ISSN 1991-3494

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫНЫҢ

Х А Б А Р Ш Ы С Ы

ВЕСТНИК

НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК
РЕСПУБЛИКИ КАЗАХСТАН

THE BULLETIN

OF THE NATIONAL ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN

1944 ЖЫЛДАН ШЫҒА БАСТАҒАН
ИЗДАЕТСЯ С 1944 ГОДА
PUBLISHED SINCE 1944

4

АЛМАТЫ
АЛМАТЫ
ALMATY

2015

ШІЛДЕ
ИЮЛЬ
JULY

«Вестник Национальной академии наук Республики Казахстан». ISSN 1991-3494
Собственник: РОО «Национальная академия наук Республики Казахстан» (г. Алматы)
Свидетельство о постановке на учет периодического печатного издания в Комитете информации и архивов Министерства культуры и информации Республики Казахстан **№5551-Ж**, выданное 01.06.2006 г.

Периодичность: 6 раз в год
Тираж: 2000 экземпляров

Адрес редакции: 050010, г. Алматы, ул. Шевченко, 28, ком. 219, 220, тел. 272-13-19, 272-13-18.
www: nauka-nanrk.kz, bulletin-science.kz

Адрес типографии: ИП «Аруна», г. Алматы, ул. Муратбаева, 75

Bulletin of the National Academy of Sciences of the Republic of Kazakhstan.
ISSN 1991-3494

Owner: RPA "National Academy of Sciences of the Republic of Kazakhstan" (Almaty)
The certificate of registration of a periodic printed publication in the Committee of Information and Archives of the Ministry of Culture and Information of the Republic of Kazakhstan N 5551-Ж, issued 01.06.2006

Periodicity: 6 times a year
Circulation: 2000 copies

Editorial address: 28, Shevchenko str., of. 219, 220, Almaty, 050010, tel. 272-13-19, 272-13-18,
http://nauka-nanrk.kz /, http://bulletin-science.kz

Address of printing house: ST "Aruna", 75, Muratbayev str, Almaty

BULLETIN OF NATIONAL ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN
ISSN 1991-3494
Volume 4, Number 356 (2015), 64 – 70

THE CALCULATION OF ENTROPY OF WEAKLY CORRELATED AND STRONGLY CORRELATED LONG BIOMETRIC CODES ON LOW TEST SAMPLES

A. I. Ivanov[1], B. B. Ahmetov[2], A. V. Bezjaev[3], K. A. Perfilov[4], Zh. K. Alimseitova[5]

[1]Penza research electrotechnical institute, Russia,
[2]International Kazakh-Turkish university of A. Yasavi, Kazakhstan, Turkestan,
[3]Penza Federal State Unitary Enterprise "STC "Atlas" branch, Russia,
[4]Penza state university, Russia,
[5]Kazakh national technical university of name K. I. Satpayev, Almaty, Kazakhstan.
E-mail: zhuldyz_al@mail.ru

Key words: biometry, evaluation of long codes entropy, the weak correlation between digits of long codes.

Annotation. It is shown that the calculation of entropy of the long weakly correlated codes according to Shannon is a very complex computational problem. There was proposed from an assessment of the probability of occurrence of long codes to go into space distance of Hamming between them. Such code conversion allows for sufficiently small test samples of 200 experiments to find the Hamming distance distribution, and it will predict the value of the entropy of weakly correlated codes. At the same time with high reliability there can be used the hypothesis of the normal distribution of discrete values of the Hamming distance. There are given limits on the average value of the coefficient module of digits correlation of the investigated codes. For strongly correlated codes of Hamming distances distribution there is proposed to describe the chi-square distribution with degrees of freedom is much less than unity.

УДК 519.2; 519.6

ВЫЧИСЛЕНИЕ ЭНТРОПИИ СЛАБО КОРРЕЛИРОВАННЫХ И СИЛЬНО КОРРЕЛИРОВАННЫХ ДЛИННЫХ БИОМЕТРИЧЕСКИХ КОДОВ НА МАЛЫХ ТЕСТОВЫХ ВЫБОРКАХ

А. И. Иванов[1], Б. Б. Ахметов[2], А. В. Безяев[3], К. А. Перфилов[4], Ж. К. Алимсеитова[5]

[1]Пензенский научно-исследовательский электротехнический институт, Россия,
[2]Международный Казахско-Турецкий университета им. А. Ясави, Казахстан, Туркестан,
[3]Пензенский филиал ФГУП «НТЦ «Атлас», Россия,
[4]Пензенский государственный университет, Россия,
[5]Казахский национальный технический университет имени К. И. Сатпаева, Алматы, Казахстан

Ключевые слова: биометрия, оценка энтропии длинных кодов, слабая корреляционная связь между разрядами длинных кодов.

Аннотация. Показано, что вычисления энтропии длинных слабо-коррелированных кодов по Шеннону является очень сложной вычислительной задачей. Предложено от оценки вероятности появления длинных кодов перейти в пространство расстояний Хэмминга между ними. Подобное преобразование кодов позволяет на достаточно малых тестовых выборках из 200 опытов находить распределение расстояний Хэмминга и по нему предсказывать значение энтропии слабо коррелированных кодов. При этом с высокой достоверностью можно пользоваться гипотезой нормального закона распределения дискретных значений расстояний Хэмминга. Даны ограничения по среднему значению модуля коэффициентов корреляции разрядов исследуемых кодов. Для сильно коррелированных кодов распределение расстояний Хэмминга предложено описывать хи-квадрат распределением с числом степеней свободы много меньше единицы.

Введение. Рассмотрим задачу измерения энтропии некоторого текста на русском языке закодированном в стандартной кодировке 92 символов клавиатуры в двух регистрах (КОИ–8). В этом случае энтропия одиночного символа текста составит:

$$H("x") \approx -\sum_{i=1}^{92} P("x_i") \cdot \log_2(P("x_i")), \quad (1)$$

где $"x_i"$ – 8 битная кодировка i-го символа; $P("x_i")$ – вероятность появления i-го символа в тексте.

Очевидно, что для достаточно надежных оценок энтропии одного символа (1) достаточно одной страницы текста на русском языке (2000 символов на страницу). Для оценки энтропии двух рядом стоящих символов русскоязычного текста затраты ресурсов растут:

$$H("x_1, x_2") \approx -\sum_{j=1}^{92}\sum_{i=1}^{92} P("x_{1,i}, x_{2,j}") \cdot \log_2(P("x_{1,i}, x_{2,j}")), \quad (2)$$

где $"x_1, x_2"$ – 16 битная кодировка пары символов; $P("x_{1,i}, x_{2,j}")$ - вероятность появления пары рядом стоящих символов в исследуемом тексте.

Для вычислений пары рядом стоящих символов нам потребуется уже не менее десятка страниц русскоязычного текста. Наблюдается экспоненциальный рост вычислительных затрат и размеров необходимого для расчетов текста. Идти по пути Шеннона при ожидании редких событий оценке многомерной энтропии весьма и весьма затратно. Необходимо создавать новые более эффективные в вычислительном отношении алгоритмы, позволяющие оценивать энтропию зависимых кодов длинной порядка 256 бит [1, 2] и выше.

Переход в пространство расстояний Хэмминга. Известно, что переход к любой иной кодировке текста не приводят к изменению энтропии, если кодировка однозначна. Перейдем от обычной кодировки знаков русского языка к кодам расстояний Хэмминга между ними:

$$h("x","c") = \sum_{i=1}^{8} ("x_i") \oplus ("c_i"). \quad (3)$$

Если речь будет идти о парах знаков, то расстояние Хэмминга будет вычисляться путем сравнения более длинных кодов:

$$h("x_1, x_2","c_1, c_2") = \sum_{i=1}^{16} ("\tilde{x}_i") \oplus ("\tilde{c}_i"), \quad (4)$$

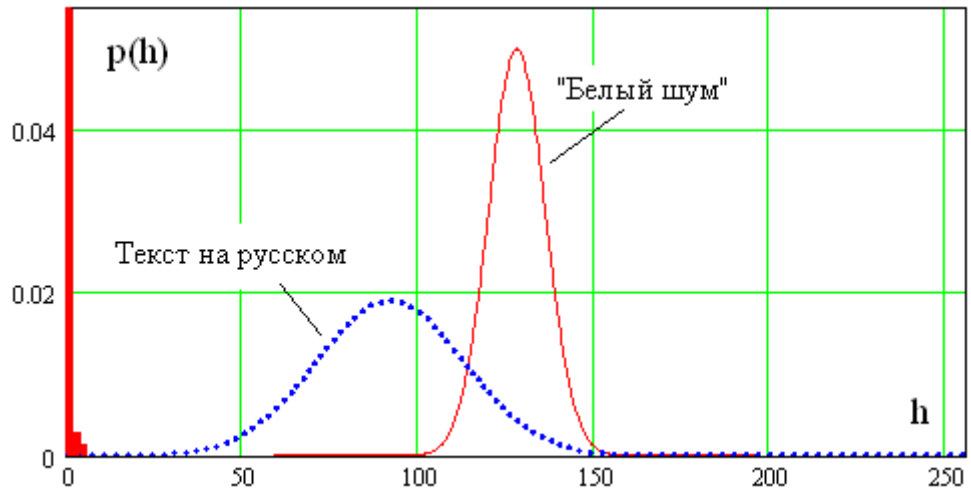
где "$\tilde{x}$" – 16-ти битный код, образованный конкатенацией двух 8-ми битных кодов "x_1, x_2" знаков кириллицы.

В случае, если мы будем сравнивать последовательности из 32 символов русскоязычного текста при вычислении расстояний Хэмминга придется сравнивать 256 разрядные коды:

$$h("x_1, x_2,..., x_{32}","c_1, c_2,..., c_{32}") = \sum_{i=1}^{256} ("\tilde{x}_i") \oplus ("\tilde{c}_i"). \quad (5)$$

Заметим, что для вычисления 2000 расстояний Хэмминга последовательности из 32 символов достаточно всего 1 страницы текста на русском языке. Преимущество перехода в пространство расстояний Хэмминга состоит в резком снижении требованиях к размерам тестовой выборки.

При оценках энтропии русскоязычных текстов для коротких последовательностей знаков по Шеннону и по Хэммингу значения расходятся. В этом случае, число состояний кодов различны и разными оказываются их статистические характеристики. Однако по мере роста длины исследуемых кодовых последовательностей наблюдается нормализация распределений расстояний Хэмминга и снижение методической погрешности оценки энтропии в разных системах отсчета. При этом распределения расстояний Хэмминга для «белого шума» и для русскоязычного текста существенно отличаются (рисунок).



Распределение расстояний Хэмминга для «белого шума» и
для 32 рядом стоящих символов осмысленного русскоязычного текста

Экономичный способ оценки энтропии «белого шума» и осмысленных парольных фраз на русском языке. Из рисунка 1 видно, что распределение расстояний Хэмминга для длинных последовательностей знаков хорошо описываются нормальным законом распределения значений. Это означает, что мы можем предсказать стойкость к атакам подбора случайной последовательности (64 бит «белого шума») и такой же осмысленной не случайной последовательности фрагмента русскоязычного текста. Для этого нам потребуется вычислить математическое ожидание – E(h) и стандартное отклонение - σ(h) двух нормальных распределений. Далее мы можем оценить вероятность подбора случайного и осмысленного пароля по следующей формуле:

$$P_2 \approx \frac{1}{\sigma(h)\sqrt{2\pi}} \int_0^1 \exp\left\{\frac{-(E(h)-u)^2}{2(\sigma(h))^2}\right\} \cdot du. \quad (2)$$

Вычисление вероятности осуществляется исходя из условия попадания расстояния Хэмминга в интервал от 0 до 1 (h=0). Далее энтропия кодовых последовательностей может быть оценена через логарифмирование:

$$H("x_1, x_2, x_3, ..., x_{32}") \approx -\log_2(P_2). \quad (3)$$

Подобные оценки являются приближенными, так как не учитывают разницу между числом возможных состояний реальных данных и много меньшим числом состояний кодов Хэмминга. Именно по этой причине должна возникать некоторая методическая погрешность. Оценить эту методическую погрешность проще всего на «белом шуме». Для «белого шума» E(h) = 128, а σ(h) = 8.0. Подстановка этих данных дает значение энтропии:

$$H("x_1, x_2, x_3, ..., x_{32}") = H("\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, ..., \tilde{x}_{256}") \approx 187{,}1 \text{ бит.} \quad (4)$$

В теории должна быть энтропия - 256 бит, оценка оказывается заниженной - 187,1 бита. Аддитивная ошибка составляет порядка 69 бит, ее можно скомпенсировать увеличив оценку в 1.27 раза.

Для осмысленной парольной фразы (распределение расстояний Хэмминга на рисунке 1 дано пунктиром) оценка энтропии, выполненная в соответствии с (2) и (3) дает значение 23,2 бита. Для учета методической ошибки занижения оценки необходимо эту величину умножить на 1.27, в итоге имеет оценку в 29.5 бита.

Получается, что оценка энтропии длинных слабо коррелированных кодов может быть осуществлена на небольшой тестовой выборке [3] в силу того, что распределение расстояний Хэмминга хорошо описывается нормальным законом.

Оценка энтропии сильно коррелированных длинных кодов. Преобразователи биометрия-код могут быть выполнены по разным технологиям. Например, может быть использована технология, так называемых «нечетких экстракторов» [4–9]. Эта технология сводится к тому, что из биометрического образа извлекаются сотни контролируемых биометрических параметров. Далее каждый из биометрических параметров подается на квантователь, дающий два выходных состояния «0» или «1». В итоге получается био-код, который как правило содержит порядка 30% ошибок, если на «нечеткий экстрактор» подавать примеры образа «Свой». Для того, чтобы био-код сделать однозначным его наиболее нестабильные разряды маскируют, далее используют какой либо избыточный код, способный обнаруживать и исправлять ошибки.

Для «нечетких экстракторов» основной проблемой является проблема доступности процедуры аутентификации. Из-за относительно низкой исправляющей способности классических самокорректирующихся кодов часто возникают ошибки нескольких разрядах в био-кода аутентификации.

Аналогичная ситуация возникает и при использовании нейросетевых преобразователей биометрия-код [1–3]. Искусственная нейронная сеть такого преобразователя обучается на конечно числе примеров образа «Свой». Как следствие, при аутентификации возникают ошибки отказа в доступе с вероятностью – P_1.

При тестировании преобразователей биометрия-код, созданных по любой технологии, возникает необходимость оценки вероятности ошибок первого рода – P_1. Если оцениваемая вероятность велика (0.1 и выше), технологических проблем не возникает. Достаточно использовать порядка 20 примеров образа «Свой», получить один или два отказа в доступе и рассчитать вероятность ошибок.

Положение резко меняется, когда требуется оценивать вероятность ошибок первого рода на уровне 0.001 и ниже. В этом случае требуется выборка из 2000 примеров образа «Свой». От пользователя средства биометрической аутентификации при тестировании требуются значительные усилия, что снижает эргономические качества биометрической технологии. В связи с этим возникает задача снижения размеров тестовой выборки примеров образа «Свой» при тестировании.

Для решения этой задачи, воспользуемся переходом в пространство расстояний Хэмминга (3). Для определенности будем считать, что имеется выборка из 20 примеров биометрического образа «Свой», которая дала 17 примеров с нужным кодом длинной 256 разрядов (h=0), 2 примера с ошибкой в 1 разряде (h=1), один пример с ошибками в 3 разрядах (h=3). Соответственно математическое ожидание расстояний Хэмминга составит E(h)=0.25, стандартное отклонение составит σ(h)=0.698.

Проведенные ранее исследования [2, 3] показали, что расстояния Хэмминга выходных кодов идеальных преобразователей хорошо описывается биномиальным законом распределения значений:

$$p(h) = \left[\frac{n!}{h!(n-h)!} \right] \cdot \tilde{P}^{h} \cdot (1-\tilde{P})^{n-h} , \tag{5}$$

где $\tilde{P}$ – средняя вероятность появления одного и того же состояния в каждом из – n разрядов био-кода.

Для идеального преобразователя биометрия-код $\tilde{P}$ = 0.5, если разряды био-кодов образов «Чужие» слабо коррелированны, плотность распределения нормальная. В случае, если мы имеем дело с кодами «Свой», параметр $\tilde{P} \approx 0.999\ldots$ В этом случае биномиальный закон дает выброс плотности распределения в близи точки h = 0.0. Из теории известно [7], что в этом случае биномиальный закон (5) хорошо приближается хи-квадрат распределением:

$$p(h) = \frac{1}{2^{\frac{m}{2}} \cdot \Gamma\left\{\frac{m}{2}\right\}} \cdot h^{\left\{\frac{m}{2}-1\right\}} \cdot \exp\left\{\frac{-h}{2}\right\}, \tag{6}$$

где m – число степеней свободы хи-квадрат распределения; $\Gamma(.)$ – гамма функция.

Случае целого числа степеней свободы m = 1, 2, 3, …. хорошо исследован [10], однако этот тип распределения плохо описывает распределения расстояний Хэмминга кодов «Свой». Для биометрических данных [11] число степеней свободы всегда оказывается не целым (фрактальным). При этом чем более коррелированными являются разряды кодов, тем меньше показатель числа степеней свободы.

Для био-кодов на выходе «нечетких экстракторов» до коррекции m = E(h)≥ 5. Если тот же показатель вычислять после корректировки ошибок, то m = E(h)≥ 0.5. В нашем случае тестирования нейросетевого преобразователя m=E(h)=0.25. Однако верить этому значению числа степеней свободы нельзя из-за малого размера тестовой выборки.

Проблема состоит в том, что для применения хи-квадрат распределения (6) одновременно должно выполняться два условия:

$$\begin{cases} m = E(h) \\ \sigma(h) = 2 \cdot m \end{cases} \tag{7}$$

Если мы принимает m=E(h)=0.25, то σ(h)=0.5, тогда как стандартное отклонение оказывается выше –0.698. Подобное расхождение будем считать ошибкой, обусловленной конечной обучающей выборкой. Для его компенсации следует найти расхождение дисперсии Δσ(h)=0.198 и компенсировать его увеличением показателя числа степеней свободы на Δm. В нашем случае следует увеличить число степеней свободы до 0.35:

$$m \approx E(h) + \frac{\Delta\sigma(h)}{2} . \tag{8}$$

После подобной коррекции оценки числа степеней свободы вероятность ошибок первого рода оценивается следующим образом:

$$P_1 = \frac{1}{2^{\frac{m}{2}} \cdot \Gamma\left\{\frac{m}{2}\right\}} \int_0^1 h^{\left\{\frac{m}{2}-1\right\}} \cdot \exp\left\{\frac{-h}{2}\right\} \cdot dh . \tag{9}$$

В нашем случае расчеты по формуле (9) дают P_1=0.105. Если бы мы пользовались обычным алгоритмом оценки вероятности, то получили бы P_1=3/20=0.15. То есть использование априорной информации и более сложных вычислений дает возможность снизить примерно на треть размер тестовой выборки. Выигрыш по размерам тестовой выборки от применения более сложных вычислений (7), (8), (9) быстро увеличивается по мере ужесточения требований к вероятности ошибок первого рода. При необходимости оценить энтропию кодов «Свой» следует применить

выражение (3), заменив в нем вероятность ошибок второго рода на вероятность ошибок первого рода.

Заключение. Переход от наблюдения длинных биометрических кодов «Чужой» и кодов «Свой» в пространство расстояний Хэмминга дает значительный выигрыш по требованиям к тестовой выборке. Наибольший выигрыш получается при слабо коррелированных кодах. Однако этот выигрыш сохраняется и при сильно коррелированных кодах. Видимо полностью независимые коды (типа «белый шум») и полностью зависимые коды дают локальные максимумы выигрыша в размерах тестовой выборки.

Предположительно, что в будущем ряд биометрических приложений придется создавать исходя из условия равных значений вероятностей ошибок первого и второго рода $P_{EE}=P_1=P_2$. В этом случае сложности оценки почти нулевой энтропии кодов «Свой» и предельно высокой энтропии кодов «все Чужие» оказываются сопоставимы. И в том и в другом случае прямые оценки вероятностей появления редких событий осуществлять не целесообразно. Гораздо более целесообразным является переход в пространство расстояний Хэмминга и учет априорной информации о законе распределения данных для примеров образа «Свой» и примеров разных образов «Чужие».

ЛИТЕРАТУРА

[1] 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)

[2] Ахметов Б.С., Волчихин В.И., Иванов А.И., Малыгин А.Ю. Алгоритмы тестирования биометрико-нейросетевых механизмов защиты информации Казахстан, Алматы, КазНТУ им. Сатпаева, 2013 г.- 152 с. ISBN 978-101-228-586-4, http://portal.kazntu.kz/files/publicate/2014-01-04-11940.pdf

[3] Ахметов Б.С., Надеев Д.Н., Фунтиков В.А., Иванов А.И., Малыгин А.Ю. Оценка рисков высоконадежной биометрии. Монография. Алматы: Из-во КазНТУ им. К.И. Сатпаева, 2014 г.- 108 с.

[4] Juels A., Wattenberg M. A Fuzzy Commitment Scheme // Proc. ACM Conf. Computer and Communications Security, 1999, p. 28–36.

[5] F. Monrose, M. Reiter, Q. Li, S. Wetzel. Cryptographic key generation from voice. In Proc. IEEE Symp. on Security and Privacy, 2001.

[6] Y. Dodis, L. Reyzin, A. Smith Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy, Data April 13, In EUROCRYPT, pages 523-540, 2004.

[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] Feng Hao, Ross Anderson, and John Daugman. Crypto with Biometrics Effectively, IEEE TRANSACTIONS ON COMPUTERS, VOL. 55, NO. 9, SEPTEMBER 2006.

[9] Чморра А.Л. Маскировка ключа с помощью биометрии «Проблемы передачи информации» 2011 № 2(47) с. 128-143.

[10] Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. М.: ФИЗМАТЛИТ, 2006 г., 816 с.

[11] Безяев А.В., Иванов А.И., Фунтикова Ю.В. Оптимизация структуры самокорректирующегося био-кода, хранящего синдромы ошибок в виде фрагментов хеш-функций. «Вестник Уральского федерального округа. Безопасность в информационной сфере» 2014 г. № 3(13) с. 4-14.

REFERENCES

[1] Ahmetov B.S., Ivanov A.I., Funtikov V.A., Bezjaev A.V., Malygina E.A. *Tehnologija ispol'zovanija bol'shih nejronnyh setej dlja preobrazovanija nechetkih biometricheskih dannyh v kod kljucha dostupa. Monografija*, Kazahstan, g. Almaty, TOO «Izdatel'stvo LEM», 2014 g., 144 c., nahoditsja v otkrytom dostupe (http://portal.kazntu.kz/files/publicate/2014-06-27-11940.pdf) (in Russ.)

[2] Ahmetov B.S., Volchihin V.I., Ivanov A.I., Malygin A.Ju. *Algoritmy testirovanija biometriko-nejrosetevyh mehanizmov zashhity informacii.* Kazahstan, Almaty, KazNTU im. Satpaeva, 2013 g., 152 s. ISBN 978-101-228-586-4, http://portal.kazntu.kz/files/publicate/2014-01-04-11940.pdf (in Russ.)

[3] Ahmetov B.S., Nadeev D.N., Funtikov V.A., Ivanov A.I., Malygin A.Ju. *Ocenka riskov vysokonadezhnoj biometrii. Monografija.* Almaty: Iz-vo KazNTU im. K.I. Satpaeva, 2014 g., 108 s. (in Russ.)

[4] Juels A., Wattenberg M. A *Fuzzy Commitment Scheme. Proc.* ACM Conf. Computer and Communications Security, 1999, p. 28–36. (in Eng.)

[5] F. Monrose, M. Reiter, Q. Li, S. Wetzel. *Cryptographic key generation from voice.* In Proc. IEEE Symp. on Security and Privacy, 2001. (in Eng.)

[6] Y. Dodis, L. Reyzin, A. Smith *Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy*, Data April 13, In EUROCRYPT, pages 523-540, 2004. (in Eng.)

[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. (in Eng.)

[8] Feng Hao, Ross Anderson, and John Daugman. *Crypto with Biometrics Effectively*, IEEE TRANSACTIONS ON COMPUTERS, VOL. 55, NO. 9, SEPTEMBER 2006. (in Eng.)

[9] Chmorra A.L. *Maskirovka kljucha s pomoshh'ju biometrii. «Problemy peredachi informacii»,* 2011 № 2(47), s. 128-143. (in Russ.)

[10] Kobzar' A.I. *Prikladnaja matematicheskaja statistika. Dlja inzhenerov i nauchnyh rabotnikov.* M.: FIZMATLIT, 2006 g., 816 s. (in Russ.)

[11] Bezjaev A.V., Ivanov A.I., Funtikova Ju.V. *Optimizacija struktury samokorrektirujushhegosja bio-koda, hranjashhego sindromy oshibok v vide fragmentov hesh-funkcij. «Vestnik Ural'skogo federal'nogo okruga. Bezopasnost' v informacionnoj sfere»* 2014 g. № 3(13) s. 4-14. (in Russ.)

АЗ МӘТІНДІК ТАҢДАУЛАРДА ӘЛСІЗ ЖӘНЕ ҚУАТТЫ КОРРЕЛЯЦИЯЛАНҒАН ҰЗЫН БИОМЕТРИЯЛЫҚ КОДТАРЫНЫҢ ЭНТРОПИЯСЫН ЕСЕПТЕУ

А. И. Иванов[1], Б. Б. Ахметов[2], А. В. Безяев[3], К. А. Перфилов[4], Ж. К. Алимсеитова[5]

[1]Пензалық ғылыми-зерттеу электротехникалық институт, Ресей,
[2]А. Ясави атындағы Халықаралық Қазақ-Түрік университеті, Қазақстан, Түркістан,
[3]Пензалық ФМБК филиалы «ҒТО «Атлас», Ресей,
[4]Пензалық мемлекеттік университет, Ресей,
[5]Қ. И. Сәтбаев атындағы Қазақ ұлттық техникалық университет, Алматы

Тірек сөздер: биометрия, ұзын кодттар энтропиясын бағалау, ұзын кодтар разрядтар арасындағы әлсіз корреляциялық байланыс.

Аннотация. Шеннон бойынша ұзын әлсіз корреляцияланған кодттардың энтропиясын есептеу өте қыйын есеп екені көрсетілген. Ұзын кодттар пайда болу ықтималдығын бағалаудан олардың арасындағы Хэмминг қашықтығы кеңістігіне өту ұсынылған. Осындай кодтарды түрлендіру жеткілікті аз мәтіндік таңдауларда 200 тәжірибе ішінен Хэмминг қашықтығы таралуын табуды және соның негізінде әлсіз корреляцияланған кодтар энтропиясының мәнін болжауға мүмкіндік береді. Сол кезде жоғары шынайлықпен Хэмминг қашықтығының дискретті мәндерін тарату қалыпты заңының гипотезасын қолдануға болады. Зерттелетін кодтар разрядтарының корреляция коэффициенттер модулінің орташа мәні бойынша шектеулер берілген. Қуатты корреляцияланған кодттар үшін Хэмминг қашықтығын таратуды бірден көп төмен бостандық дәрежелер санымен хи-квадрат таратумен жазбалау ұсынылған.

Поступила 22.05.2015 г.

Publication Ethics and Publication Malpractice
in the journals of the National Academy of Sciences of the Republic of Kazakhstan

For information on Ethics in publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайте:

www:nauka-nanrk.kz

http://www.bulletin-science.kz/index.php/ru/

Редакторы *М. С. Ахметова, Д. С. Аленов*
Верстка на компьютере *Д. Н. Калкабековой*

Подписано в печать 21.07.2015.
Формат 60x881/8. Бумага офсетная. Печать – ризограф.
12,9 п.л. Тираж 2000. Заказ 4.