Труды научно-технической конференции кластера пензенских предприятий,
обеспечивающих **БЕЗОПАСНОСТЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.**
Том 8, Пенза-2012, с. 63-64 (http://пниэи.рф/activity/science/BIT/T8-p63.pdf)

ОЦЕНКА ВЕРОЯТНОСТИ ОШИБОЧНОГО ОТКАЗА В ДОСТУПЕ «СВОЕМУ» НА МАЛОМ ЧИСЛЕ ТЕСТОВЫХ ПРИМЕРОВ

Фунтикова Ю.В., Захаров О.С. (г. Пенза)

Перед нейросетевыми преобразователями биометрия-код стоит задача высокоточного преобразования биометрических данных образа «Свой» в код ключа доступа, удовлетворяющего требованиям базового стандарта ГОСТ Р 52633.0-2006 [1]. Для того чтобы убедиться в способности обученного преобразователя биометрия-код узнавать образ «Свой», необходимо осуществлять его тестирование после каждого обучения. В случае, когда вероятность ошибок первого рода мала, требуется достаточно большое количество тестовых попыток, которые, как правило, дают нулевое расстояние Хемминга (все биты выходного кода совпадают с заданным при обучении кодом «Свой»).

Для обычных преобразователей с допустимым значением вероятности $P1 = 0{,}1$ проблем с тестированием не возникает, достаточно базы из 20 тестовых примеров. Ситуация меняется, если по техническому заданию необходимо обеспечить $P1 = 0{,}001$. Для подтверждения такой вероятности пользователю «Свой» потребуется воспроизвести не менее 2 000 рукописных парольных слов, на что может уйти несколько часов рабочего времени. Возникает задача ускоренной достоверной оценки малых значений вероятности на ограниченной тестовой выборке из 20 примеров. В этой ситуации достигается стократное снижение трудозатрат пользователя на тестирование.

Достичь столь существенного сокращения трудозатрат удаётся, опираясь на факт нулевой энтропии и предельно высокой корреляции кодов «Свой», а также практически нулевого числа степеней свободы системы:

$$\begin{cases} E(|r_{i,j}|) \approx 1 \\ H(n) \approx 0 \\ \omega \approx 0 \end{cases}, \qquad (1)$$

где $r_{i,j}$ – коэффициент парной корреляции между i-тым и j-тым разрядами, контролируемых кодов «Свой» длиной – n; $H(n)$ – энтропия кодов «Свой»; $E(.)$ – оператор вычисления математического ожидания; ω – остаточное число степеней свободы исследуемой системы.

Число степеней свободы оценивается следующим образом:

$$\omega = E(h) \approx \frac{1}{k}\sum_{i=1}^{k} h_i , \qquad (2)$$

где h_i – расстояние Хемминга между кодом «Свой», использованным при обучении нейронной сети, и i-тым примером из k примеров тестовой выборки.

Очевидно, что выражение (2) применимо только в том случае, когда в серии из k опытов обнаружена хотя бы одна ошибка с ненулевым расстоянием Хемминга. В случае если в серии из k опытов все расстояния Хемминга нулевые (нет ошибок), оценку числа степеней свободы ведут исходя из предположения, что в следующем (k+1) опыте будет обнаружена ошибка с минимальным расстоянием Хемминга:

$$\omega = E(h) \approx \frac{1}{k+1} \qquad (3)$$

Исходя из гипотезы распределения χ^2 с очень малым числом степеней свободы, вероятность ошибок первого рода может быть вычислена по следующей формуле:

$$P_1 \approx \int_1^{\infty} \frac{1}{2^{\frac{\omega}{2}} \cdot \Gamma\left(\frac{\omega}{2}\right)} \cdot x^{\frac{\omega}{2}-1} \cdot e^{-\frac{x}{2}} \cdot dx, \qquad (4)$$

где $\Gamma(.)$ – гамма функция.

Примеры оценки вероятности ошибок первого рода по формуле (3) иллюстрируют распределения, приведённые на рисунке 1.

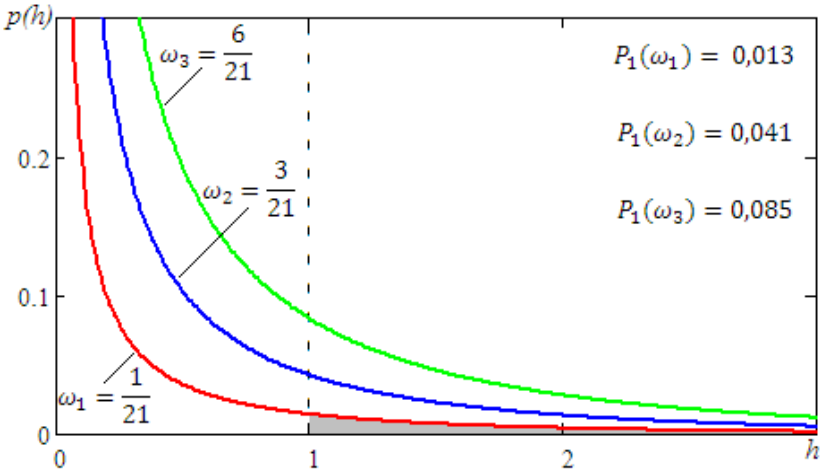

Рисунок 1 – Примеры оценки вероятностей ошибок первого рода по распределению χ^2 с малым числом степеней свободы.

Приближённая оценка малых значений корреляции может вычисляться следующим образом:

$$E_{256}(|r_{i,j}|) = 1 - 2 \cdot \left\{\frac{E(h)}{256}\right\}^1 \qquad (5)$$

Таким образом, классическое распределение χ^2 может быть обобщено под случай сильно зависимых биометрических данных, при этом приходится применять дробные показатели числа степеней свободы [3], которые также оказываются почти нулевыми (очень близкими к нулю).

Литература:

1 ГОСТ Р 52633.0-2006 «Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации».

2 Язов Ю.К., Волчихин В.И., Фунтиков В.А., Иванов А.И., Назаров И.Г. Нейросетевая защита персональных биометрических данных. М.: Радиотехника. 2012 г., 160 с.

3 Захаров О.С., Иванов А.И. Учет корреляционных связей биометрических данных через дробный показатель степеней свободы закона распределения значений хи-квадрат. Инфокоммуникационные технологии Том 6, № 1, 2008 г., с. 12-15.

Материалы поступили 21.11.2012, опубликовано в Интернет 12.12.2012 по положительной рецензии д.т.н., профессора Малыгина А.Ю. (Пенза).