УДК 621.391.25

КОНФИДЕНЦИАЛЬНАЯ ИДЕНТИФИКАЦИЯ ДВОИЧНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Докт. техн. наук, проф. ГОЛИКОВ В. Ф.,
асп. АБДОЛЬВАНД Ф.

Белорусский национальный технический университет

При решении некоторых задач криптографической защиты информации возникает необходимость оценивать степень близости двоичных последовательностей, сохраняя при этом конфиденциальность последних. Например, при формировании общего ключа симметричной криптосистемы с использованием квантового канала передачи информации после окончания сеанса передачи одиночных поляризованных фотонов от абонента A к абоненту B и удаления из ключевой последовательности бит, принятых в несогласованных базисах [1, 2], оставшаяся часть последовательности (сырой ключ) проверяется на наличие отличий.

Эти отличия носят случайный характер и возникают вследствие шумовых эффектов в канале связи или в результате «прослушивания» квантового канала злоумышленником. В этом случае возникшие отличия можно считать ошибками, считая двоичную последовательность A правильной, а последовательность B искаженной. Необходимость определения уровня ошибок возникает, с одной стороны, для обнаружения факта «прослушивания» квантового канала, с другой – для принятия решения о целесообразности дальнейшей процедуры согласования последовательностей. Действительно, при «прослушивании» канала злоумышленник узнает некоторую часть будущего криптографического ключа, но при этом вносит дополнительные ошибки в передаваемую последовательность. Поэтому при обнаружении факта прослушивания сеанс формирования общего ключа может быть отменен. Кроме того, сеанс формирования общего ключа может закончиться неудачей и при отсутствии прослушивания, если уровень ошибок по естественным причинам достаточно высок. Как показано в [3], устранение ошибок (согласование последовательностей) сопровождается уменьшением конфиденциальности общего ключа, которое тем больше, чем больше уровень ошибок. Если обозначить: длительность последовательности n_0, вероятность ошибки p_0, то для устранения всех ошибок потребуется огласить не менее r бит

$$r = -n_0(p_0\log_2 p_0 + (1-p_0)(1-\log_2 p_0)).$$

Зависимость $u(p_0) = \frac{r}{n_0}\cdot 100$ изображена на рис. 1.

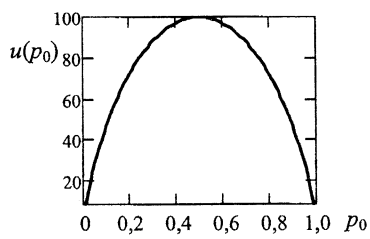

Рис. 1. Зависимость доли открываемых бит $\frac{r}{n_0}\cdot 100$ от вероятности ошибки p_0

Как видно из рис. 1, например при уровне ошибок $p_0 > 30$ % для их устранения требуется открыть не менее 89 % бит ключа. Таким образом, возникает задача перед началом процедуры устранения ошибок определить уровень отличия в паре последовательностей или при наличии возможности выбора таких пар отобрать пары с минимальным уровнем отличий.

Пусть абонент A сформировал некоторым способом двоичную последовательность K^A =

$= \{k_i^A\}$, $k_i^A = \{1, 0\}$ – элемент двоичной последовательности, $i = \overline{1, n_0}$. Аналогичная последовательность $K^B = \{k_i^B\}$ имеется у абонента B. Это может быть реализовано путем передачи последовательности K^A абоненту B по защищенному каналу связи с ошибками (квантовый канал связи), либо абонент B формирует K^B автономно, получив от A некоторую информацию о K^A по открытому каналу связи. В дальнейшем независимо от способа получения последовательностей K^A и K^B будем считать, что A и B имеют возможность обмениваться информацией по открытому каналу связи (по каналу, доступному криптоаналитику C) и им необходимо выявить уровень отличий между K^A и K^B.

Пусть каждый бит k_i^A с вероятностью p_A принимает значение 1, либо 0 с вероятностью $q_A = 1 - p_A$, аналогично для k_i^B: p_B, $q_B = 1 - p_B$, тогда вероятность того, что i-е биты последовательностей K^A и K^B будут одинаковы, равна

$$P_c = P\left(k_i^A = k_i^B\right) = P\left(k_i^A = 1,\ k_i^B = 1\right) + P\left(k_i^A = 0,\ k_i^B = 0\right).$$

Считая k_i^A и k_i^B независимыми, получим

$$P_c = P\left(k_i^A = 1\right)P\left(k_i^B = 1\right) + P\left(k_i^A = 0\right)P\left(k_i^B = 0\right) = p_A p_B + q_A q_B.$$

Аналогично, вероятность того, что i-е биты последовательностей K^A и K^B будут противоположны, равна:

$$P_0 = P\left(k_i^A = 1\right)P\left(k_i^B = 0\right) + P\left(k_i^A = 0\right)P\left(k_i^B = 1\right) = p_A q_B + q_A p_B.$$

Очевидно, что $P_c + P_0 = 1$. Если от последовательностей K^A и K^B перейти к некой виртуальной последовательности $K^0 = \{k_i^A\}$, где $k_i^0 = \{1, 0\}$ – элемент двоичной последовательности, причем $k_i^0 = 1$, если $k_i^A = k_i^B$, и $k_i^0 = 0$, если $k_i^A \neq k_i^B$, то количество нулей в K^0 равно числу несовпадающих бит в последовательностях K^A и K^B, а количество единиц – числу совпадающих бит. Обозначим количество нулей через s, тогда вероятность того, что эта случайная величина примет значение D, равна

$$P(s = D) = \binom{n_0}{D} P_0^D P_c^{n_0 - D}.$$

P_c Последнее соотношение показывает, что количество несовпадений в последовательностях K^A и K^B имеет биномиальное распределение, параметры которого P_c и P_0 зависят от p_A и p_B (рис. 2), которые в свою очередь зависят от способа формирования последовательностей K^A и K^B.

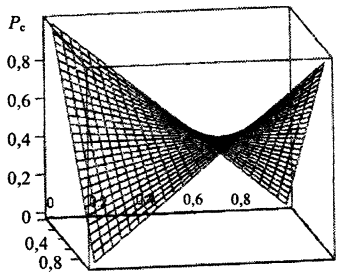

Рис. 2. Зависимость P_c от p_A и p_B

Таким образом, зная способ формирования последовательностей K^A и K^B, можно рассчитать среднее значение или диапазон возможных значений для D. Однако для последовательностей конечной длины, например в пределах ста бит, реальное значение D может сильно отличаться от расчетного. Поэтому необходимо использовать методы анализа конечных совокупностей.

Для оценки количества несовпадений в последовательностях K^A и K^B будем рассматривать виртуальную последовательность K^0, содержащую D несовпадающих бит. Если из K^0, считая ее некой генеральной совокупностью с параметрами n_0 и D, извлечь случайную выборку с параметрами n и d, где n – объем выборки, d – количество нулей в выборке, то задачу оценки количества несовпадений в последовательностях K^A и K^B можно свести к задаче оценки параметров распределения генеральной совокупности по результатам ее выборочного исследования. Поскольку генеральная совокупность K^0 существует виртуально, для создания выборки из нее абонент A извлекает из K^A

n элементов k_i^A, где $i = I_j^A$, $I_j^A \in \overline{1, n_0}$, где $j = 1, 2, \ldots, n$, и пересылает эту выборку абоненту B. Абонент B в свою очередь извлекает из K^B n элементов k_i^B, где $i = I_j^B$, $I_j^B \in \overline{1, n_0}$, где $j = 1, 2, \ldots, n$. Сравнивая элементы с одинаковыми номерами из выборок, абонент B формирует выборку, которая получилась бы, если бы она была извлечена из виртуальной генеральной совокупности K^0. Обозначим ее как $K^V = \{k_i^v\}$, $k_i^v = \{1,0\}$ – элемент выборки, $k_i^v = 1$, если $k_i^A = k_i^B$, и $k_i^v = 0$, если $k_i^A \neq k_i^B$. Таким образом, наблюдая в выборке объемом n_0 количество несовпадений (нулей) d, можно сделать определенные выводы о количестве несовпадений (нулей) D в генеральной совокупности объемом n_0. Ограничимся рассмотрением одноступенчатого плана выборочного контроля по альтернативному признаку. Согласно [4] вероятность обнаружения в выборке d нулей при условии, что число нулей в генеральной совокупности равно D, равна

$$h_{n_0,D}^{n,D} = \frac{\binom{D}{d}\binom{n_0 - D}{n - d}}{\binom{n_0}{n}},$$

что соответствует гипергеометрическому распределению вероятностей. Одноступенчатым планом выборочного контроля с параметрами (n, w) для рассматриваемой задачи называется такое правило, при котором из n_0 берется случайная выборка без возвращения объемом n. Если число d обнаруженных в выборке нулей больше w, то генеральная совокупность содержит нулей больше D, если же $d \le w$, то нулей меньше D. Оперативной характеристикой одноступенчатого плана называется вероятность того, что $d \le w$, рассматриваемая как функция числа D, обозначим ее как $z(D) = h_{n_0,D}^{n,w}$. При планировании параметров выборочного контроля необходимо определить значения n, w, а также задаться граничными уровнями ошибок D_1, D_2 ($D_1 < D_2$). Будем выбирать параметры n, w таким образом, что если $D = D_1$, то вероятность события $d \le w$ будет не менее $1 - \alpha$, а если $D = D_2$, то вероятность события $d \le w$ будет не более β. Числа α и β – риски первого и второго рода. В рассматриваемой задаче они имеют соответственно смысл вероятности того, что при числе несовпадений в последовательностях K^A и K^B, равном D_1 и менее, будет сделан вывод о том, что число несовпадений более D_1, а при числе несовпадений в последовательностях K^A и K^B, равном D_2 и более, будет сделан противоположный вывод. В соответствии с изложенным выше w и n найдем как решение системы уравнений:

$$\begin{cases} \sum_{i=0}^{w} h_{n_0,D_1}^{n,i} \ge 1 - \alpha; \\ \sum_{i=0}^{w} h_{n_0,D_2}^{n,i} \le \beta. \end{cases} \quad (1)$$

Данная система решается численными методами. Например, если сравниваются две последовательности длиной $n_0 = 120$ и требуется определить, не превышает ли уровень несовпадений 30 %, то выбрав $D_1 = 20$, $D_2 = 40$, $\alpha = 0{,}9$, $\beta = 0{,}1$, получим $n = 31$, $w = 7$. Решение (1) поясняется на рис. 3, где $HD_1 = \sum_{i=0}^{w} h_{n_0,D_1}^{n,i}$, $HD_2 = \sum_{i=0}^{w} h_{n_0,D_2}^{n,i}$.

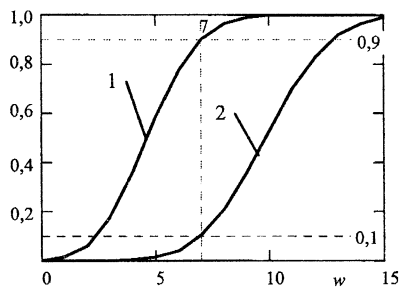

Рис. 3. Решение системы уравнений: $1 - HD_1(w)$; $2 - HD_2(w)$

Таким образом, если в выборке объемом $n = 31$ из виртуальной последовательности объемом $n_0 = 120$ получим $d \le 7$, то уровень несовпадений не превышает 30 %. Точность идентификации можно повысить, сближая D_1 и D_2, однако при этом существенно возрастает объем необходимой выборки (при прочих равных условиях, если выбрать $D_1 = 20$, $D_2 = 30$, то получим $n = 65$, $w = 13$). А поскольку биты, во-

шедшие в выборку, исключаются из будущего ключа, необходимо находить разумный компромисс между точностью идентификации и потерями.

ВЫВОД

Предложенный способ идентификации двоичных последовательностей с использованием процедуры выборочного контроля по альтернативному признаку позволяет достаточно эффективно решать задачу при умеренных потерях конфиденциальности элементов последовательностей и может с успехом использоваться для формирования общего ключа симметричных криптосистем.

ЛИТЕРАТУРА

1. **Bennet, C. H., Brassard, G.** // Proc. IEEE Int. Conference on Computers, Systems and Signal Processing, IEEE. – New York, 1984.

2. **Bennet, C. H.** Phys. Rev. // Lett. 68, 3121. – 1992.

3. **Брассар, Ж.** Современная криптология / Ж. Брассар. – М.: Полимед, 1999.

4. **Беляев, Ю. К.** Вероятностные методы выборочного контроля / Ю. К. Беляев. – М.: Наука, 1975.

Поступила 24.04.2009

УДК 004.8.032.26: 631.472.6

РАЗРАБОТКА НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ МИГРАЦИИ ХИМИЧЕСКИХ ВЕЩЕСТВ В ПОЧВЕ И АЛГОРИТМОВ ИХ ОБУЧЕНИЯ

Докт. техн. наук, проф. КУНДАС С. П., КОВАЛЕНКО В. И., ХИЛЬКО О. С.

Международный государственный экологический университет имени А. Д. Сахарова

Анализ процессов [1–5], влияющих на миграцию химических веществ в почве, показывает, что для осуществления моделирования и прогнозирования в этой области необходимо учитывать большое количество факторов, связанных как с особенностями самого загрязняющего вещества, так и с условиями его распространения. В силу этого большинство разработанных моделей либо значительно упрощают реальные процессы, не учитывая при этом отдельные факторы, либо для уменьшения погрешности моделирования учитывают все наиболее влияющие факторы, что приводит к их значительному усложнению. Однако усложнение модели связано со значительными проблемами в формировании их входных параметров и граничных условий. Причем некоторые входные параметры и в сложных феноменологических моделях можно получать только экспериментальным путем (например, изотерма сорбции водяного пара), некоторые требуют проведения дополнительных расчетов (тепловая задача, расчет движения влаги и т. п.). А это в свою очередь приводит к проблеме численной реализации и применимости модели, так как даже если предложенная модель и позволяет получать корректные результаты, для ее применения необходимо иметь достаточно объемную базу данных входных параметров, что не всегда представляется возможным.

Проведенный анализ существующих математических моделей миграции веществ в почве показывает, что наибольшее распространение получили два класса моделей: эмпирические модели [1–3 и др.], основанные на решении уравнения конвективной диффузии, и модели,