УДК 004.056.55

ПРИМЕР ПОСТРОЕНИЯ АЛГОРИТМА ШИФРОВАНИЯ

М.А. Протько
Белорусский государственный университет информатики и радиоэлектроники, г. Минск.
Научный руководитель: О.Ф. Борисенко, канд. физ.-мат. наук, доцент

Введение. Что составляет любой базовый криптографический алгоритм? По сути, это два соответствия: базовый/шифрованный текст. Связь между первым и вторым происходит по некой функции F с приблизительно следующими свойствами:

$F(B) = A$ – легко рассчитываемая функция.

$B = F^{-1}(A)$ – не вычисляемая функция доступными средствами.

То есть задача построения алгоритма шифрования будет соответствовать следующей формулировке:

Пусть K – пространство ключей, e и d – ключи шифрования и расшифрования соответственно. E_e – односторонняя функция шифрования для произвольного ключа $e \in K$, такая, что $E_e(t)=c$, $c \in C$, C - пространство шифротекстов, $t \in T$, T – пространство сообщений. D_d – функция расшифрования, такая, что $D_d(c) = t$. Каждая пара (E,D) имеет свойство: зная E_e невозможно найти $E_e(t)=c$.

Учитывая необходимость в вычислительной сложности таковых функций, простой перебор всех возможных значений и применение принципа индукции могут не дать доказательства их верности. Для более качественной оценки по-

лученных алгоритмов необходимо четко обозначить условия, из которых они вытекают, а также те параметры, которым они обязаны удовлетворять.

Условия, удовлетворяемые шифром

1) $T = D(E(T))$ – где T – пространство сообщений. D и E – функции расшифрования и шифрования соответственно.

Иначе: применив функции D или E, мы получим однозначный связяный текст без потери информации в обоих случаях. Мы можем выполнять эту операцию сколь угодно долго.

2) $H(E) \geq H(T)$ – где $H(E)$ и $H(T)$ – неопределенности зашифрованного и изначального сообщения.

Иными словами, вероятность предсказания следующего символа перехваченного сообщения (нахождения закономерности) значительно уменьшается при шифровании.

3) Удовлетворяет принципу Керкго́ффса

То есть, даже зная процедуру (функцию F), невозможно получить изначальный текст. Необходим ключ, чтобы это стало возможным.

Допущения:

– У злоумышленника нет никаких априорных сведений о зашифрованном сообщении (все исходные варианты равновероятны).

– Отсутствуют специальные символы

Пример работы алгоритма:

Воспользуемся RSA алгоритмом.

Для построения RSA алгоритма нам будет необходимо:

1) Выбрать два случайных простых числа p и q
2) Найти модуль: $n = pq$
3) Найти функцию Эйлера $\varphi(n) = (p-1)(q-1)$
4) Найти простое число e, $1 < e < \varphi(n)$, где e – взаимно простое по отношению к φ(n)
5) Найти d из уравнения $de \equiv 1 (\mathrm{mod}\, \varphi(n))$

В итоге, получим пары (e,n) – открытый ключ, (d,n) – закрытый ключ.

Допустим, у нас есть ряд чисел, которые необходимо зашифровать

T: 2 1 4 3 8 5 6 n = 10 e = 3 d = 3

$$E(m) = m^e \qquad (1)$$

$$c = E(m) \bmod n \qquad (2)$$

$$D(c) = c^d \qquad (3)$$

$$m = D(c) \bmod n \qquad (4)$$

Воспользовавшись формулами 1-4 получим:

$E(m)$:	8	1	64	27	512	125	216
c:	8	1	4	7	2	5	6
$D(c)$	512	1	64	343	8	125	216
m:	2	1	4	3	8	5	6

Оценка полученного алгоритма

Для качественной оценки алгоритма нам будут необходимы формулы 5-7:

$$H(T) = -\sum_{i=1}^{n} p_i \log_2 p_i \quad (5)$$

где H(T) – мера неопределенности сообщения/шифра. T – множество возможных отправленных сообщений, p_i – вероятность отправки соответствующего сообщения T_i, n – количество битов шифротекста. Физический смысл найденной величины: количество битов информации, которое необходимо в среднем передать, чтобы полностью устранить неопределенность.

После перехвата полученного сообщения получим следующее значение неопределенности (в литературе называемое условным):

$$H(T \mid T') = -\sum_{i=1}^{n} p_i(T \mid T') \log_2 p_i(T \mid T') \quad (6)$$

Где $p(T_i \mid T')$ – вероятность того, что исходное сообщение есть T_i при условии, что результат шифрования T'.

Отсюда найдем значимую характеристику нашего шифра:

$$I = H(T) - H(T \mid T') \quad (7)$$

Где I - информация об исходном тексте, которую злоумышленник может извлечь из перехваченного шифротекста. Необходимо обеспечить, чтобы $H(T) \to H(T \mid T')$ ($I \to 0$).

Т.е, чем меньше I, тем меньше вероятность однозначного дешифрования без знания ключа.

Качественная оценка полученного алгоритма

Для начала предположим, что один бит сообщения T – одна цифра. Передается по порядку следования. Никаких манипуляций, связанных с непоследовательной передачей сообщения, не было. Также не передается «шум» - какие-либо последовательности, не имеющие логического смысла.

Предположим, что злоумышленнику известно, что предаются цифры от 1 до 9.

Посчитаем H(T) из вышеописанных условий:

$p_i = 0{,}11111$ (1 к 9 , где событие – выбор конкретной цифры.)

Поскольку мы рассматриваем априорный случай, предыдущие сообщения никак не влияют на последующие:

$$H(T) = -\sum_{i=1}^{7} 0{,}11111 \log_2 0{,}11111 = -2{,}4654 \quad (8)$$

Если никаких апостериорных сведений не было получено:

$H(T) = H(T \mid T')$, I=0 – самый лучший исход, возможный только теоретически.

Если предположить, что цифры не повторяются:

$$H(T \mid T') - \sum_{i=1}^{n} \frac{1}{n-i+1} \log_2 \left(\frac{1}{n-i+1} \right) = -3{,}0518 \quad (9)$$

$$I = -2{,}4654 - (-3{,}0518) = 0{,}5864 \quad (10)$$

Вывод

Из вышеизложенного следует, что для построения шифра с высокими показателями криптографической стойкости необходимо учитывать вычислительную сложность алгоритма шифрования, условия передачи информации и изна-

чальные показатели незашифрованного сообщения, поскольку из знания последних будет возможно найти уязвимость, которая существенно влияет на защиту информации.

Список использованных источников:

1. Липницкий, В. А. Современная прикладная алгебра. Математические основы защиты информации от помех и несанкционированного доступа: учеб.-метод. Пособие / В. А. Липницкий. – Минск: БГУИР, 2006.-87 с.