Труды научно-технической конференции кластера пензенских предприятий,
обеспечивающих **БЕЗОПАСНОСТЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.**
Том 10, Пенза-2016, с. 15-18 (http://пниэи.рф/activity/science/BIT/T10-p15.pdf)

УДК: 519.2, 612.087, 621.319.7

Безяев А.В., Иванов А.И., Корнеев О.В.

Типовая схема защиты нейросетевых архивов биометрических данных не криптографическим хешированием через применение линейной рекурренты подсчета контрольных сумм CRC-4

В России и Казахстане [1, 2] активно ведутся работы по созданию технологии нейросетевой биометрический аутентификации личности. Требования к нейросетевым преобразователя биометрии в код доступа определяет базовый стандарт [3] их обучение осуществляется алгоритмом ГОСТ Р 52633.5 [4].

В США, Канаде, Евросоюзе проблему биометрической аутентификации личности пытаются решать через использование «нечетких экстракторов» [5, 6, 7, 8]. К сожалению, «нечеткие экстракторы» и нейросетевые преобразователи биометрия-код нельзя хранить открыто на сервере биометрической аутентификации [9, 10]. Созданы специальные средства [11], позволяющие извлекать знания из обученных нейронных сетей и «нечетких экстракторов.

Расширить область применения пока удается только для нейросетевых преобразователей биометрия-код. При этом следует использовать типовую схему защиты нейросетевых архивов, приведенную на рисунке 1.

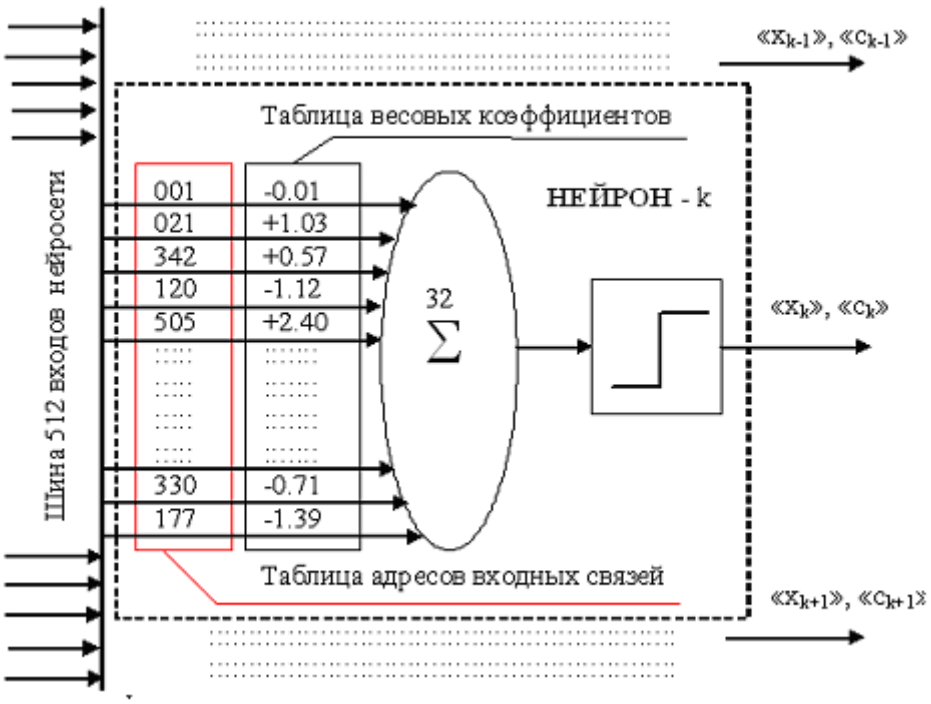

Рис. 1. Защита нейросетевых архивов хешированием данных образов «Чужой»

Защита нейросетевых архивов биометрических данных осуществляется хешированием части уже полученного выходного кода. Таблицы связей для последующих нейронов получают хешированием конкатенации соли, пароля и выходных данных предшествующих нейронов. Такая схема защиты не мешает распознаванию образа «Свой», так как его выходные коды оказываются одними и

теми же при естественной нестабильности биометрических данных. Для образов «Чужой» положение коренным образом меняется. Первая же ошибка в выходном коде или пароле приводит к эффекту размножения ошибок. Хеш-функции, из которых строятся все последующие таблицы связей нейронов, становятся случайными и не повторяют связи обученной нейронной сети на образе «Свой». Возникает эффект хеширования (перемешивания) данных образов «Чужой».

Схема защиты данных связей нейронов оказывается сильной хеширующей функцией для данных образа «Чужой» или неправильно набранного пароля. Разделить совместное хеширование пароля и биометрических данных нельзя, если все таблицы (i = 1, 2, …, 32) случайных связей каждого нейрона (9 битные адреса) получать следующим образом:

$$\begin{cases} "a_{1,i}" = hash\{соль, пароль, i\}, \\ "a_{2,i}" = hash\{соль, пароль, c_1, i\}, \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ "a_{k,i}" = hash\{соль, пароль, c_1, c_2, ..., c_{k-1}, i\} \end{cases} \quad (1),$$

где «c_i» - выходное состояние i-го нейрона, заданное при обучении нейронной сети.

Вычисление адресов по формуле (1) выполняется на этапе обучения нейронной сети. На этапе распаковки защищенного архива вместо параметров «c_i» осуществляют подстановку выходных состояний уже отработавших нейронов «c_i» для образа «Свой» или случайного состояния «x_i» для образа «Чужой».

Адреса связей нейронов полностью не совпадают с (1), если пользователь ошибся при вводе хотя бы одного символа пароля:

$$\begin{cases} "\widetilde{a}_{1,i}" = hash\{соль, парольX, i\}, \\ "\widetilde{a}_{2,i}" = hash\{соль, парольX, c_1, i\}, \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ "\widetilde{a}_{k,i}" = hash\{соль, парольX, c_1, c_2, ..., c_{k-1}, i\} \end{cases} \quad (2).$$

Если пароль скомпрометирован, то злоумышленник не может восстановить верные адреса связей из-за того, что его биометрический образ дает иное чем нужно состояние выходов у нейронной сети:

$$\begin{cases} "\hat{a}_{1,i}" = hash\{соль, пароль, i\}, \\ "\hat{a}_{2,i}" = hash\{соль, пароль, x_1, i\}, \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ "\hat{a}_{k,i}" = hash\{соль, пароль, x_1, x_2, ..., x_{k-1}, i\} \end{cases} \quad (3).$$

Очевидно, что в выражениях (1), (2), (3) может быть использована криптографическая хеш-функция, вызванная программой реализации нейросетевой защиты. Это вполне возможно, когда вычисления осуществляются под управлением операционной системы, уже имеющей встроенный криптопровайдер. Положение меняется, когда необходимо использовать маломощное аппаратно-программное средство. В этом случае применение универсального криптографического провайдера не оправдано. Полноценное криптографическое хеширование данных в выражениях (1), (2), (3) является избыточным.

В данной статье мы рассмотрим возможность применения не криптографических хеш-функций, построенных на линейной рекурренте подсчета контрольной суммы CRC-4.

Реализуем рекурренту CRC-4 с использованием регистра сдвига с обратными связями, схема которого приведена на рисунке 2.

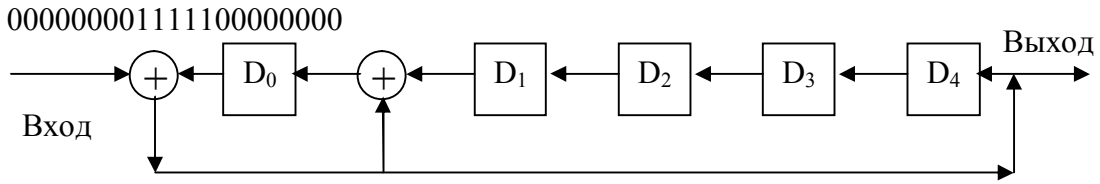

Рис. 2 Реализация рекуррентной процедуры вычисления контрольной суммы CRC-4 с порождающим многочленом $P(x) = x^4 + x^1 + x^0$

Если рекуррента CRC-4 реализована программно, то потребуется использовать 5 ячеек памяти. Далее мы можем использовать эту программу как некоторую хеш-функцию. Эта хеш-функция способна подсчитывать контрольную сумму бинарной последовательности любой длинны. При подсчетах мы будем обнулять регистры и подавать на вход рекурренты данные. Тогда на выходе последовательно будет появляться хеш-отклик. Число тактов сдвига должно быть равно числу бит свертываебмой двоичной последовательности.

Очевидно, что если подать на вход преобразователя последовательность из одних нулей «000000….00000», то на выходе регистра CRC-4 возникнет псевдослучайная последовательность с периодом $(2^4 - 1) = 15$ бит. Если с такой последовательности снимать 15 битные данные со сдвигом в 1 бит мы получим поток одинаковых адресов у всех нейронов, имеющих по 15 входов.

В предложенной схеме (рис. 1) сумматоры имеют 32 входа. 32 не кратно 15, что говорит о работоспособности схемы защиты. Если же мы в каждой из 15 входных состояний, изменим хоты бы 1 бит, то будем получать полноценную псевдослучайную последовательность без периода. Для выполнения этого условия в выражениях (1), (2), (3) в каждом такте меняется состояние «i» от состояния «00000» до состояния «11111». То есть в каждом цикле вычисления адреса входа «i» по предложенной схеме изменяется от одного до 5 состояний двоичного номера адреса. Этого более чем достаточно для получения ациклической псевдослучайной последовательности необходимой длины 9 бит по 32 раза.

Гарантией того, что любая пара разных нейронов будут иметь случайные (независимые, некоррелированные) адреса является то, что для каждого нейрона адреса вычисляются хешированием двоичной последовательности разной длины. Длина хешируемой последовательности, используемой для вычисления адресов связей для каждого следующего нейрона увеличивается на один бит.

Кроме того, дополнительной гарантией апериодичности используемой псевдостучайной последовательности является то, что состояния разрядов «$x_1,x_2,\ldots,x_k$» случайно для образа «Чужой» и не совпадает с состоянием разрядов «$c_1,c_2,\ldots,c_k$» ранее использованного в выражении (1).

Таким образом, для образа «Свой» мы имеем систему случайных адресов связей, вычисленных путем использования соотношения (1). Эти адреса являются случайными даже если конкатенация бинарных последовательностей «соль,пароль» состоит только из нулей. Любая ошибка при вводе пароля или при появлении состояния «x_i» в место верного состояния нейрона «c_i» приводит к лавинообразному размножению ошибок из-за расхождения всех последующих адресов связей нейронов.

В силу того, что последовательность данных «пароль» и «$c_1,c_2,\ldots,c_{256}$» неизвестна злоумышленнику Длинна CRC-k рекурренты не имеет значения, однако желательно чтобы период рекуррентны $(2^k - 1)$ был не кратен числу входов у нейронов. В нашем случае это условие выполняется.

То есть в типовой схеме защиты, приведенной на рисунке 1 могут быть использованы как полноценные криптографические хеш-функции, так и не криптографические хеш-функции, построенные на линейных рекуррентах CRC-1, CRC-4, CRC-5, CRC-6, …. В итоге получается своеобразная хеш-функция пароля доступа и биометрических данных человека. Архивы нейросетевых преобразователей биометрия-код уже защищенные, описанным выше способом, можно хранить не сервере биометрической аутентификации пользователей. Так же как мы безопасно храним на сервере хеш от пароля доступа, та том же сервере можно хранить защищенный архив с хешами, соли, паролей, параметров нейронной сети, заранее обученной преобразовывать биометрию пользователя в его длинный пароль доступа или его личный ключ криптографической аутентификации.

ЛИТЕРАТУРА:

1. Язов Ю.К. и др. Нейросетевая защита персональных биометрических данных. //Ю.К.Язов (редактор и автор), соавторы В.И. Волчихин, А.И. Иванов, В.А. Фунтиков, И.Г. Назаров // М.: Радиотехника, 2012 г. 157 с.
2. Ахметов Б.С., Иванов А.И., Фунтиков В.А., Безяев А.В., Малыгина Е.А. Технология использования больших нейронных сетей для преобразования нечетких биометрических данных в код ключа доступа. Монография, Казахстан, г. Алматы, ТОО «Издательство LEM», 2014 г. -144 с., находится в открытом доступе (http://portal.kazntu.kz/files/publicate/2014-06-27-11940.pdf).
3. ГОСТ Р 52633.0-2006 «Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации».
4. ГОСТ Р 52633.5-2011 «Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа».
5. Dodis Y., Reyzin L., Smith A. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy // Proc. EUROCRYPT, April 13, pages 523-540, 2004.
6. Monrose F., Reiter M., Li Q., Wetzel S. Cryptographic key generation from voice. //Proc. IEEE Symp. on Security and Privacy, pp. 202-213, 2001.
7. Ramírez-Ruiz J., Pfeiffer C., Nolazco-Flores J. Cryptographic Keys Generation Using FingerCodes. //Advances in Artificial Intelligence - IBERAMIA-SBIA 2006 (LNCS 4140), p. 178-187, 2006
8. Hao F., Anderson R., Daugman J. Crypto with Biometrics Effectively //IEEE TRANSACTIONS ON COMPUTERS, VOL. 55, NO. 9, SEPTEMBER, Page(s):1073 – 1074, 2006.
9. Иванов А.И. Нечеткие экстракторы: проблема использования в биометрии и криптографии. // Первая миля. № 1, 2015 г. с. 40-47.
10. Иванов А.И. Сопоставительный анализ показателей конкурирующих технологий биометрико-криптографической аутентификации личности. «Защита информации. ИНСАЙД» № 3 2014 г., с. 32-39.
11. Иванов А.И. Квантовые компьютеры: прошлое, настоящее, будущее. // "Защита информации. INSAID" № 2 2015 г. с.. 29-32.

Статья поступила 15.03.2016, опубликована 23.05.2016 по положительной рецензии д.т.н. Малыгина А.Ю.