Труды научно-технической конференции кластера пензенских предприятий,
обеспечивающих **БЕЗОПАСНОСТЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.**
Том 8, Пенза-2012, С 32-35 (http://пниэи.рф/activity/science/BIT/T8-p32.pdf)

СОКРАЩЕНИЕ ЗАТРАТ НА НЕЧЕТКУЮ ВЗАИМНУЮ АДРЕСАЦИЮ БИОМЕТРИЧЕСКИХ ОБРАЗОВ ЧЕРЕЗ ИСПОЛЬЗОВАНИЕ ВЗВЕШЕННОЙ МЕРЫ ХЭММИНГА

Андреев Д. Ю. (г. Пенза)

Одной из главных проблем при тестировании средств высоконадежной биометрической аутентификации является сбор и контроль тестовой базы биометрических образов.

Очевидно, что для оценки стойкости средств высоконадежной биометрико-криптографической аутентификации необходима тестовая база биометрических образов «Чужой», не только содержащая достаточное количество биометрических образов, но и равномерно заполненная ими. На практике, при сборе естественных биометрических образов у людей-доноров тестовая биометрическая база получается неравномерной – в ней наблюдаются области «сгущения» (повышенной плотности заполнения) и «разрежения» (пониженной плотности заполнения) биометрических образов. При оценке стойкости с использованием такой базы биометрических образов могут возникать значительные ошибки, связанные с попаданием в область «разрежения» – область поля биометрических образов, не представленную в базе. К примеру, автомат направленного перебора биометрических образов может не обнаружить коллизию даже в том случае, если биометрический образ «Свой» находится в этой области, что даст сильно завышенные результаты стойкости к атакам подбора. Двумерные графические интерпретации равномерно и неравномерно заполненных баз приведены на рисунке 1.

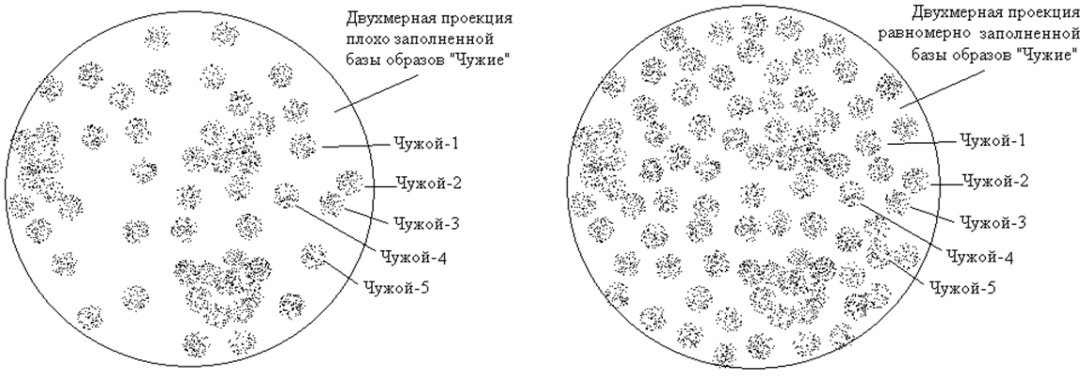

Рисунок 1 – Двумерная графическая интерпретация равномерности заполнения тестовой базы биометрических образов.

Простейшим способом обеспечения равномерности тестовой базы биометрических образов является увеличение размеров базы. Так, в случае идеальной равномерной тестовой базы минимальный ее размер, обеспечивающий достоверное тестирование, составляет $N_{min} = 1/P_2$, где P_2 – предполагаемая вероятность ошибки второго рода. При этом для реальных тестовых баз ГОСТ Р 52633.1 устанавливает необходимость 100-кратного превышения количества образов по отношению к необходимому для достоверного тестирования, т.е. $N = 100 / P_2$.

Очевидно, что при этом значительно увеличивается трудоемкость сбора и использования подобной базы. К примеру, проведение оценки стойкости средства биометрической аутентификации с защищенным биометрическим контейнером, которое возможно провести только методом полного перебора биометрических образов, может занимать до нескольких десятков лет

Более оптимальным способом обеспечения равномерности тестовой базы биометрических образов является аналитическое определение ее равномерности. Для этого необходима система адресации биометрических образов в базе, позволяющая через оценку заполнения адресного пространства тестовой базы находить области «сгущения» и «разрежения».

Для адресации биометрических образов необходимо использовать метрики биометрических образов, позволяющие определять взаимные расстояния между биометрическими образами.

Простейшим способом определения расстояния между биометрическими образами является непосредственное сравнение значений биометрических параметров.

Однако такой способ упорядочивания является крайне неэффективным: биометрические образы, используемые при высоконадежной биометрико-криптографической аутентификации, являются высокоразмерными – количество параметров одного биометрического образа составляет несколько сотен (обычно порядка 400 – 500), а учитывать все параметры одновременно становится технически невыполнимо уже для нескольких десятков параметров (так называемое «проклятие размерности»).

Единственным возможным способом избежать «проклятия размерности», описанного выше, является переход из пространства входных непрерывных высокоразмерных биометрических образов в пространство выходных дискретных кодовых откликов. При этом сортировка биометрических образов становится линейной и одномерной, а работа автомата упорядоченного перебора биометрических образов – тривиальной.

Простейшей метрикой в пространстве выходных кодовых откликов является мера Хэмминга – количество не совпавших разрядов кодовых откликов преобразователя биометрия-код на изучаемые образы:

$$h = \sum_{i=1}^{k} x_i \oplus y_i = \sum_{i=1}^{k} h_i \qquad (1)$$

где x_i – i-тый разряд первого из сравниваемых двоичных кодов, y_i – i-тый разряд второго из сравниваемых двоичных кодов, k – длина сравниваемых двоичных кодов, h_i - i-тый разряд кода Хэмминга, возникающего при вычислении расстояний Хэмминга.

По аналогии с обычной мерой Хэмминга может быть введена инверсная мера:

$$\overline{h} = \sum_{i=1}^{k} \overline{x}_i \oplus y_i = \sum_{i=1}^{k} \overline{h}_i = k - h \qquad (2).$$

Инверсную меру Хэмминга получают инвертируя один из сравниваемых кодов или инвертируя результат сложения по модулю два. Для обычных абсолютно стабильных кодов вычисление прямой и инверсной метрики Хэмминга нецелесообразно, однако этот прием выгоден при исследовании нестабильных биометрических кодов.

При помощи метрики Хэмминга обычно устанавливают меру близости между собой двух биометрических образов «Чужой», либо меру близости биометрического образа «Чужой» к биометрическому образу «Свой», для

которого было произведено обучение средства высоконадежной биометрической аутентификации.

Однако, классическая мера Хэмминга не является оптимальной метрикой при определении расстояний между биометрическими образами, так как разряды кодовых откликов преобразователя биометрия-код могут быть нестабильны, и этот факт необходимо учитывать при сравнении кодовых откликов между собой.

Рассмотрим работу преобразователя биометрия-код относительно нескольких примеров одного биометрического образа. Для каждого примера биометрического образа «Чужой» кодовые отклики будут различаться, но при этом они не будут случайными и равновероятными, так как биометрические примеры одного образа по определению должны являться очень близкими друг к другу. Не равновероятность кодовых откликов можно оценить, анализируя их поразрядно. Для равномерно распределенных кодовых откликов вероятность появления в каждом разряде значений «0» и «1» будет равна 0,5. На практике эта вероятность будет уменьшаться и увеличиваться. При этом важно не столько направление изменения вероятности, сколько сам факт такого изменения и его модуль. Определим показатель стабильности разряда кодового отклика как модуль отклонения вероятности от его математического ожидания по формуле:

$$w_i = 2 \cdot \left| P_i - 0.5 \right| \qquad (3),$$

где P_i- вероятность появления единицы в i-м разряде кодового отклика.

Показатель стабильности разрядов двоичного кода (3) позволяет более точно оценивать расстояние Хэмминга между кодом образа «Свой» и кодом образа «Чужой». Код образа «Свой» по определению всегда стабилен, то есть все показатели стабильности его разрядов единичные. Иначе обстоит дело со стабильностью разрядов образа «Чужой». Для разрядов кода «Чужой» все показатели стабильности оказываются разными, а взвешенная мера Хэмминга вычисляется следующим образом:

$$sh = \sum_{i=1}^{k} w_{yi} \cdot (x_i \oplus y_i) = \sum_{i=1}^{k} w_{yi} \cdot h_i \qquad (4),$$

где y_i - i-тый разряд кода «Чужой», w_{yi} – показатель стабильности i-го разряда кода «Чужой».

Выражение (4) является поразрядным произведением вектора показателей стабильности кода «Чужой» с кодом сравнения по Хэммингу.

В том случае, если сравниваются между собой два кода образов «Чужой-1» и «Чужой-2» ври вычислении взвешенной мерники Хэмминга приходится учитывать среднее геометрическое показателей стабильности, разрядов сравниваемых кодов:

$$sh = \sum_{i=1}^{k} \sqrt{w_{yi} \cdot w_{xi}} (x_i \oplus y_i) = \sum_{i=1}^{k} \sqrt{w_{yi} \cdot w_{xi}} \cdot h_i \qquad (5).$$

По аналогии с обычной взвешенной мерой Хэмминга целесообразно использовать дополняющую ее инверсную меру Хэмминга, которая может быть получена из выражения (5) путем введения операции инвертирования значения кодов сравнения Хэмминга:

$$\overline{sh} = \sum_{i=1}^{k} \sqrt{w_{yi} \cdot w_{xi}} (\overline{x}_i \oplus y_i) = \sum_{i=1}^{k} \sqrt{w_{yi} \cdot w_{xi}} \cdot \overline{h}_i \qquad (6).$$

Очевидно, что распределение значений взвешенной меры Хемминга (прямой или инверсной) для биометрических образов «Чужой» всегда будет меньшее, чем

аналогичные значения обычных мер Хэмминга $\overline{sh} \le \overline{h}, \quad sh \le h$. При необходимости При необходимости взвешенная метрика Хэмминга и инверсная взвешенная метрика Хэмминга могут рассматриваться как реальная и мнимая части некоторой обобщенной комплексной метрики Хэмминга:

$$sH = sh + j \cdot \overline{sh} \qquad (7).$$

Для оценки эффективности взвешенной меры Хэмминга по сравнению с классической мерой Хэмминга проведем эксперимент по сбору статистических закономерностей распределения значений этих мер. В ходе эксперимента для нейросетевого преобразователя биометрия-код с длиной кодовых откликов 256 бит были получены распределения значений метрик, представленные на рисунке 2 (в этом случае математические ожидания классической и взвешенной меры Хэмминга равны E(h) = 119 и E(sh) = 96 соответственно).

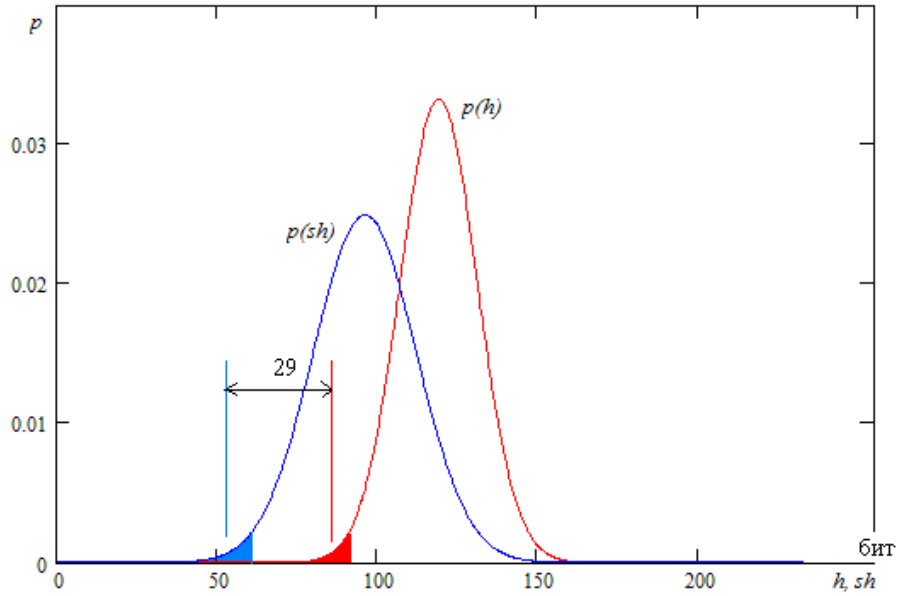

Рисунок 2 – Сравнение плотностей распределений классической и взвешенной меры Хэмминга «Свой»/«Чужой»

Как видно из рисунка 2, распределение взвешенной меры Хэмминга имеет меньшие значения математического ожидания, а ее среднеквадратическое отклонение увеличивается, что равнозначно сокращению адресуемого этой метрикой пространства. Данный эксперимент показал сокращение пространства на 29 бит, что равнозначно сокращению адресуемого пространства в $2^{29} \approx 10^9$ раз. Сокращение адресуемого пространства происходит за счет того, что «нестабильные» разряды не учитываются при вычислении метрики и происходит фактическое уменьшение значения эффективной длины кода.

В свою очередь сокращение адресного пространства позволяет упростить задачу обращения матриц нейросетевых функционалов генетическими алгоритмами подбора [1]. Так же удается увеличить стабильность результатов тестирования, выполненных по ГОСТ Р 52633.3-2011.

ЛИТЕРАТУРА:

1. Язов Ю.К. и др. Нейросетевая защита персональных биометрических данных. //Ю.К.Язов (редактор и автор), В.И. Волчихин, А.И. Иванов, В.А. Фунтиков, И.Г. Назаров // М.: Радиотехника, 2012 г. 210 с

Материалы поступили 12.02.2012. Опубликовано 21.04.2012, рецензент д.т.н. Иванов А.И.