УДК 621.391.25

В.Ф. Голиков, Ф. Абдольванд

ОЦЕНКА ПОТЕРЬ КОНФИДЕНЦИАЛЬНОСТИ ПРИ НЕКЛАССИЧЕСКИХ СПОСОБАХ ФОРМИРОВАНИЯ КРИПТОГРАФИЧЕСКОГО КЛЮЧА

Рассматривается оценка потерь конфиденциальности общей ключевой бинарной последовательности, формируемой без использования однонаправленных функций. Потери конфиденциальности возникают вследствие обмена прямой или косвенной информацией о формируемой последовательности по открытому каналу связи с целью устранения имеющихся несовпадений. В качестве меры конфиденциальности используется энтропия последовательности.

Введение

Одной из главных проблем, которую необходимо решать для достижения правильного функционирования симметричной криптосистемы, является проблема обеспечения абонентов системы общим секретным ключом. Эта проблема в настоящее время решается по-разному. Для подобного рода систем уже длительное время с успехом используется алгоритм открытого распределения ключей Диффи – Хеллмана или аналогичные процедуры, базирующиеся на использовании односторонних функций. Однако в последние годы бурное развитие физики, электроники, математики и информатики сделало вполне реальным появление квантового компьютера, одним из возможных применений которого является «взлом» традиционных односторонних функций с последующим вычислением общего ключа, формируемого по схеме Диффи – Хеллмана. В альтернативных способах формирования общего ключа без использования классических однонаправленных функций [1, 2] у сторон возникает необходимость обмениваться информацией об имеющихся текущих значениях «сырого ключа». Поскольку обмен такой информацией осуществляется по открытому каналу, существует потенциальная угроза, что эта информация будет доступна криптоаналитику, «прослушивающему» открытый канал. В связи с этим возникает необходимость оценить либо утечку конфиденциальности формируемого ключа, либо его остаточную конфиденциальность.

1. Постановка задачи

При решении данной задачи конфиденциальность формируемой общей случайной последовательности будем оценивать, используя энтропийную трактовку степени неопределенности случайного числа [3]. Известно, что неопределенность дискретной случайной величины X, принимающей значения x_i, $i = 1, 2, 3, \ldots, M$, можно оценить энтропией

$$H = -\sum_{i=1}^{M} P(x_i) * \log\left(P(x_i), 2\right),$$

где $P(x_i)$ – вероятность того, что $X = x_i$. Если X – бинарный вектор, т. е. $X = \{x_1, x_2, \ldots, x_n\}$, где $x_i \in \{0,1\}$, $P(x_i = 1) = p$, $P(x_i = 0) = 1 - p = q$, то в силу независимости компонент этого вектора

$$H(X) = H(\{x_1, x_2, \ldots, x_n\}) = \sum_{i=1}^{n} H(x_i),$$

где $H(x_i) = -\left(p \log(p, 2) + q \log(q, 2)\right)$. Окончательно получаем

$$H(X) = -\sum_{i=1}^{n} \left(p \log(p, 2) + q \log(q, 2)\right) = -n\left(p \log(p, 2) + q \log(q, 2)\right). \quad (1)$$

Зависимость (1) изображена на рис. 1.

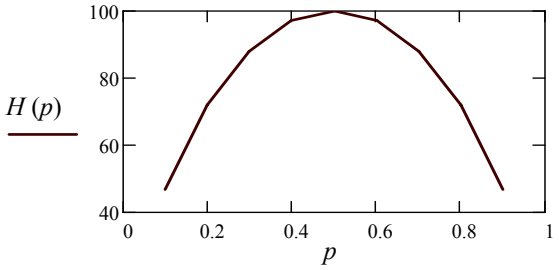

Рис. 1. Зависимость энтропии бинарной последовательности H от величины p

Для ключевой последовательности необходимо, чтобы ее энтропия имела максимальное значение. Энтропия будет максимальна при $p = 0{,}5$. Действительно, при $p = q = 0{,}5$ получим

$$H_{\max} = -n\left(\frac{1}{2}\log\left(\frac{1}{2}, 2\right) + \frac{1}{2}\log\left(\frac{1}{2}, 2\right)\right) = n.$$

При утечке некоторой информации о ключевой последовательности длиной n ее энтропия будет уменьшаться, что эквивалентно уменьшению длины последовательности. Будем обозначать эту длину через $n_э$.

Пусть стороны A и B сформировали случайные бинарные последовательности X_A и X_B длиной n независимо друг от друга. Обмениваясь определенной информацией об этих последовательностях, стороны согласовывают свои последовательности, устраняя несовпадения (ошибки), т. е. формируют одинаковую последовательность X_{AB}. Будем считать, что криптоаналитик перехватывает эту информацию и пытается вычислить X_{AB}.

Энтропия последовательности X_{AB} при отсутствии прослушивания равна n. Рассмотрим, как снижается энтропия бинарной последовательности при различных утечках информации о ней к криптоаналитику.

2. Оценка потерь конфиденциальности при различных видах разглашения информации

Выделяют несколько видов разглашения информации:

1. Стороны A и B огласили значения и порядковые номера k бит $(k < n)$ своей последовательности.

Криптоаналитик конструирует свою последовательность X_E, в которой k бит ему известны и являются правильными, а $(n-k)$ являются неизвестными. Нетрудно видеть, что число возможных значений угадываемой последовательности $M_{oct} = 2^{(n-k)}$, остаточная энтропия и эквивалентная длина последовательности $H_{oct} = n_э = n - k$. На рис. 2 изображена зависимость $W(k) = \dfrac{H}{H_{oct}}$ при $n = 100$.

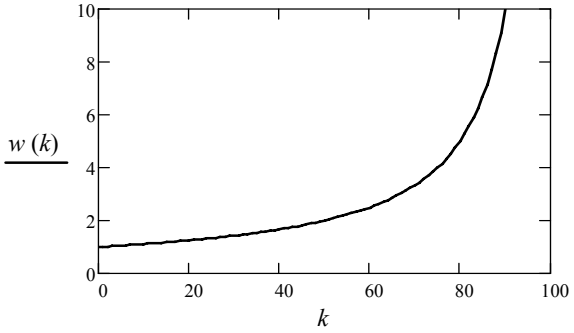

Рис. 2. Относительные потери энтропии при оглашении значения и порядковых номеров k бит

2. Криптоаналитику становится известным, что в его последовательности k бит являются правильными, а $(n-k)$ – неправильными (противоположными), однако порядковые номера и тех и других бит ему неизвестны.

Число комбинаций, в которых k бит остаются без изменения (криптоаналитик считает их правильными), а остальные $(n-k)$ бит меняются на противоположные, равно $M_{oct} = C_k^n$. Очевидно, что $C_k^n < 2^n$. Так как все значения неизвестной последовательности из множества C_k^n равновероятны, вероятность каждой комбинации равна $1/C_k^n$, а энтропия

$$H_{oct} = -\sum_{i=1}^{C_k^n} 1/C_k^n * \log\left(\frac{1}{C_k^n}, 2\right) = \log(C_k^n, 2).$$

Из последнего равенства можно определить эквивалентную длину последовательности $n_э = \log(C_k^n, 2)$.

На рис. 3 изображена зависимость $W(k) = \dfrac{H}{H_{oct}} = \dfrac{n}{n_э}$ при $n = 100$. Видно, что наибольшие потери энтропии наблюдаются при малых значениях k и значениях k , близких к n . Действительно, при $k = 0$ все биты в X_E правильные, т. е. искомая последовательность полностью известна. Аналогично при $k = n$ все биты в X_E неправильные и после их инвертирования искомая последовательность полностью известна.

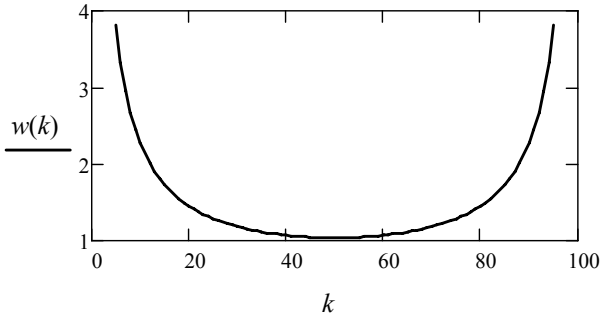

Рис. 3. Относительные потери энтропии для случая 2

3. Криптоаналитику становится известным, что в его последовательности k бит являются правильными, при этом порядковые номера этих бит ему неизвестны, остальные биты могут быть любыми.

Такая информация о X_{AB} позволяет сделать заключение, что число правильных бит в X_E определяется неравенством $k \le i \le n$, где $i-$ возможное значение числа правильных бит. Считая для каждого i , что остальные $(n-i)$ бит являются неправильными, и перебрав все возможные значения i , получим, что число возможных значений X_E

$$M_{oct} = \sum_{i=k}^{n} C_i^n .$$

Из последнего равенства можно определить эквивалентную длину последовательности $n_э = \log(\sum_{i=k}^{n} C_i^n, 2)$. Остаточная энтропия $H_{oct} = n_э = \log(\sum_{i=k}^{n} C_i^n, 2)$.

На рис. 4 изображена зависимость $W(k) = \dfrac{H}{H_{oct}} = \dfrac{n}{n_э}$ при $n = 100$. Видно, что потери энтропии в данном случае меньше, чем в предыдущем, так как полученная информация о X_{AB} носит менее определенный характер.

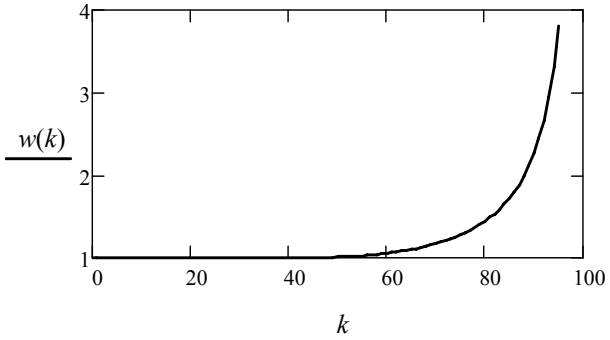

Рис. 4. Относительные потери энтропии для случая 3

4. Криптоаналитику становится известной четность (нечетность) суммы фрагментов X_{AB}:

$$x_1^{(1)} \oplus x_2^{(1)} \oplus x_3^{(1)} \oplus \ldots \oplus x_{n_2}^{(1)} = c_1;$$
$$x_1^{(2)} \oplus x_2^{(2)} \oplus x_3^{(2)} \oplus \ldots \oplus x_{n_2}^{(2)} = c_2;$$
$$x_1^{(s)} \oplus x_2^{(s)} \oplus x_3^{(s)} \oplus \ldots \oplus x_{n_S}^{(s)} = c_s,$$

где $x_j^{(i)} - j$-й бит i-го фрагмента, $j = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, s$; $n_i -$ число бит в i-м фрагменте; $s -$ число фрагментов; $c_i = \{0,1\}$; $\oplus -$ сложение по mod 2. Из каждого уравнения системы можно выразить один из битов, например:

$$x_1^{(1)} = c_1 \oplus x_2^{(1)} \oplus x_3^{(1)} \oplus \ldots \oplus x_{n_2}^{(1)};$$
$$x_1^{(2)} = c_2 \oplus x_2^{(2)} \oplus x_3^{(2)} \oplus \ldots \oplus x_{n_2}^{(2)};$$
$$x_1^{(s)} = c_s \oplus x_2^{(s)} \oplus x_3^{(s)} \oplus \ldots \oplus x_{n_S}^{(s)}.$$

Это эквивалентно знанию s бит. Следовательно, остаточная энтропия $H_{oct} = n - s,$ т. е. знание четности (нечетности) одного фрагмента эквивалентно знанию одного бита.

5. Криптоаналитику E становятся известными значения и порядковые номера k бит в последовательности X_{AB}. Сторонам A и B об этом известно, однако они не знают, какие биты из X_{AB} известны E. Будем считать, что известные E биты распределены в последовательности X_{AB} случайным образом, обозначим их $a_1, a_2, \ldots, a_k$.

Если A и B не предпринимают специальных мер по увеличению конфиденциальности X_{AB}, то энтропия X_{AB} согласно п. 1 равна $H_{oct} = n - k.$ Однако в ряде работ [4] предлагается использовать процедуру «усиления секретности» сформированного ключа. Суть одного из ее вариантов заключается в согласованном преобразовании X_{AB} сторонами A и B, при котором каждый бит итоговой последовательности $Y = \{y_1, y_2 \ldots, y_{n/2}\}$ вычисляется по формуле $y_i = x_u \oplus x_v,$ где $u, v = 1, 2, \ldots, n$; $x_u, x_v -$ биты исходной последовательности, выбранные случайным образом $(u \neq v)$.

Будем считать, что алгоритм этого преобразования известен E. Таким образом, каждый бит итоговой последовательности является суммой либо двух бит, неизвестных E, либо одного известного E бита и одного неизвестного, либо двух известных E битов. Для пар первых двух типов суммарный бит известен E с вероятностью 0,5. Для пары третьего типа суммарный бит известен с вероятностью 1, так как это сумма известных ему бит. Уменьшение энтропии итоговой последовательности Y происходит за счет того, что она оказывается в два раза короче X_{AB} и за

счет того, что по-прежнему часть ее бит остается известной E. Оценим сокращение энтропии за счет сохранения бит третьего типа, являющихся известными для E и неизвестными для A и B.

Обозначим число итоговых бит первого, второго и третьего типов соответственно через m_0, m_1, m_2 (индексы выбраны по числу бит, входящих в сумму и известных E). Величины m_0, m_1, m_2 являются случайными и зависят от значений n и k. Найдем закон распределения вероятностей m_2. Очевидно, что

$$m_0 + m_1 + m_2 = n/2;$$
$$m_1 + 2m_2 = k.$$

Вероятность того, что y_1 будет состоять из двух неизвестных для E бит,

$$P_0 = P(x_u = H, x_v = H) = P(x_u = H) * P(x_v = H / x_u = H) = \frac{n-k}{n} * \frac{n-1-k}{(n-1)},$$

где $H-$ неизвестный для E бит.

Вероятность того, что y_1 будет состоять из одного неизвестного бита и одного известного,

$$P_1 = P(x_u = H, x_v = И) = P(x_u = И, x_v = H) = P(x_u = H) * P(x_v = И / x_u = H) +$$
$$+P(x_u = И) * P(x_v = H / x_u = И) = \frac{n-k}{n} * \frac{k}{(n-1)} + \frac{k}{n} * \frac{n-1-(k-1)}{(n-1)},$$

где $И-$ известный для E бит.

Вероятность того, что y_1 будет состоять из двух известных битов,

$$P_2 = P(x_u = И, x_v = И) = P(x_u = И) * P(x_v = И / x_u = И) = \frac{k}{n} * \frac{k-1}{(n-1)}.$$

Из выражений для P_0, P_1, P_2 видно, что при конечных значениях k, n вероятность попадания в $y_i, 0, 1, 2$ известных бит зависит от того, какие биты попали в предыдущие итоговые биты, т. е. P_0, P_1, P_2 являются функциями i. В связи с этим процесс образования y_i относится к классу дискретных случайных процессов с памятью и его точный анализ достаточно сложен. Поэтому для упрощения решения поставленной задачи вероятности P_0, P_1, P_2 считались независимыми от i. В результате были получены выражения для закона распределения вероятностей m_2 и его математического ожидания:

$$B(i) = P(m_2 = i) = \sum_{j=\frac{k}{2}-i}^{\frac{n}{2}-i} \frac{(n/2)!}{i!\, j!\, (\frac{n}{2} - j - i)!} P_0^{(\frac{n}{2}-j-i)} P_1^{j} P_2^{i};$$

$$M m_2 = \sum_{i=0}^{k/2} i * B(i).$$

Вероятности $B(i)$ при различных значениях n, k изображены на рис. 5 ($B1(i)$ – для $n = 200, k = 20$; $B2(i)$ – для $n = 200, k = 40$; $B3(i)$ – для $n = 200, k = 80$).

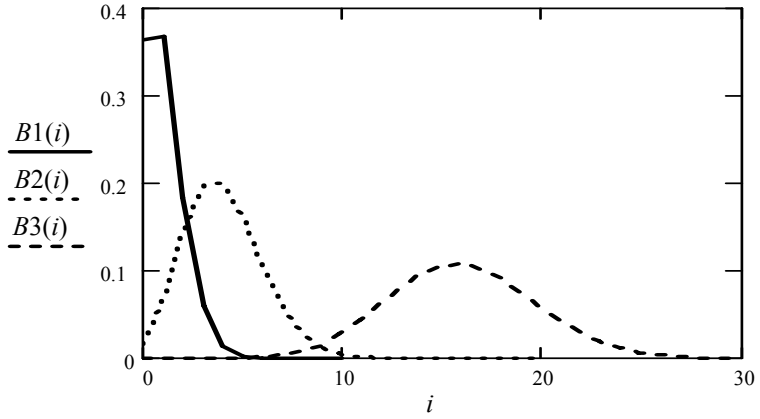

Рис. 5. Графики зависимости $B(i)$

Соответствующие математические ожидания $M1m_2 = 1, M2m_2 = 4, M3m_2 = 16.$

Полученные результаты, по мнению авторов, следует интерпретировать следующим образом. С точки зрения лица, применяющего процедуру усиления секретности для исключения известных E бит, энтропия итоговой последовательности за счет суммирования пар бит исходной последовательности уменьшится в два раза. В итоговой последовательности попрежнему остаются биты, известные E, количество которых можно спрогнозировать с учетом приведенных выше результатов. В рассматриваемом примере только при $k < 10$ с вероятностью, близкой к 1, можно утверждать, что $m_2 = 0$, в то время как для E энтропия итоговой последовательности $H_{ост} = n/2 - m_2^*$, где $m_2^* -$ полученное значение m_2.

Заключение

Процесс формирования криптографических ключей неклассическими способами сопровождается снижением их конфиденциальности. Величина этого снижения зависит от вида разглашаемой информации о формируемом ключе. Для количественного измерения конфиденциальности наиболее удобно использовать энтропию сформированного ключа. Применение изложенного подхода позволило получить оценки снижения энтропии ключевой последовательности для типичных видов утечки информации в процессе формирования ключа.

Список литературы

1. Kinzel, W. Interacting neural networks and cryptography / W. Kinzel, I. Kanter // Advances in Solid State Physics ; ed. B. Kramer. – Springer Verlag, 2002. – 122 p.

2. Голиков, В.Ф. О распределении ключевой информации в современных информационных системах / В.Ф. Голиков, Ф. Абдольванд // 14-я Междунар. конф. «Комплексная защита информации». – Могилев, 2009. – С. 77–79.

3. Основы криптологии / Ю.С. Харин [и др.]. – Минск : Новое знание, 2003. – 382 с.

4. Боумейстер, Д. Физика квантовой информации / Д. Боумейстер, А. Экерт, А. Цайлингер. – М. : Постмаркет, 2002. – 276 с.

Поступила 13.04.11

Белорусский национальный технический университет,
Минск, пр-т Независимости, 65
e-mail: vgolikov@bntu.by
faridabdolvand@gmail.com

V.F. Golikov, F. Abdolvand

EVALUATION OF LOSSES OF CONFIDENTIALITY BY A NON-CLASSICAL METHOD OF FORMING CRYPTOGRAPHIC KEYS

In this paper, an evaluation of losses of confidentiality of the general key binary sequence formed without use of unidirectional functions is considered. Confidentiality losses arise as a result of the exchange of direct or indirect information on the formed sequence via an open communication channel in order to eliminate existing discrepancies. As a confidentiality measure, an entropy measure of sequence is used.