О некоторых проблемах в задачах распределения криптографических ключей с помощью искусственных нейронных сетей

Голиков В.Ф., Брич Н.В., Пивоваров В. Л.

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 (рисунок 1).

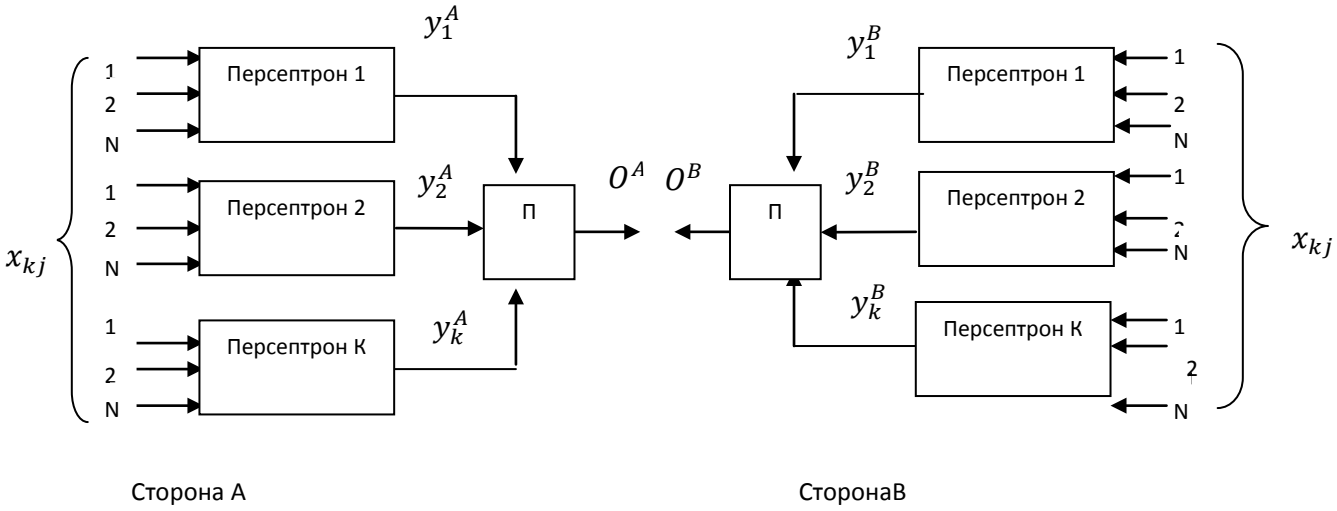

Рисунок 1 - Синхронизируемые ИНС

Архитектура на стороне отправителя и получателя представляет собой двуслойный персептрон (TPM-архитектура), состоящий из K внутренних персептронов, каждый из которых имеет N входов. Значения дискретной входной величины с равномерным распределением обозначено как $x_{kj} = \pm 1$, где $k = 1, 2, \dots, K$, $j = 1, 2, \dots, N$. Значение на выходе k-го внутреннего персептрона отправителя (получателя) обозначено как $y_k^{A/B}$.

Индекс A/Bозначает, 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, то векторы весов тех персептронов, для которых $(O^{A/B} \cdot y_k^{A/B}) > 0$, производят шаг в следующем направлении:

$$w_{kj}^{A/B}[i+1] = w_{kj}^{A/B}[i] + x_{kj}[i] \cdot O^{A/B}[i]. \quad (3)$$

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 $O^{A/B}$.

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 ВК персептронов W^A, W^B, 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 ВК).

Сходимость процесса. В [5] была доказана теорема о том, что элементарный персептрон, обучаемый по методу коррекции ошибки (с квантованием или без него), независимо от начального состояния весовых коэффициентов и последовательности появления стимулов всегда приведет к достижению решения за конечный промежуток времени.

Действительно, набор возможных значений весов конечен, поскольку весовые коэффициенты принадлежат конечному дискретному множеству ($|w_{kj}^{A/B}| \leq L$). В результате серии итераций сети придут к состоянию полного синхронизма, т.е. полная синхронизация происходит за конечное число шагов, поскольку веса нормируются на каждом шаге. Теорема не распространяется на случай многослойных персептронов - в этом случае о сходимости процесса синхронизации достоверно можно говорить лишь в том случае, если веса корректируются для персептрона последнего слоя. Для сетей с архитектурой, представленной на рисунке 1, авторам удалось найти, по крайней мере, одно множество значений ВК, для которого полная синхронизация недостижима. Например, если для сетей с $K = 3$ изначально либо в процессе синхронизации установится, что выходные величины двух внутренних персептронов равны: $y_1^A = y_1^B$, $y_2^A = y_2^B$, а третьего - противоположны $y_3^A = -y_3^B$, то, независимо от входных значений x_{kj}, сети остаются с несогласованными ВК бесконечно долго.

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* (рисунок 2).

Как видно из рисунка 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 $P(W^A(0) = W^B(0))=1/2^d$.

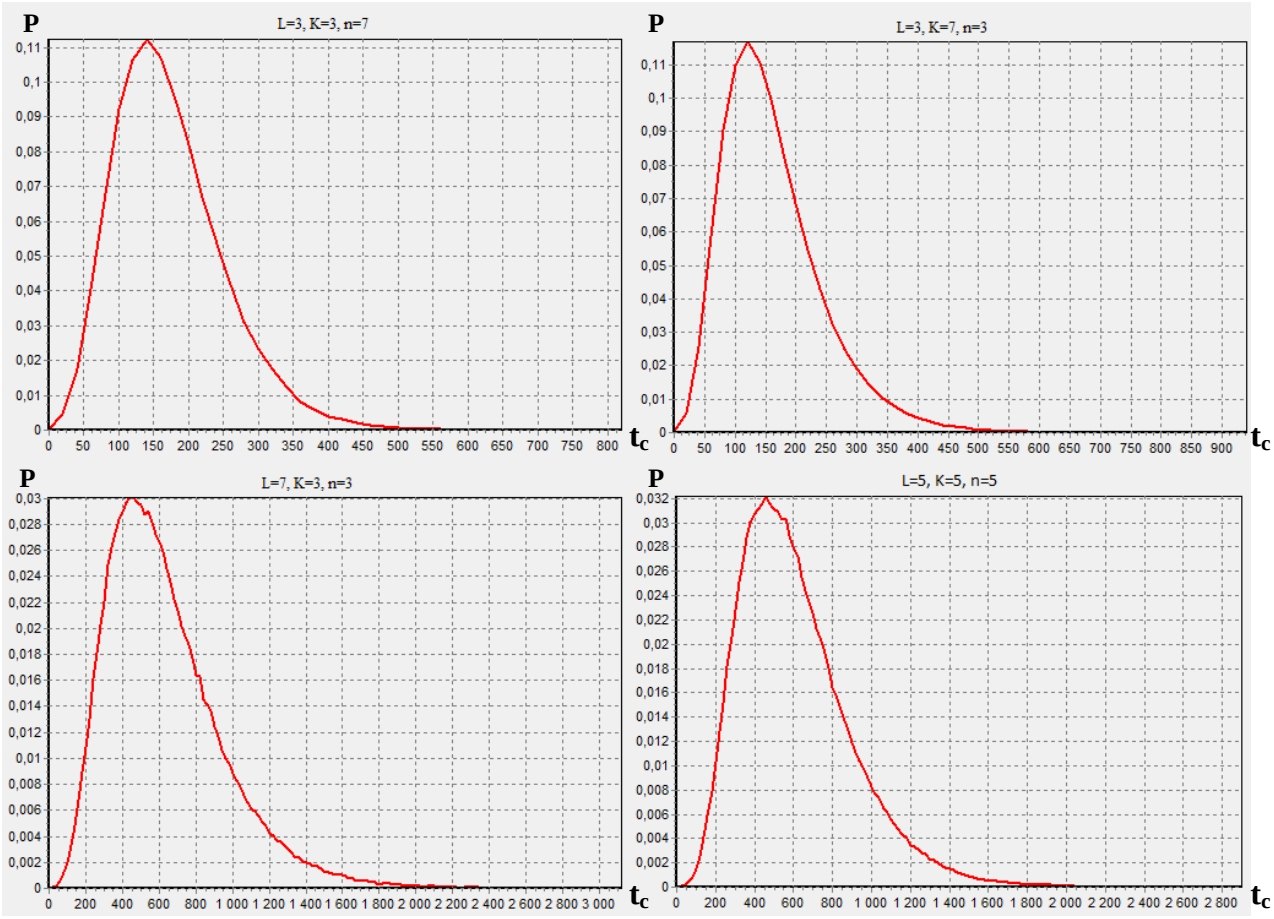

Рисунок 2 – Закон распределения числа тактов синхронизации, необходимых для наступления равенства ВК

Неопределенность абонентов относительно наступления полной синхронизации ВК приводит к тому, что процесс синхронизации может продолжаться уже после выравнивания ВК. Это предоставляет злоумышленнику время на реализцию одной из возможных атак на формируемые ключи [4]. К сожалению, в известных нам работах вопрос об остановке процесса синхронизации ВК не рассматривается. Косвенно создается впечатление, что необходимое число тактов ориентировочно выбирается, исходя из результатов моделирования. Поэтому поиск ответов на поставленные вопросы представляется актуальным.

Исходя из принципа Керкхоффа будем считать, что криптоаналитику, прослушивающему канал связи системы до начала сеанса связи между абонентами известна архитектура сетей A и B и их параметры, за исключением $W^{A/B}(0)$, а также правила коррекции весов. Кроме того, прослушивая канал связи, криптоаналитик получает в каждом i- том такте синхронизации значения $O^{A/B}(i)$. Такой же информацией располагают и абоненты A и B. На первый взгляд, судить о степени синхронизации можно только по «поведению» значений $O^{A/B}(i)$. Действительно, при полной синхронизации сетей выполняется $W^A(i) = W^B(i)$, а, следовательно, $O^A(i) = O^B(i)$, и, наблюдая за

значениями $O^{A/B}(i)$, можно заметить, что, начиная с некоторого такта, это событие многократно повторяется (это может свидетельствовать о достижении $W^A(0) = W^B(0)$). Однако, как показали эксперименты, использование этого события в качестве индикатора синхронизации приводит к большим ошибкам. Довольно часто встречаются отрезки тактов длиной до нескольких сотен и более совпадений $O^A(i) = O^B(i)$, однако синхронизация еще не достигнута (рисунок 3). Следовательно, только на основании информации о выходных значениях сетей задачу своевременной остановки синхронизации решить не удастся.

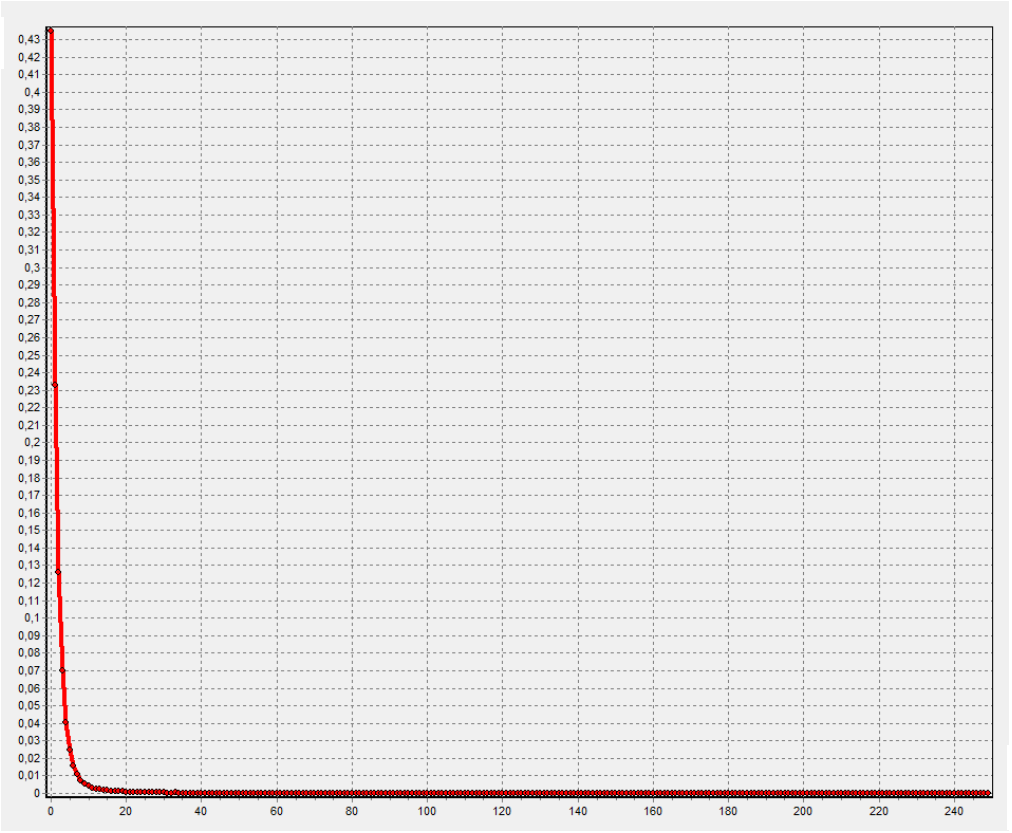

Рисунок 3 – Закон распределения длин отрезков тактов, в которых $O^A(i) = O^B(i)$

Вместе с тем каждой легитимной стороне известны значения внутренних персептронов своей (и только своей ИНС) и их ВК. Эти значения не известны другому абоненту и криптоаналитику. ВК сетей должны храниться в тайне от криптоаналитика, поэтому использовать их в открытом виде, обмениваясь значениями по открытому каналу связи, нельзя. Предлагается обмениваться значениями некоторой функции$f\left(W^{A/B}(i)\right)$, которая должна обладать следующими свойствами:

1. Если $f\left(W^A(i)\right) = f\left(W^B(i)\right)$, то $W^A(i) = W^B(i)$.
2. Зная $f\left(W^{A/B}(i)\right)$, вычисление $W^{A/B}(i)$ представляет собой задачу огромной вычислительной сложности.
3. Зная $W^{A/B}(i)$, относительно легко вычисляется $f\left(W^{A/B}(i)\right)$.

Этим условиям наиболее полно соответствуют функции, относящиеся к классу хэш-функций и широко используемые в криптографии. Они, как правило, стандартизованы и хорошо исследованы. Однако, если выполнение первых двух требований не вызывает сомнения, то многократное хэширование ВК в процессе синхронизации может существенно замедлить его процесс (с учетом того, что количество тактов синхронизации может достигать десятков тысяч). Поэтому актуальной является задача рационального подхода в выборе и применении хэш-функций, что позволит сократить время вычисления одной итерации и количество итераций. Например, можно назначить периодичность хэширования либо соединить контроль по значениям $f\left(W^{A/B}(i)\right)$ и $O^{A/B}(i)$ и т.д. Выбор наилучшего варианта составляет предмет дальнейших исследований.

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.

5. Розенблатт, Ф. Принципы нейродинамики: перцептроны и теория механизмов мозга. - М.: «Мир», 1965.