

УДК 621.391.25

АТАКА НА СИНХРОНИЗИРУЕМЫЕ ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ, ФОРМИРУЮЩИЕ ОБЩИЙ СЕКРЕТ, МЕТОДОМ ОТЛОЖЕННОГО ПЕРЕБОРА

В.Ф. ГОЛИКОВ, А.Ю. КСЕНЕВИЧ

Белорусский национальный технический университет, Республика Беларусь

Поступила в редакцию 19 октября 2017

Аннотация. Рассматриваются применение синхронизируемых искусственных нейронных сетей для генерации общего криптографического ключа, а также атака методом отложенного перебора на эти сети с целью получения этого ключа.

Ключевые слова: искусственная нейронная сеть, криптографический ключ, синхронизация.

Abstract. The application of synchronized artificial neural networks for generation of a common cryptographic key and an attack by deferred search on these networks with the purpose of obtaining this key are considered.

Keywords: artificial neural network, cryptographic key, synchronization.

Doklady BGUIR. 2017, Vol. 110, No. 8, pp. 48-53

Attack on synchronized artificial neural networks,

forming a common secret by deferred search

V.F. Golikov, A.Y. Ksenevich

Введение

Безопасная пересылка данных является одним из самых важных вопросов в современном информационном мире. Ключевую роль в этой сфере играет криптография, позволяющая шифровать важную информацию, защищая тем самым ее от несанкционированного доступа.

Безопасность криптографических систем зависит от трудности решения проблем теории чисел (например, проблема факторизации, проблема дискретного логарифмирования). Однако рост вычислительной мощности современных компьютеров требует применения все более длинных целых чисел, составляющих параметры криптографических систем. С другой стороны, использование все более длинных целых чисел ведет к росту вычислительной сложности криптографических алгоритмов. Поэтому так интересны любые новаторские криптографические системы, не использующие теорию чисел. Одной из таких новых идей является применение нейронных сетей для формирования общего криптографического ключа при обмене информацией по незащищенным каналам связи.

Общая схема синхронизируемых искусственных нейронных сетей

В работах [1–3] предложен способ формирования общего криптографического ключа с помощью двух искусственных нейронных сетей (ИНС), соединенных открытым каналом связи и синхронизируемых общими случайными воздействиями.

Наименее доказанным в этом способе, на наш взгляд, является стойкость формируемого ключа к возможным атакам на него со стороны третьей стороны, «прослушивающей» канал связи, по которому синхронизируемые сети обмениваются информацией. Это объясняется отсутствием строгих математических моделей, адекватно описывающих процесс синхронизации, ввиду высокой сложности процесса изменения весовых коэффициентов сетей как дискретных временных рядов.

Анализ процессов, происходящих при этом, возникающие проблемы и возможные атаки третьей стороны описаны в [4–7]. В этих работах рассматриваются атаки, организованные следующим образом.

Если обозначить сети абонентов, формирующих общий криптографический ключ, A и B , а третьего абонента, тайно пытающегося узнать этот ключ, – E , то схема взаимодействия сетей абонентов представляется, как на рис. 1.

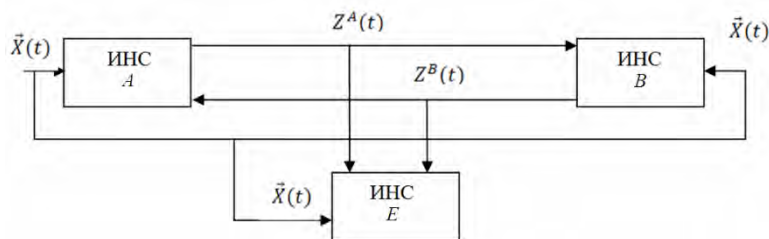


Рис. 1. Схема взаимодействия сетей абонентов

На рис. 1 t – номер такта синхронизации, $\tilde{X}(t)$ – вектор синхронизирующих случайных воздействий, $Z^A(t), Z^B(t)$ – выходные величины сетей A и B соответственно. Архитектура и параметры всех сетей идентичны (рис. 2). На рис. 2 каждый персептрон имеет n входов, на каждый из которых поступает случайное число $x_{ij}(t) \in [-1, 1]$ (одна из компонент $\tilde{X}(t), j = 1, 2, \dots, n; i = 1, 2, \dots, K$). Каждый персептрон описывается вектором весовых коэффициентов $\vec{W}_i(t)$ с компонентами $w_{ij}(t) \in [-L, L]$, где L – целое положительное число.

Выходные величины персептронов $Y_i(t) \in [-1, 1]$ перемножаются и образуют выходы сетей $Z(t) \in [-1, 1]$.

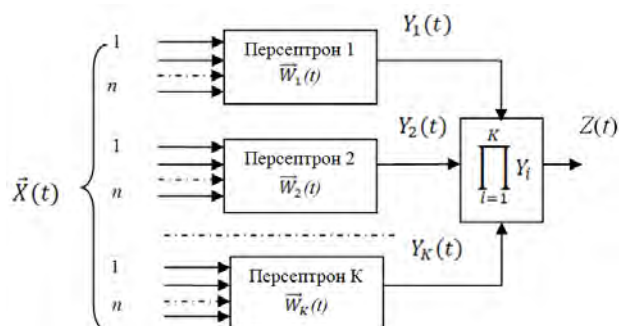


Рис. 2. Архитектура сети

Начальные значения весовых коэффициентов персептронов сетей A и B $\vec{W}^A(0), \vec{W}^B(0)$ выбираются абонентами случайно, независимо друг от друга и сохраняются в секрете [6]. Подавая синхронно на входы своих сетей вектор $\tilde{X}(t)$, абоненты A и B вычисляют выходные величины сетей $Z^A(t)$ и $Z^B(t)$, обмениваются ими и корректируют значения весовых коэффициентов персептронов своих сетей таким образом, что через некоторое число тактов t_{AB} наступает равенство

$$\vec{W}^A(t) = \vec{W}^B(t). \quad (1)$$

Атакующая сеть E , используя значения $\tilde{X}(t)$, вычисляет $Z^E(t)$ и сравнивает его с перехваченными $Z^A(t)$ и $Z^B(t)$, корректирует значения весовых коэффициентов персептронов своей сети по определенному алгоритму и через некоторое число тактов t_{AB} добивается равенства

$$\vec{W}^E(t) = \vec{W}^A(t) = \vec{W}^B(t). \quad (2)$$

В зависимости от выбранного абонентом E алгоритма коррекции различают несколько видов атак. Наиболее эффективной считается «геометрическая атака» [4]. Исследования показали, что независимо от вида атаки обеспечивается

$$P(t_{AB} \leq d) > P(t_{AE} \leq d), \quad (3)$$

где d – назначенное сторонами A и B предельное число тактов, достаточное для наступления полного синхронизма их сетей. Выражение (3), однако, совсем не означает, что в процессе атаки обязательно произойдет событие $t_{AB} \leq d \leq t_{AE}$. То есть могут иметь место успешные

атаки, при которых окажется $t_{AE} \leq d, t_{AB} \leq d$, что приведет к выполнению (2). С наличием таких реализаций и связаны основные сомнения в безопасности анализируемого метода открытого формирования общего секрета. На рис. 3 приведены зависимости $P(t_{AB} \leq d), P(t_{AE} \leq d)$ от d при $n = 1000, K = 3, L = 8$, полученные методом имитационного моделирования.

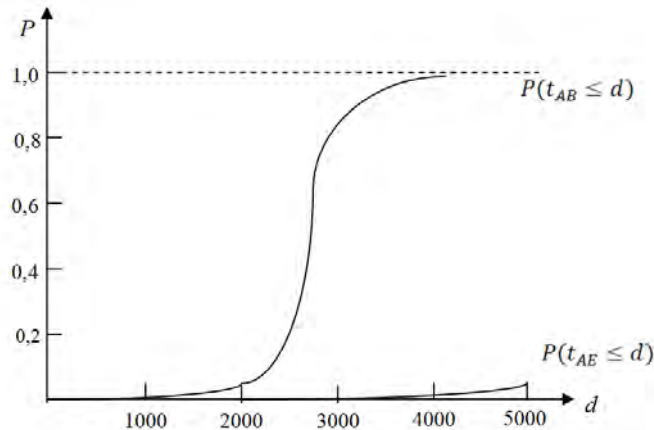


Рис. 3. Зависимости вероятностей синхронизации

Из приведенных графиков видно, что для ИНС с выбранными параметрами для достижения синхронизма с вероятностью не менее 0,95 следует выбрать $d \approx 3500$. При этом $P(t_{AE} \leq d) \approx 0,045$. То есть в четырех процентах реализаций атакующая сеть E успевает войти в синхронизм с сетью A за отведенное число тактов аналогично сети B . Однако в [4] показано, что вероятность успешной атаки можно снизить за счет увеличения параметра L . Например, при $n = 1000, K = 3, L = 57$ удается обеспечить $P(t_{AE} \leq d) \approx 10^{-4}$. Но при этом для того, чтобы обеспечить синхронизм защищаемых сетей, необходимо назначить $d = 1,6 \cdot 10^5$. С практической точки зрения столь длительный процесс обмена данными делает анализируемый метод бесперспективным, тем более, что достигаемая малая величина вероятности успешной атаки отнюдь не является безопасной.

Атака методом отложенного перебора

Проанализировав используемый процесс синхронизации сетей, можно предложить атаку, эффективную практически при любых параметрах сетей.

Как уже указывалось выше, в процессе синхронизации сетей за счет общих входных воздействий и специального алгоритма коррекции происходит постепенное сближение значений весовых коэффициентов персептронов сетей. Количество тактов синхронизации, за которое наступит равенство весов, зависит от их начальных значений и от последовательности $\vec{X}(t)$. Общепринято считать, что атака абонента E методом перебора начальных значений весовых коэффициентов персептронов своей сети обречена на неудачу, так как количество этих значений даже для сети с небольшими значениями n, K, L требует очень больших временных затрат. Можно показать [6], что количество возможных начальных значений вектора весовых коэффициентов персептронов сети не менее $M = (2L + 1)^{nK}$. Поэтому, например, для формирования двоичного общего секретного числа размером 256 бит нужно выбрать $L = 4, nK = 84$, при этом получим $M = 9^{84} \approx 1,43 \cdot 10^{80} \approx 2,145 \cdot 2^{380}$. То есть абоненту E следует создать M ИНС и попытаться синхронизировать их с сетями A и B . При указанных выше количествах подобная задача не реализуема. Однако изучение процесса синхронизации сетей показало, что абоненту E совершенно необязательно угадать истинное значение $\vec{W}^A(0)$ или $\vec{W}^B(0)$, так как существует достаточно большое множество начальных значений вектора весовых коэффициентов $\vec{W}^E(0)$, движение из которых при благоприятных траекториях $\vec{X}(t)$ позволяет обеспечить $t_{AB} \leq t_{AE} \leq d$ (рис. 4). На этом рисунке показаны условные траектории изменения векторов весовых коэффициентов сетей A, B в процессе синхронизации. Сеть E представлена тремя траекториями. Траектории $\vec{W}^A(t)$ и $\vec{W}^B(t)$ совпали при $t = d - 2$, траектории $\vec{W}^A(t)$ и $\vec{W}_1^E(t)$ не совпали за

назначенное время синхронизации d , траектории $\vec{W}^A(t)$ и $\vec{W}_2^E(t), \vec{W}_3^E(t)$ совпали, причем время достижения синхронизации второй сети E меньше, чем у сетей A и B .

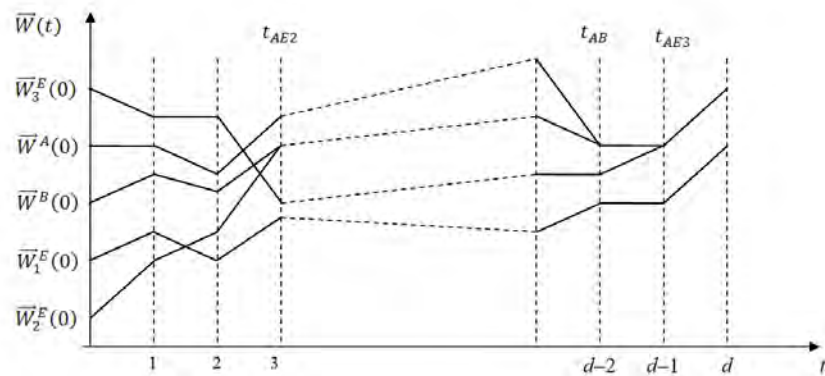


Рис. 4. Условные траектории изменения векторов весовых коэффициентов

Кроме того, абоненту E нет необходимости строить модель, состоящую из M ИНС, используя для этого огромные вычислительные ресурсы. Атаку на сформированное A и B общее число можно организовать, располагая одной или относительно небольшим числом ИНС, в «отложенном» по времени режиме. Для этого абонент E , прослушивая канал связи между A и B , запоминает значения $\vec{X}(t), Z^A(t), Z^B(t), d$. Затем случайным образом генерирует $\vec{W}_1^E(0)$, используя $\vec{X}(1)$, формирует $Z^E(1)$ и сравнивает его с $Z^A(1)$, проводит коррекцию $\vec{W}_1^E(0) \rightarrow \vec{W}_1^E(1)$ в соответствии с выбранным алгоритмом. Этот процесс продолжается такт за тактом.

Если к моменту назначенного конечного такта $t = d$ устанавливается факт наступления синхронизации по одному из принятых критериев [7], то E принимается решение о совпадении $\vec{W}_1^E(d) = \vec{W}_1^A(d) = \vec{W}_1^B(d)$. В противном случае следует выбрать другое значение $\vec{W}^E(0)$ и вновь провести синхронизацию. Процесс повторяется до первой успеха.

Оценим необходимый объем отложенного моделирования. Обозначим вероятность успешной синхронизации абонента E в одной попытке P_{AE}^1 . С учетом предыдущих обозначений $P_{AE}^1 = P(t_{AE} < d)$. Вероятность того, что в m попытках событие $t_{AE} < d$ произойдет не менее одного раза, равна

$$P_m(i \geq 1) = 1 - P_m(i = 0) = 1 - (1 - P_{AE}^1)^m. \quad (4)$$

Потребуем, чтобы эта вероятность была не менее заданной γ . Тогда

$$1 - (1 - P_{AE}^1)^m \geq \gamma. \quad (5)$$

Из (5) найдем

$$m = \frac{\ln(1 - \gamma)}{\ln(1 - P_{AE}^1)}. \quad (6)$$

Используя (6), можно оценить объем отложенного моделирования, например, для выше рассмотренных данных, в которых $P_{AE}^1 = 10^{-4}$. Зададимся $\gamma = 0,98$, получим

$$m = \frac{\ln(1 - 0,98)}{\ln(1 - 10^{-4})} \approx 3,9 \cdot 10^4.$$

Такое количество экспериментов легко реализуется за очень короткое время. Из (6) несложно вывести значение P_{AE}^1 , при котором объем отложенного моделирования будет соизмерим с объемом перебора значений ключа длиной 256 бит симметричного алгоритма шифрования, то есть сложность вычисления атакуемого ключа соизмерима со сложностью взлома симметричного алгоритма полным перебором. Для этого упростим (6). Известно, что при $x < 1$ справедливо $\ln(1 - x) \approx -x$, т. е. $\ln(1 - P_{AE}^1) \approx -P_{AE}^1$. С учетом этого из (6) получим $P_{AE}^1 = -\frac{\ln(1 - \gamma)}{m}$. При $m = 2^{256}$ имеем $P_{AE}^1 = 3,9 \cdot 2^{-256}$. Такое значение вероятности P_{AE}^1 практически невозможно обеспечить подбором параметров ИНС.

Заклучение

Проведенный анализ позволяет сделать вывод о высокой эффективности атаки с отложенным перебором. Противостояние такой атаке является весьма затруднительным. Следовательно, вопрос о повышении стойкости формируемого ключа к возможным атакам на него с третьей стороны остается открытым и требует дальнейшего исследования, как и вопрос о более эффективных атаках.

Список литературы

1. Kanter I., Kinzel W. The Theory of Neural Networks and Cryptography, Quantum Computers and Computing. 2005. Vol. 5, № 1. P. 130–140.
2. Kinzel W., Kanter I. Neural Cryptography // 9th International Conference on Neural Information Processing. Singapore, 2002.
3. Kanter I., Kinzel W., Kanter E. Secure exchange of information by synchronization of neural networks // arxiv: cond/0202112v1 [cond-mat.stat-mech]. 2002.
4. Ruttor A., Kanter I., Kinzel W. Dynamics of neural cryptography // Phys. Rev. E. 2007. 75(5):056104.
5. Голиков, В.Ф., Брич Н.В. Механизм синхронизации весовых коэффициентов в искусственных нейронных сетях Кинцеля и проблемы безопасности // Электроника ИНФО. № 6 (96). С. 185–188.
6. Голиков В.Ф., Брич Н.В. Вероятностные свойства начальных значений весовых коэффициентов в синхронизируемых искусственных нейронных сетях Кинцеля // Системный анализ и прикладная информатика. 2013. № 1–2. С. 33–37.
7. Голиков В.Ф., Брич Н.В., Пивоваров В.Л. О некоторых проблемах в задачах распределения криптографических ключей с помощью искусственных нейронных сетей // Системный анализ и прикладная информатика. 2014. № 1–3. С. 42–46.

References

1. Kanter I., Kinzel W. The Theory of Neural Networks and Cryptography, Quantum Computers and Computing. 2005. Vol. 5, № 1. P. 130–140.
2. Kinzel W., Kanter I. Neural Cryptography // 9th International Conference on Neural Information Processing. Singapore, 2002.
3. Kanter I., Kinzel W., Kanter E. Secure exchange of information by synchronization of neural networks // arxiv: cond/0202112v1 [cond-mat.stat-mech]. 2002.
4. Ruttor A., Kanter I., Kinzel W. Dynamics of neural cryptography // Phys. Rev. E. 2007. 75(5):056104.
5. Golikov, V.F., Brich N.V. Mehanizm sinhronizacii vesovyh kojefficientov v iskusstvennyh nejronnyh setjah Kincelja i problemy bezopasnosti // Jelektronika INFO. № 6 (96). S. 185–188. (in Russ.)
6. Golikov V.F., Brich N.V. Verojatnostnye svojstva nachal'nyh znachenij vesovyh kojefficientov v sinhroniziruemyh iskusstvennyh nejronnyh setjah Kincelja // Sistemnyj analiz i prikladnaja informatika. 2013. № 1–2. S. 33–37. (in Russ.)
7. Golikov V.F., Brich N.V., Pivovarov V.L. O nekotoryh problemah v zadachah raspredelenija kriptograficheskikh ključež s pomoshh'ju iskusstvennyh nejronnyh setej // Sistemnyj analiz i prikladnaja informatika. 2014. № 1–3. S. 42–46. (in Russ.)

Сведения об авторах

Голиков В.Ф., д.т.н., профессор, заведующий кафедрой информационных технологий в управлении Белорусского национального технического университета.

Ксеневиц А.Ю., магистрант Белорусского национального технического университета.

Information about the authors

Golikov V.F., D.Sci., professor, head of the department of information technologies in the management of Belarusian national technical university.

Ksenevich A.Y., master student of Belarusian national technical university.

Адрес для корреспонденции

220114, Республика Беларусь,
г. Минск, ул. Ф. Скорины, д. 25/3,
Белорусский национальный
технический университет
тел. +375-33-649-45-74;
e-mail: vgolikov@bntu.by
Голиков Владимир Федорович

Address for correspondence

220114, Republic of Belarus,
Minsk, F. Skaryna str., 25/3,
Belarusian national
technical university
tel. +375-33-649-45-74;
e-mail: vgolikov@bntu.by
Golikov Vladimir Fedorovich