров учащиеся успешно изучают современные блочные алгоритмы шифрования, им становятся доступными идеи линейного и дифференциального криптоанализа.

Авторы сочли необходимым теоретические сведения дополнить подробно изложенными примерами выполнения заданий. После изучения теории и ознакомления с образцами решений заданий студент должен выполнить свой вариант лабораторной работы. Мы не приводим ответы к задачам, дабы не лишать обучающихся удовольствия от самостоятельного решения. Заинтересовавшиеся коллеги могут получить ответы по адресу: onzhdanov@mail.ru.

КЛАССИЧЕСКИЕ ШИФРЫ

Разработкой методов преобразования (шифрования) информации с целью ее защиты от незаконных пользователей занимается криптография. Такие методы и способы преобразования информации называются шифрами.

Шифрование (зашифрование) — процесс применения шифра к защищаемой информации, т. е. преобразование защищаемой информации (открытого текста) в шифрованное сообщение (шифртекст, криптограмму) 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.

Криптография — прикладная наука, она использует самые последние достижения фундаментальных наук и, в первую очередь, математики. С другой стороны, все конкретные задачи криптографии существенно зависят от уровня развития техники и технологии, от применяемых средств связи и способов передачи информации.

Современная криптография является областью знаний, связанной с решением таких проблем безопасности информации, как конфиденциальность, целостность, аутентификация и невозможность отказа сторон от авторства. Достижение этих требований безопасности информационного взаимодействия и составляет основные цели криптографии. Они определяются следующим образом.

Обеспечение конфиденциальности — решение проблемы защиты информации от ознакомления с ее содержанием со стороны лиц, не имеющих права доступа к ней. В зависимости от контекста вместо термина "конфиденциальная" информация могут выступать термины "секретная", "частная", "ограниченного доступа" информация.

Обеспечение целостности — гарантирование невозможности несанкционированного изменения информации. Для гарантии целостности необходим простой и надежный критерий обнаружения любых манипуляций с данными. Манипуляции с данными включают вставку, удаление и замену.

Обеспечение аутентификации — разработка методов подтверждения подлинности сторон (идентификация) и самой информации в процессе информационного взаимодействия. Информация, передаваемая по каналу связи, должна быть аутентифицирована по источнику, времени создания, содержанию данных, времени пересылки и т. д.

Обеспечение невозможности отказа от авторства — предотвращение возможности отказа субъектов от некоторых из совершенных ими действий. Рассмотрим средства для достижения этих целей более подробно.

Традиционной задачей криптографии является проблема обеспечения конфиденциальности информации при передаче сообщений по контролируемому противником каналу связи. В простейшем случае эта задача описывается взаимодействием трех субъектов (сторон). Владелец информации, называемый обычно отправителем, осуществляет преобразование исходной (открытой)

информации (сам процесс преобразования называется шифрованием) в форму передаваемых получателюпо открытому каналу связи шифрованных сообщений с целью ее защиты от противника.

Под противником понимается любой субъект, не имеющий права ознакомления с содержанием передаваемой информации. В качестве противника может выступать криптоаналитик, владеющий методами раскрытия шифров. Законный получатель информации осуществляет расшифрование полученных сообщений. Противник пытается овладеть защищаемой информацией (его действия обычно называют атаками). При этом он может совершать как пассивные, так и активные действия. Пассивные атаки связаны с прослушиванием, анализом трафика, перехватом, записью передаваемых шифрованных сообщений, дешифрованием, то есть попытками "взломать" защиту с целью овладения информацией.

При проведении активных атак противник может прерывать процесс передачи сообщений, создавать поддельные (сфабрикованные) или модифицировать передаваемые шифрованные сообщения. Эти активные действия называют по-пытками имитации и подмены соответственно.

Под шифром обычно понимается семейство обратимых преобразований, каждое из которых определяется некоторым параметром, называемым ключом, а также порядком применения данного преобразования, называемым режимом шиф-рования.

Ключ — это важнейший компонент шифра, отвечающий за выбор преобразования, применяемого для зашифрования конкретного сообщения. Обычно ключ представляет собой некоторую буквенную или числовую последовательность. Эта последовательность как бы "настраивает" алгоритм шифрования.

Каждое преобразование однозначно определяется ключом и описывается некоторым криптографическим алгоритмом. Один и тот же криптографический алгоритм может применяться для шифрования в различных режимах. Тем самым реализуются различные способы шифрования (простая замена, гаммирование и т. п.). Каждый режим шифрования имеет как свои преимущества, так и недостатки.

Поэтому выбор режима зависит от конкретной ситуации. При расшифровании используется криптографический алгоритм, который в общем случае может отличаться от алгоритма, применяемого для зашифрования сообщения. Соответственно могут различаться ключи зашифрования и расшифрования. Пару алгоритмов зашифрования и расшифрования обычно называют криптосистемой (шифрсистемой), а реализующие их устройства — шифртехникой.

Если обозначить через M открытое, а через C шифрованное сообщения, то процессы зашифрования и расшифрования можно записать в виде равенств

Ek1(M)=C

Dk2(C)=M

в которых алгоритмы зашифрования E и расшифрования D должны удовлетворять равенству

Dk2(Ek1(M))=M

Наряду с конфиденциальностью не менее важной задачей является обеспечение целостности информации, другими словами, — неизменности ее в

процессе передачи или хранении. Решение этой задачи предполагает разработку средств, позволяющих обнаруживать не столько случайные искажения (для этой цели вполне подходят методы теории кодирования с обнаружением и исправлением ошибок), сколько целенаправленное навязывание противником ложной информации. Для этого в передаваемую информацию вносится избыточность. Как правило, это достигается добавлением к сообщению некоторой проверочной комбинации, вычисляемой с помощью специального алгоритма и играющей роль контрольной суммы для проверки целостности полученного сообщения. Главное отличие такого метода от методов теории кодирования состоит в том, что алгоритм выработки проверочной комбинации является "криптографическим", то есть зависящим от секретного ключа. Без знания секретного ключа вероятность успешного навязывания противником искаженной или ложной информации мала. Такая вероятность служит мерой имитостойкости шифра, то есть способности самого шифра противостоять активным атакам со стороны противника.

Итак, для проверки целостности к сообщению M добавляется проверочная комбинация S, называемая кодом аутентификации сообщения (сокращенно — КАС) или имитовставкой. В этом случае по каналу связи передается пара C = (M, S). При получении сообщения M пользователь вычисляет значение проверочной комбинации и сравнивает его с полученным контрольным значением S Несовпадение говорит о том, что данные были изменены.

Как правило, код аутентификации является значением некоторой (зависящей от секретного ключа) криптографической хеш-функции от данного сообщения: hk(M) = S К кодам аутентификации предъявляются определенные требования. К ним относятся:

— невозможность вычисления значения hk(M) = S для заданного сообщения M без знания ключа k,

— невозможность подбора для заданного сообщения M с известным значением hk(M)=S другого сообщения M1 с известным значением hk (M1) = S1, без знания ключа k.

Первое требование направлено против создания поддельных (сфабрикованных) сообщений при атаках типа имитация; второе — против

модификации передаваемых сообщений при атаках типа подмена.

Аутентификация — установление подлинности. В общем случае этот термин может относиться ко всем аспектам информационного взаимодействия: сеансу связи, сторонам, передаваемым сообщениям и т. д.

Установление подлинности (то есть проверка и подтверждение) всех аспектов информационного взаимодействия является важной составной частью проблемы обеспечения достоверности получаемой информации. Особенно остро эта проблема стоит в случае не доверяющих друг другу сторон, когда источником угроз может служить не только третья сторона (противник), но и сторона, с которой осуществляется взаимодействие.

Применительно к сеансу связи аутентификация означает проверку: целостности соединения, невозможности повторной передачи данных противником и своевременности передачи данных. Для этого, как правило, используют

дополнительные параметры, позволяющие "сцепить" передаваемые данные в легко проверяемую последовательность. Это достигается, например, путем вставки в сообщения некоторых специальных чисел или меток времени. Они позволяют предотвратить попытки повторной передачи, изменения порядка следования или обратной отсылки части переданных сообщений. При этом такие вставки в передаваемом сообщении необходимо защищать (например, с помощью шифрования) от возможных подделок и искажений.

Применительно к сторонам взаимодействия аутентификация означает проверку одной из сторон того, что взаимодействующая с ней сторона — именно та, за которую она себя выдает. Часто аутентификацию сторон называют также идентификацией.

Основным средством для проведения идентификации являются протоколы идентификации, позволяющие осуществлять идентификацию (и аутентификацию) каждой из участвующих во взаимодействии и не доверяющих друг другу сторон. Различают протоколы односторонней и взаимной идентификации.

Протокол — это распределенный алгоритм, определяющий последовательность действий каждой из сторон. В процессе выполнения протокола идентификации каждая из сторон не передает никакой информации о своем секретном ключе, а хранит его у себя и использует для формирования ответных сообщений на запросы, поступающие при выполнении протокола.

Наконец, применительно к самой информации аутентификация означает проверку того, что информация, передаваемая по каналу, является подлинной по содержанию, источнику, времени создания, времени пересылки и т. д.

Проверка подлинности содержания информации сводится, по сути, к проверке ее неизменности (с момента создания) в процессе передачи или хранения, то есть проверке целостности.

Аутентификация источника данных означает подтверждение того, что исходный документ был создан именно заявленным источником.

Заметим, что если стороны доверяют друг другу и обладают общим секретным ключом, то аутентификацию сторон можно обеспечить применением кода аутентификации. Действительно, каждое успешно декодированное получателем сообщение может быть создано только отправителем, так как только он знает их общий секретный ключ. Для не доверяющих друг другу сторон решение

подобных задач с использованием общего секретного ключа становится невозможным. Поэтому при аутентификации источника данных нужен механизм цифровой подписи, который будет рассмотрен ниже.

В целом, аутентификация источника данных выполняет ту же роль, что и протокол идентификации. Отличие заключается только в том, что в первом случае имеется некоторая передаваемая информация, авторство которой требуется установить, а во втором требуется просто установить сторону, с шторой осуществляется взаимодействие.

Математические модели открытого текста

Потребность в математических моделях открытого текста продиктована, прежде всего, следующими соображениями. Во-первых, даже при отсутствии ограничений на временные и материальные затраты по выявлению закономерностей, имеющих место в открытых текстах, нельзя гарантировать того, что такие свойства указаны с достаточной полнотой. Например, хорошо известно, что частотные свойства текстов в значительной степени зависят от их характера. Поэтому при математических исследованиях свойств шифров прибегают к упрощающему моделированию, в частности, реальный открытый текст заменяется его моделью, отражающей наиболее важные его свойства. Во-вторых, при автоматизации методов криптоанализа, связанных с перебором ключей, требуется "научить" ЭВМ отличать открытый текст от случайной последовательности знаков. Ясно, что соответствующий критерий может выявить лишь адекватность последовательности знаков некоторой модели открытого текста.

Один из естественных подходов к моделированию открытых текстов связан с учетом их частотных характеристик, приближения для которых можно вычислить с нужной точностью, исследуя тексты достаточной длины. Основанием для такого подхода является устойчивость частот к -грамм или целых словоформ реальных языков человеческого общения (то есть отдельных букв, слогов, слов и некоторых словосочетаний). Основанием для построения модели может служить также и теоретико-информационный подход, развитый в работах К. Шеннона.

Учет частот k-грамм приводит к следующей модели открытого текста. Пусть $P^{(k)}(A)$ представляет собой массив, состоящий из приближений для вероятностей $p(b_1,b_2,...,b_k)$ появления k-грамм $b_1b_2...b_k$ в открытом тексте, $k \in N$,

$A = (a_1,...,a_n)$ — алфавит открытого текста, $b_i \in A$, $i = 1,k$.

Тогда источник "открытого текста" генерирует последовательность $c_1,c_2,...,c_k,c_{k+1},...$ знаков алфавита A, в которой k-грамма $c_1c_2...c_k$ появляется с вероятностью $p(c_1c_2...c_k) \in P^{(k)}(A)$, следующая k-грамма $c_1c_2...c_{k+1}$ появляется с вероятность $p(c_2c_3...c_{k+1}) \in P^{(k)}(A)$ и т. д. Назовем построенную модель открытого текста вероятностной моделью k-го приближения.

Таким образом, простейшая модель открытого текста -вероятностная модель первого приближения - представляет собой последовательность знаков $c_1,c_2,...,$ в которой каждый знак c_i, $i = 1,2,...,$ появляется с вероятностью $p(c_i) \in P^{(1)}(A)$,

независимо от других знаков. Будем называть также эту модель позначной моделью открытого текста. В такой модели открытый текст c1c2...c1 имеет вероятность

$$p(c1c2...cl) = \prod_{i=1}^{l} p(ci).$$

В вероятностной модели второго приближения первый знак c1 имеет вероятность p(c1) ∈ P(1)(A), а каждый следующий знак ci зависит от предыдущего и появляется с вероятностью

$$p(ci / ci-1) = \frac{p(ci-1ci)}{p(ci-1)},$$

где p(ci-1ci) ∈ P(2)(A), p(ci-1) ∈ P(1)(A), i = 2,3,.... Другими словами, модель
открытого текста второго приближения представляет собой простую однородную цепь Маркова. В такой модели открытый текст c1c2...c1 имеет вероятность

$$p(c1c2...cl) = p(c1) \cdot \prod_{i=2}^{l} p(ci / ci-1).$$

Модели открытого текста более высоких приближений учитывают зависимость каждого знака от большего числа предыдущих знаков. Ясно, что чем выше степень приближения, тем более "читаемыми" являются соответствующие модели. Проводились эксперименты по моделированию открытых текстов с помощью ЭВМ.

Отметим, что с более общих позиций открытый текст может рассматриваться как реализация стационарного эргодического случайного процесса с дискретным временем и конечным числом состояний.

Критерии распознавания открытого текста

Заменив реальный открытый текст его моделью, мы можем теперь построить критерий распознавания открытого текста. При этом можно воспользоваться либо стандартными методами различения статистических гипотез, либо наличием в открытых текстах некоторых запретов, таких, например, как биграмма ЪЪ в русском тексте. Проиллюстрируем первый подход при распознавании позначной модели открытого текста.

Итак, согласно нашей договоренности, открытый текст представляет собой реализацию независимых испытаний случайной величины, значениями которой являются буквы алфавита A = {a1,...,aп}, появляющиеся в соответствии с распределением вероятностей P(1)(A) = (p(a1),...,p(aп)). Требуется ' определить, является ли
случайная последовательность c1c2...cl букв алфавита A открытым текстом или нет.

Пусть H0 — гипотеза, состоящая в том, что данная последовательность — открытый текст, H1 — альтернативная гипотеза. В простейшем случае

последовательность c1c2...cl можно рассматривать при гипотезе H1 как случайную и равновероятную. Эта альтернатива отвечает субъективному представлению о том, что при расшифровании криптограммы с помощью ложного ключа получается "бессмысленная" последовательность знаков. В более общем случае можно считать, что при гипотезе H1 последовательность c1c2...cl представляет собой реализацию независимых испытаний некоторой случайной величины, значениями которой являются буквы алфавита A = {a1,...,ап}, появляющиеся в соответствии с распределением вероятностей Q(1)(A)= (q(al),...,q(an)). При таких договоренностях можно применить, например, наиболее мощный критерий различения двух простых гипотез, который дает лемма Неймана—Пирсона.

В силу своего вероятностного характера такой критерий может совершать ошибки двух родов. Критерий может принять открытый текст за случайный набор знаков. Такая ошибка обычно называется ошибкой первого рода, ее вероятность равна α = p{H1/H0}. Аналогично вводится ошибка второго рода и ее вероятность β = p{H0/H1} . Эти ошибки определяют качество работы критерия. В криптографических исследованиях естественно минимизировать вероятность

ошибки первого рода, чтобы не "пропустить" открытый текст. Лемма Неймана—Пирсона при заданной вероятности первого рода минимизирует также вероятность ошибки второго рода.

Критерии на открытый текст, использующие запретные сочетания знаков, например к -граммы подряд идущих букв, будем называть критериями запретных k-грамм. Они устроены чрезвычайно просто. Отбирается некоторое число s редких k-грамм, которые объявляются запретными. Теперь, просматривая последовательно k-грамму за k-граммой анализируемой последовательности c1c2...cl, мы объявляем ее случайной, как только в ней встретится одна из запретных k-грамм, и открытым текстом в противном случае. Такие критерии также могут совершать ошибки в принятии решения. В простейших случаях их можно рассчитать. Несмотря на свою простоту, критерии запретных k-грамм являются весьма эффективными.

Классификация шифров

В качестве первичного признака, по которому производится классификация шифров, используется тип преобразования, осуществляемого с открытым текстом при шифровании. Если фрагменты открытого текста (отдельные буквы или группы букв) заменяются некоторыми их эквивалентами в шифртексте, то соответствующий шифр относится к классу шифров замены. Если буквы открытого текста при шифровании лишь меняются местами друг с другом, то мы имеем дело с шифром перестановки. С целью повышения надежности шифрования шифрованный текст, полученный применением некоторого шифра, может быть еще раз зашифрован с помощью другого шифра. Всевозможные такие композиции различных шифров приводят к третьему классу шифров, которые обычно называют композиционными шифрами. Заметим, что композиционный шифр может не входить ни в класс шифров замены, ни в класс шифров перестановки (рис. 1).

Шифры

Шифры Шифры Композиционные

замены перестановки шифры

Рисунок 1. Классификация шифров

Шифры перестановки

Шифры перестановки, или транспозиции, изменяют только порядок следования символов или других элементов исходного текста. Классическим примером такого шифра является система, использующая карточку с отверстиями – решетку Кардано, которая при наложении на лист бумаги оставляет открытыми лишь некоторые его части. При зашифровке буквы сообщения вписываются в эти отверстия. При расшифровке сообщение вписывается в диаграмму нужных размеров, затем накладывается решетка, после чего на виду оказываются только буквы открытого текста.

Решетки можно использовать двумя различными способами. В первом случае

Страницы ← предыдущая следующая →

1 2 3 4 5 6 7 8 9 10