В.Ф. ГОЛИКОВ, Н.В. БРИЧ, В. Л. ПИВОВАРОВ

О НЕКОТОРЫХ ПРОБЛЕМАХ В ЗАДАЧАХ РАСПРЕДЕЛЕНИЯ КРИПТОГРАФИЧЕСКИХ КЛЮЧЕЙ С ПОМОЩЬЮ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Введение

В работах [1,2] предложено использовать синхронизируемые искусственные нейронные сети (ИНС) для решения задачи распределения ключей криптографической системы между двумя абонентами, имеющими не защищенный от прослушивания канал связи и не обладающими общим секретом. Предложенный метод в дальнейшем анализировался многими исследователями, в том числе и авторами этой работы [3], особенно тщательный анализ представлен в [4]. Однако, на наш взгляд, до сих пор не найдены ответы на некоторые очень важные вопросы, такие как:

– всегда ли процесс синхронизации по входам взаимодействующих ИНС заканчивается выравниванием векторов весовых коэффициентов персептронов, входящих в ИНС (значения изначально задаются абонентами случайно и независимо друг от друга), т.е. является ли процесс сходящимся;

– как определить момент наступления полной синхронизации сетей (полного равенства векторов весовых коэффициентов).

Основная часть

Пусть имеются ИНС, состоящие из K внутренних персептронов (рис. 1). Каждый персептрон имеет N входов.

Значения дискретной входной величины с равномерным распределением обозначено как $x_{kj} = \pm 1$, где $k = 1,2,\ldots,K$, $j = 1,2,\ldots,N$. Значение на выходе k-го внутреннего персептрона отправителя (получателя) обозначено как $^{A\ B}$.

Индекс A/B означает, что операция касается обеих сетей A и B, а единичный индекс – что операция касается одной сети соответственно. Выходная величина O каждой ИНС:

$$O^{A/B} = \prod_{k=1}^{K} y_k^{A/B} = \prod_{k=1}^{K} \sigma(\alpha_k^{A/B}) = \prod_{k=1}^{K} \sigma\left(\sum_{j=1}^{N} w_{kj}^{A/B} x_{kj} \right), \quad (1)$$

где $w_{kj}^{A/B}$ – весовые коэффициенты (ВК) персептронов; $\sigma(\alpha_k^{A/B})$ – модифицированная функция знака.

Модифицированная функция знака

$$\sigma(\alpha_k^{A/B}) = \begin{cases} 1, \sigma(\alpha_k^{A/B}) \geq 0, \\ -1, \sigma(\alpha_k^{A/B}) < 0. \end{cases} \quad (2)$$

Если выходы обеих сетей идентичны

$$O^A = O^B, \quad (3)$$

то векторы весов тех персептронов, для которых $(O^{A/B} \cdot y_k^{A/B}) > 0$, корректируются в соответствии с правилом:

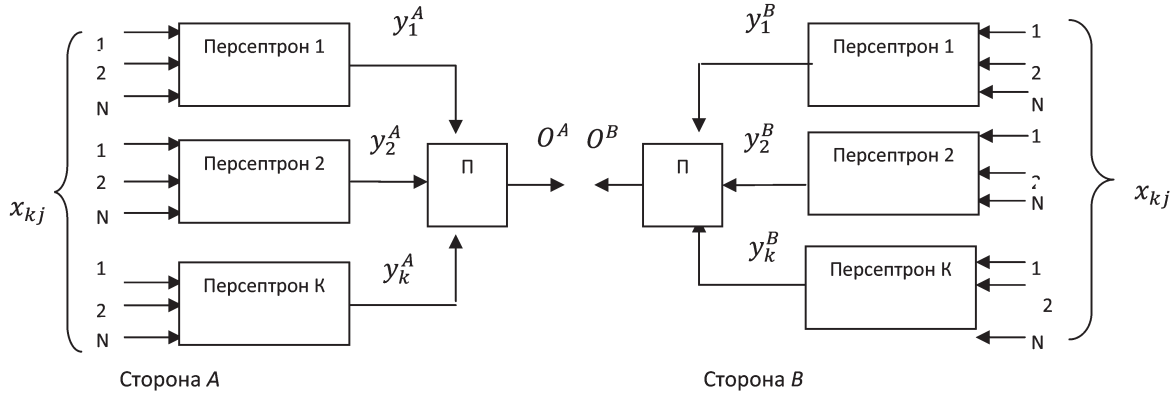

Рис. 1. Синхронизируемые ИНС

$$w_{kj}^{A/B}[i+1] = w_{kj}^{A/B}[i] + x_{kj}[i]O^{A/B}[i]. \quad (4)$$

Если $|w_{kj}^{A/B}| > L$, тогда $w_{kj}^{A/B} = L$ с соответствующим знаком.

В процессе синхронизации по открытому каналу передаются только значения входного вектора x_{kj} и выходных значений $O^{A/B}$.

Сходимость процесса

Для сетей с архитектурой, представленной на рис. 1, строгое доказательство сходимости отсутствует и является весьма сложной теоретической задачей. На наш взгляд, гораздо проще решать эту задачу от противного, т.е. доказать, что существуют такие наборы параметров синхронизируемых сетей, при которых достижение полной синхронизации становится невозможным. В многочисленных экспериментах с сетями (рис. 1) были зафиксированы редкие реализации, в которых в течение длительного времени полная синхронизация так и не наступила. Детальный анализ эволюционирующих величин сетей показал, что в этих реализациях наступает эволюционный «тупик» – полная синхронизация не достигнута, а используемое правило корректировки ВК (соотношение (3)) запрещает ее проведение при любых дальнейших наборах x_{kj}.

Пусть синхронизируемые сети состоят из одного персептрона с N входами. Согласно коррекция ВК персептронов не проводится, если $O^A \neq O^B$ или, что тоже самое, если $y_1^A \neq y_1^B$, что равносильно

$$\begin{cases} \sum_{j=1}^{N} w_{1j}^A x_{1j} \geq 0 \\ \sum_{j=1}^{N} w_{1j}^B x_{1j} < 0 \end{cases} \text{ или } \begin{cases} \sum_{j=1}^{N} w_{1j}^A x_{1j} < 0 \\ \sum_{j=1}^{N} w_{1j}^B x_{1j} \geq 0. \end{cases} \quad (5)$$

Если существуют такие значения и w_{1j}^B, при которых независимо от значений x_{1j} хотя бы одна из систем неравенств (5) справедлива, то коррекция ВК производится не будет, таким образом возникает указанный выше «тупик». Наборы x_{1j} обновляются, а ВК остаются неизменными т.е. несогласованными. Найдем частное решение (5). Например, пусть:

$$w_{1j}^A = w_{1j}^B = 0, \; j=2,3,\ldots,N. \quad (6)$$

Тогда (5) примет вид

$$\begin{cases} w_{11}^A x_{11} \geq 0 \\ w_{11}^B x_{11} < 0. \end{cases} \text{ или } \begin{cases} w_{11}^A x_{11} < 0 \\ w_{11}^B x_{11} \geq 0. \end{cases} \quad (7)$$

Для любых возможных значений x_{11} (± 1) одна из систем (7) будет справедлива, если имеет место

$$w_{11}^A * w_{11}^B < 0. \quad (8)$$

Таким образом, «тупик» возникает для значений ВК, определяемых (7) и (8). Эти значения весовые коэффициенты могут принять как при генерации начальных значений, так и в процессе синхронизации.

Найдем вероятность возникновения «тупика» при генерации начальных значений ВК. Это сложное событие, состоящее из суммы N событий, каждое из которых заключается в том, что один из ВК отличен от нуля, а остальные равны нулю. Вероятность этого события равна

$$P = \sum_{j=1}^{N} P_j,$$

где P_j – вероятность того, что j-тый весовой коэффициент не равен 0, а остальные равны 0.

Вероятность того, что $w_{11}^A \neq 0, w_{11}^B \neq 0$, а $w_{1j}^A = w_{1j}^B = 0$, $j=2,3,\ldots,N$, равна

$$P_1 = P\left(w_{12}^A = 0, w_{13}^A = 0, \ldots, w_{1N}^A = 0, w_{11}^A \geq 0, w_{12}^B = 0, w_{13}^B = 0, \ldots, w_{1N}^B = 0, w_{11}^B < 0\right) + P\left(w_{12}^A = 0, w_{13}^A = 0, \ldots, w_{1N}^A = 0, w_{11}^A < 0, w_{12}^B = 0, w_{13}^B = 0, \ldots, w_{1N}^B = 0, w_{11}^B \geq 0\right).$$

Поскольку все перечисленные события независимы, то

$$P_1 = P\left(w_{12}^A = 0\right)P\left(w_{13}^A = 0\right)\ldots P\left(w_{1N}^A = 0\right)P\left(w_{11}^A \geq 0\right)P\left(w_{12}^B = 0\right)P\left(w_{13}^B = 0\right)\ldots P\left(w_{1N}^B = 0\right)P\left(w_{11}^B < 0\right) + P\left(w_{12}^A = 0\right)P\left(w_{13}^A = 0\right)\ldots P\left(w_{1N}^A = 0\right)P\left(w_{11}^A < 0\right)P\left(w_{12}^B = 0\right)P\left(w_{13}^B = 0\right)\ldots P\left(w_{1N}^B = 0\right)P(w_{11}^B \geq 0).$$

Так как для ИНС, состоящей из одного персептрона, каждый ВК может принять d значений, где $d = 2L+1$, то

$$P\left(w_{1j}^{A/B} = 0\right) = d^{-1}, \; P\left(w_{11}^{A/B} \geq 0\right) = (L+1)d^{-1},$$

$$P\left(w_{11}^{\frac{A}{B}} < 0\right) = Ld^{-1},$$

и окончательно имеем

$$P_1 = 2L(L+1)(2L+1)^{-2N}.$$

Аналогично найдем вероятность P_2. Это вероятность аналогичного сложного события при котором все ВК персептрона равны 0 за исключением второго. Аналогично находим P_3, P_4, ..., P_N. Окончательно вероятность искомого «тупика» равна

$$P = \sum_{j=1}^{N} P_j = 2NL(L+1)(2L+1)^{-2N}. \qquad (9)$$

Для ИНС, состоящей из K персептронов, где K – нечетное число, несложно найти тупиковые значения ВК. Пусть ВК первого персептрона каждой сети выбраны согласно (6), (8), что обеспечивает $y_1^A \neq y_1^B$ при любых наборах x_{1j}, следовательно, для остальных К-1 персептронов необходимо обеспечить $y_i^A = y_i^B$, где $i = 2,3,\ldots,K$. В этом случае $O^A \neq O^B$ и коррекция ВК не должна проводиться. Для этого можно задать, например, следующие значения ВК

$$w_{ij}^A = w_{ij}^B, \quad i = 2,3,\ldots,K \;;\; j=1,2,\ldots,N.$$

Рассуждая аналогично, можно найти значения ВК, при которых не будет осуществляться их коррекция и при четном количестве персептронов.

Определение момента наступления полной синхронизации

Экспериментальное исследование, приведенное с помощью имитационной модели, показало, что число тактов синхронизации t_c, необходимое для наступления равенства ВК, является случайной величиной с законом распределения, зависящим от параметров сетей – *L, K, n* (рис. 2).

Как видно из рис. 2, величина t_c изменяется от 0 до нескольких тысяч тактов. Действительно, $t_c = 0$, если векторы $W^A(0), W^B(0)$ изначально оказались равными. Естественно, вероятность этого события очень мала. Например, если длина векторов в битах равна *d*, то вероятность равна $P(W^A(0) = W^B(0)) = 1/2^d$.

Неопределенность относительно наступления полной синхронизации ВК для абонентов приводит к тому, что процесс синхронизации может продолжаться уже после выравнивания ВК. Это предоставляет злоумышленнику время на реализацию одной из возможных атак на формируемые ключи [4]. К сожале-

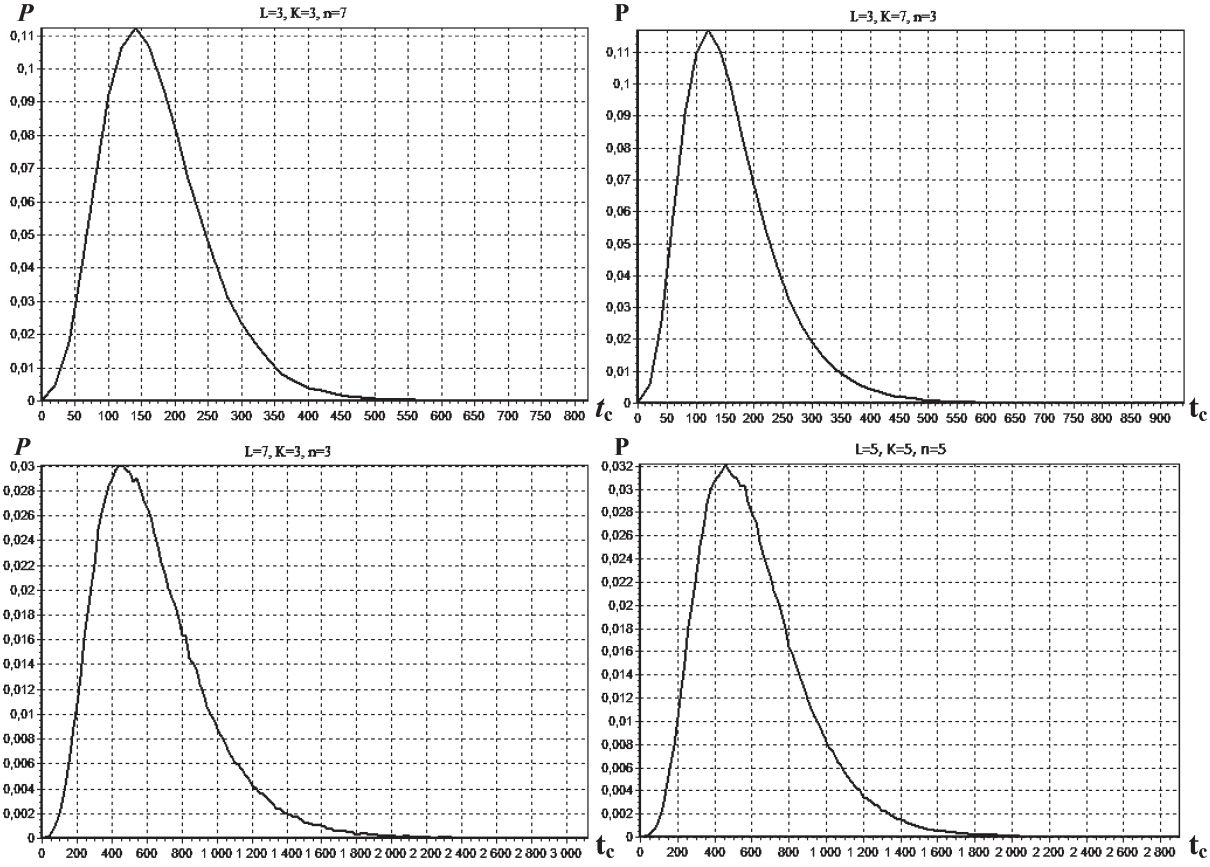

Рис. 2. Закон распределения числа тактов синхронизации, необходимых для наступления равенства ВК

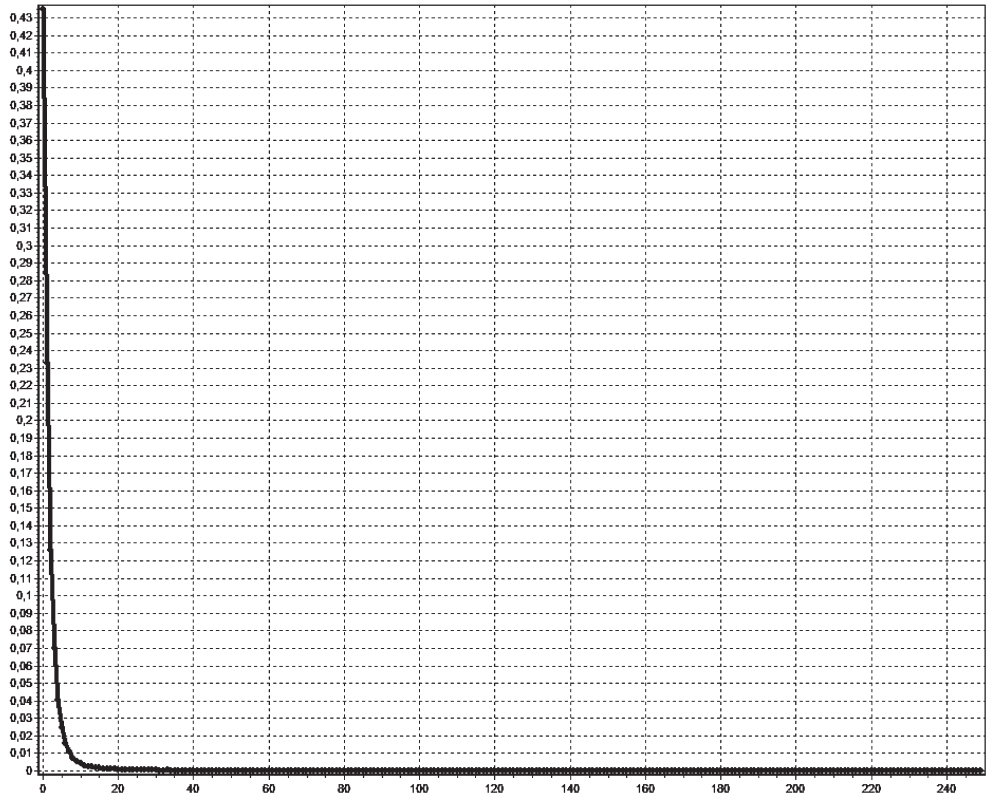

Рис. 3. Закон распределения *t*

нию, в известных нам работах вопрос об остановке процесса синхронизации ВК не рассматривается.

Для решения этой задачи возможен следующий подход. В процессе синхронизации пока $W^A(i) \neq W^B(i)$, наблюдаются такты, в которых $O^A(i) \neq O^B(i)$, и такты, в которых $O^A(i) = O^B(i)$. Последние могут образовывать непрерывные последовательности и свидетельствовать о наступлении полной синхронизации с некоторой вероятностью. Действительно, как показали эксперименты, довольно часто встречаются отрезки тактов длиной до нескольких сотен и более совпадений $O^A(i) = O^B(i)$, однако синхронизация еще не достигнута. Введем в рассмотрение случайную величину t – длину непрерывной последовательности тактов, в которых $O^A(i) = O^B(i)$, но полной синхронизации не наступило. Моделирование показывает, что эта величина изменяется от 2 до нескольких сотен. Огибающая закона распределения этой величины представлен на рис. 3.

В качестве предельно значения *t*, при котором возможно наступила полная синхронизация, можно выбрать квантиль распределения уровня $\mathcal{E}$ случайной величины *t*. Обозначим эту величину $t_{\text{синхр}}$. Таким образом, вероятность того, что случайная величина t не превысит $t_{\text{синхр}}$, должна быть не менее $\mathcal{E}$

$$\sum_{t=2}^{t_{\text{синхр}}} P(t) \geq \mathcal{E}. \qquad (10)$$

Решая (10) относительно $t_{\text{синхр}}$, получим пороговое значение для принятия решения об остановке процесса синхронизации.

Рассмотрим пример. Пусть синхронизируемые ИНС имеют параметры: $L = 5$, $K = 5$, $N = 5$. Зададимся двумя уровнями квантилей $\varepsilon_1 = 0,95, \varepsilon_2 = 0,01$ и вычислим значения $t_{\text{синх1}}$ и $t_{\text{синх2}}$, соответствующие этим уровням. Для этого промоделируем процесс синхронизации и построим закон распределения длин отрезков *t* для сетей с заданными параметрами $P(t)$ (рис. 4). При этом оказалось, что: $t_{\text{синх1}} = 61$, $t_{\text{синх2}} = 82$. Т.е. чтобы обнаружить наступление полной синхронизации ИНС с вероятностью ошибки 5% или 1% нужно остановить процесс синхронизации, когда достигнута длина отрезка тактов, в которых $O^A(i) = O^B(i)$, равная 61 или 82 соответственно.

Поскольку предлагаемый способ носит вероятностный характер, то в некоторых случаях процесс формирования общего ключа будет

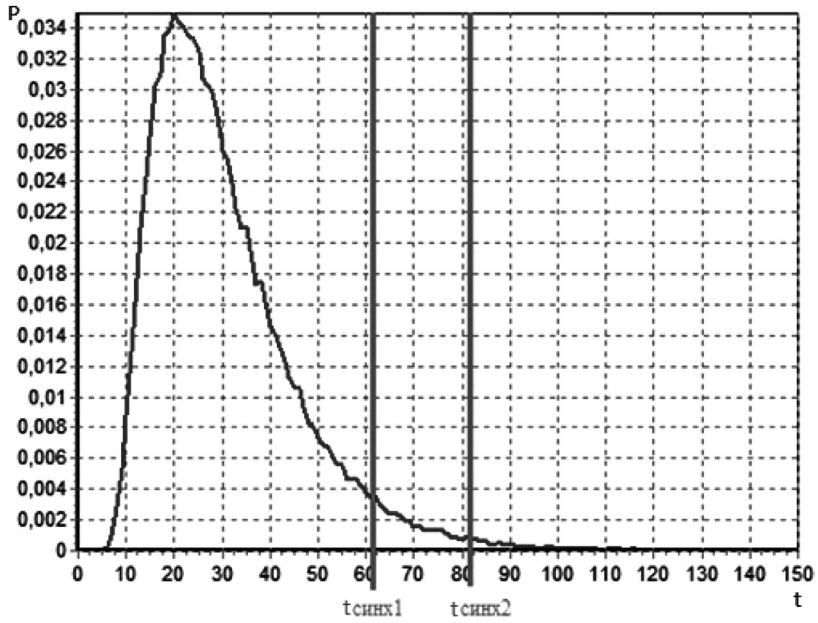

Рис. 4. Определение момента наступления полной синхронизации

остановлен слишком рано, и абоненты будут иметь несовпадающие ключевые последовательности. Поэтому предлагаемый метод может быть дополнен механизмом проверки $W^A(i) = W^B(i)$. Так как значения $W^A(i)$, $W^B(i)$ являются секретными, и обмениваться ими в открытом виде нельзя, то предлагается обмениваться значениями некоторой функции $f(W^{A/B}(i))$, которая должна обладать следующими свойствами:

1. Если $f(W^A(i)) = f(W^B(i))$, то $W^A(i) = W^B(i)$.

2. Зная $f(W^{A/B}(i))$, вычисление $W^{A/B}(i)$ представляет собой задачу огромной вычислительной сложности.

3. Зная $W^{A/B}(i)$, относительно легко вычисляется $f(W^{A/B}(i))$.

Этим условиям наиболее полно соответствуют функции, относящиеся к классу хэш-функций и широко используемые в криптографии. Они, как правило, стандартизованы и хорошо исследованы.

Литература

1. **Kanter, I.** The Theory of Neural Networks and Cryptography, Quantum Computers and Computing / I. Kanter, W.Kinzel. − 2005. Vol. 5, n. 1. – P. 130−140.

2. **Kinzel, W.** Neural Cryptography / W.Kinzel, / I. Kanter // 9th International Conference on Neural Information Processing, Singapore, 2002.

3. **Голиков, В. Ф.** Механизм синхронизации весовых коэффициентов в искусственных нейронных сетях Кинцеля и проблемы безопасности / Н.В. Брич, В.Ф. Голиков // Электроника ИНФО. – №6(96). – С.185-188.

4. **Ruttor, A.** Dynamics of neural cryptography / A. Ruttor, I. Kanter, and W. Kinzel // Phys. Rev. E, 75(5):056104, 2007.