Лекция 2. Стойкость шифров

Цель лекции: ознакомиться с понятием стойкости шифра, рассмотреть конкретные примеры.

План лекции:
Введение.
1 Стойкость шифров
Заключение
Контрольные вопросы

Ключевые слова: [выбрать самостоятельно].

Содержание лекции:

Введение

Данное в предыдущей лекции определение шифра не разграничивает хорошие и плохие шифры. Однако главная цель шифра – надежная защита скрываемой информации. Поэтому качество шифра оценивается прежде всего понятием надежности защиты. В криптографии принят термин «стойкость» шифра. Под этим понимается устойчивость защиты перед атаками противника, получившего какие-либо сведения, связанные с шифром или информацией, закрываемой с помощью шифра, или просто сам шифртекст.

1 Стойкость шифров

Рассмотрим подробнее ситуацию, в которой ведется криптоанализ. В определение шифра входят следующие параметры T, x, y, k, T^{-1}. Защищаемая информация описывается параметром x. Сама идея криптоанализа основана на том, что противник перехватил и имеет на руках y. Классифицируем основные условия криптоанализа.

1. Известны только один или несколько шифртекстов y. В этих условиях решают следующие задачи:

А) найти T (определение типа шифра)

В) найти T, T^{-1}, x (дешифрование по шифртексту).

2. Известны одна или несколько пар (x, y). В этих условиях надо определить вид шифра T (T^{-1}) и найти k. Это редкая ситуация.

3. Известны вид шифра T(T^{-1}), один или несколько шифртекстов y. Найти:

А) x (бесключевое чтение);

В) к, x (дешифрование по шифртексту при известной шифрсистеме).

4. Известны: вид шифра T, T^{-1}, одна или несколько пар (x, y). Найти k. Это типичные условия криптоанализа. Стойкость шифров оценивается именно в этих условиях. Иногда отдельно возникают ситуации, когда известно много пар (x, y) так, что можно подобрать x или y, удовлетворяющие некоторым дополнительным условиям. Тогда говорят об атаке с использованием выбранного открытого или шифрованного текстов.

5. Известны T, T^{-1}, шифртекст y или пары (x, y), некоторая форма преобразования T(·, k), но неизвестны k и T^{-1}(·, k). Допустимость таких постановок задач рассматривалась Диффи и Хеллманом в 1976 г. Системы шифрования, где допускается возможность знания T(. , к) без раскрытия k и T^{-1} получили название систем с открытым ключом.

Разрешимость для противника указанных задач определяет стойкость шифра в соответствующих условиях, а методы решения этих задач называются методами криптоанализа. Универсальные методы криптоанализа – это такие методы, которые с определенными вариациями применимы ко многим шифрам.

В криптографии разработаны два базовых подхода в оценке стойкости шифров. Основы первого подхода (совершенная секретность) К. Шеннон изложил в своей работе [[1], с. 360] в 1948 г. Рассмотрим этот подход. К. Шеннон предложил следующую модель обращения с параметрами, описывающую шифрование и действия противника. Пусть X = $\{x_1, ..., x_n\}$, Y = $\{y_1, ..., y_m\}$ – множества открытых сообщений и возможных шифртекстов. Открытый текст для передачи в шифрованном виде выбирается случайно в соответствии с распределением

$$\left\{P(x_1), \dots, P(x_n), P(x_i) \geq 0, \sum_{i=1}^{n} P(x_i) = 1\right\}.$$

Тогда при решении задачи вскрытия x_i по известному y_j противник может вычислить апостериорные вероятности сообщений, которые могли быть посланы $P(x_s|y_j)$, s = 1, ..., n. Если задача дешифрования решена и открытый текст найден, то апостериорное распределение вырождено:

$$P(x_i|y_j) = 1, \;\; P(x_s|y_j) = 0 \;\text{ при } s \neq i.$$

Таким образом в апостериорных вероятностях {P(x|y), x ∈ X} отражаются даже частичные сведения об открытом тексте, полученном при перехвате криптограммы. Шеннон определил совершенную секретность (стойкость) шифра условием: апостериорные распределения на открытых текстах при любом y ∈ Y совпадают с априорным распределением на X, т.е. P(x|y) = P(x) для любых x ∈ X, y ∈ Y.

Пример 1. Пусть $X = \{0, 1\}^r$, $Y = \{0, 1\}^r$, $K = \{0, 1\}^r$, r – натуральное и

$$T(x, k) = (x \oplus k)(mod\ 2), \;\; T^{-1}(y, k) = (y \oplus k)(mod\ 2)$$

т.е. рассматривается гаммирование в двоичном алфавите. Пусть также ($P_X(x)$, x ∈ X) – произвольное распределение вероятностей на X. Предположим, что k выбирается случайно и равновероятно из K, то есть ∀ k ∈ K, $P_K(k) = 1/2^r$. Иными словами, знаки гаммы выбираются из множества {0, 1} независимо друг от друга и от открытого текста с равными вероятностями P(0) = P(1) = 1/2. Тогда для любого y ∈ Y

$$P(y) = \sum_{\substack{x \in X \\ k \in K \\ y=(x \oplus k)(mod\ 2)}} P(x,k) = \sum_{x \in X} P_X(x) \sum_{\substack{k \in K \\ 0y=(x \oplus k)(mod\ 2)}} P(k|x) =$$

$$= \sum_{x \in X} P_X(x) \frac{1}{2^r} = \frac{1}{2^r},$$

$$P(y|x) = \frac{P(x,y)}{P(x)} = \frac{P_X(x) P_K\big((x \oplus k)(mod\ 2)\big)}{P_X(x)} = \frac{1}{2^r}.$$

По формуле Байеса

$$P(x|y) = \frac{P_X(x) P(y|x)}{P(y)} = \frac{P_X(x)}{\frac{1}{2^r} 2^r} = P_X(x).$$

[1] Шеннон К. Работы по теории информации и кибернетике, М.: Иностранная литература, 1963.

Следовательно, гаммирование при помощи последовательности, полученной при помощи реализации независимых равновероятных случайных векторов, является совершенным шифром.

Подход Шеннона превратил криптографию из искусства в науку, т.к. появилась возможность доказывать защищенность информации при помощи шифра. Однако реализовать последовательность независимых и равновероятных случайных величин оказалось делом сложным. Кроме того, объем ключа в совершенном шифре совпадает с объемом сообщения, что затрудняет выполнение условия секретности ключа.

Следующим шагом в теоретическом подходе к оценке стойкости шифров послужила оценка числа открытых текстов, которые мы можем получить, применяя все возможные способы преобразования имеющихся шифр текстов при неизвестном заранее ключе. Смысл этого подхода поясним на примере, не связанном с криптографией.

Пример 2. Пусть мы играем в игру: надо угадать слово по имеющимся в порядке их следования некоторым буквам этого слова. Предположим, что дана буква «п». Если нет каких-либо ограничений на возможное слово, позволяющее однозначно назвать допустимое слово по одной известной из него букве, то существует много слов, имеющих букву «п», и мы не можем даже приблизительно указать слово, которое нам предложили угадать. То есть мы не получаем достаточно информации, спрятанной за этим словом. В аналогичных условиях мы не можем восстановить слово, даже зная несколько букв. Например, нам известно, что первыми тремя буквами слова являются буквы «при». Без дополнительных ограничений мы можем указать множество различных по смыслу слов, имеющих приставку «при». Поэтому это буквосочетание почти ничто не дает к пониманию смысла неизвестного слова.

Аналогично, если при дешифровании уравнение $T(x, k) = y$ имеет много решений x, а ключ может принимать любые допустимые значения, то возможно много осмысленных текстов x, удовлетворяющих имеющимся в нашем распоряжении ограничениям и имеющих примерно одинаковую вероятность их появления в качестве открытых текстов. Тогда нет алгоритма, решающего задачу дешифрования (в смысле однозначного нахождения открытого текста и ключа). Наоборот, во многих моделях шифров можно доказать, что такое решение единственно.

Другой подход к определению стойкости берет начало в той же работе Шеннона [[2], с. 387] и называется практической стойкостью или сложностный подход к стойкости. Традиционно рассматриваются ситуации 3.б или 4. Рассмотрим выражение $T(x, k) = y$ как уравнение относительно x и k. Тогда решение этого уравнения предполагает существование алгоритма, для которого в математике определено понятие сложности [[3].]. Сложность характеризуется двумя параметрами: число операций для вычисления результата (трудоемкость алгоритма) и объем необходимой памяти.

Число операций при данном уровне развития вычислительной техники связано со временем работы алгоритма, поэтому стойкость можно выразить в терминах времени работы алгоритма дешифрования. Естественное требование надежности шифра – высокая сложность всех возможных алгоритмов дешифрования.

Замечание 1. Возможно повышать сложность алгоритмов дешифрования, повышая сложность преобразования $T(x, k) = y$. Однако, если сложность вычисления $T(x, k)$ и $T^{-1}(y, k)$ при известном k будут велики, то практически такой шифр трудно использовать. Поэтому наряду с требованием сложности решения уравнения $T(x, k) = y$ при неизвестном k, привлекают естественное требование простоты вычисления $T(x, k)$ и $T^{-1}(y, k)$ при известном k.

---

[2] Шеннон К. Работы по теории информации и кибернетике, М.: Иностранная литература, 1963.

[3] Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. М., 1982.

Замечание 2. Требование высокой сложности при всех возможных алгоритмах дешифрования неконструктивно, т.к. опирается на потенциальную возможность перечисления всех возможных алгоритмов дешифрования. Практически [4] это требование заменяется на реализуемое условие высокой трудоемкости при всех известных создателям шифра методах дешифрования. Большинство универсальных методов опубликовано в литературе и мы будем рассматривать задачу синтеза шифров, стойких относительно универсальных методов.

Заключение

Криптограф, оценивая стойкость шифра, как правило, имитирует атаку на шифр со стороны криптоаналитика противника. Для этого он строит модель действий и возможностей противника, в которой максимально учитываются интеллектуальные, вычислительные, технические, агентурные и другие возможности противника. Примером такого подхода может служить случай в США в конце 70-х годов. Криптографы не нашли практически приемлемого алгоритма дешифрования «DES-алгоритма». Но небольшой размер ключа DES-алгоритма не позволил прогнозировать его практическую стойкость как достаточную на длительный срок, что привело к решению отказаться от использования DES-алгоритма в государственных учреждениях для защиты информации.

Контрольные вопросы

Смотри руководство по организации самостоятельной работы магистрантов.

[4] Rueppel R. Good Stream Ciphers are Hard to Design. ICCST, Zurich, 1989, c. 163-173.