УДК 004.8; 004.032.26
DOI: 10.22213/2410-9304-2022-3-88-93

Перспектива совместного использования двоичных и троичных искусственных нейронов при анализе качества «белого» шума в пространстве сверток Хэмминга, вычисленных по разным модулям

А. И. Иванов, доктор технических наук, Пензенский научно-исследовательский электротехнический институт, Пенза, Россия
А. П. Юнин, Пензенский научно-исследовательский электротехнический институт, Пенза, Россия
М. А. Бояршинов, кандидат технических наук, ИжГТУ имени М. Т. Калашникова, Ижевск, Россия
А. П. Иванов, кандидат технических наук, Пензенский государственный университет, Пенза, Россия

Работа посвящена задаче упрощения тестирования качества бинарных последовательностей, полученных либо от физических генераторов шума, либо из нестабильной части биометрических данных человека. Решаемая проблема порождена тем, что часто используемый тест «случайного нажатия клавиш обезьяной» имеет экспоненциальную вычислительную сложность. Как результат, такой тест может быть реализован только на достаточно мощной вычислительной машине.

В работе показано, что самый простой тест качества «белого» шума выполняется путем подсчета единиц. Его можно рассматривать как частный случай свертки Хэмминга, вычисленной по модулю два при сравнении с единичным кодом «111...11». Дано обобщение этого теста на его аналоги, использующие свертки Хэмминга, вычисленные по модулям 2 и 256. Приводятся программы на языке MathCAD для выполнения численных экспериментов. Даны распределения амплитуд вероятности спектральных линий Хэмминга. Предложено усилить проверку гипотезы «белого» шума дополнительным анализом с использованием троичных искусственных нейронов. Приводится оценка выигрыша от перехода к совместному использованию бинарных и троичных искусственных нейронов при проверке гипотезы «белого» шума для той или иной бинарной последовательности. Рассматриваемые в статье вычислительные процедуры не требуют значимых вычислительных ресурсов. Они ориентированы на реализацию в маломощных низко потребляющих доверенных контроллерах SIM-карт или микро-SD-карт. Это обстоятельство обусловлено линейной вычислительной сложностью всех тестов, построенных на использовании сверток Хэмминга, независимо от значений модулей, по которым они вычисляются. Также не влияет на вычислительную сложность число уровней выходных квантователей того или иного искусственного нейрона.

Ключевые слова: проверка гипотезы «белого» шума, двоичный нейрон, троичный нейрон, избыточный код с обнаружением и исправлением ошибок.

Введение

В соответствии с Указом Президента В. В. Путина № 490 от 10.10.19 «О развитии искусственного интеллекта в РФ» создан технический комитет (ТК) по стандартизации № 164 «Искусственный интеллект» (ТК164 «ИИ»). В ТК164 работает группа РГ03 «Качество технологий искусственного интеллекта». Одной из важнейших характеристик качества приложений искусственного интеллекта является уровень их защищенности. Защитить приложения искусственного интеллекта возможно гомоморфным шифрованием [1–3], тогда безопасность приложений обеспечивается тем, что решающие правила искусственного интеллекта работают с зашифрованными данными. Это делает возможным выполнение доверенных вычислений даже в «облаках» [4–6], что нашло отражение также в международном стандарте ISO/IEC 18033-6:2019.

Второй путь защиты приложений искусственного интеллекта – это использование аппаратно-программных решений, выполненных в форме самостоятельного компактного модуля [7] доверенной вычислительной среды.

Доверенная вычислительная среда искусственного интеллекта оказывается эффективной, если все криптографические операции с ключами выполняются внутри нее. В том числе внутри нее желательно выполнять синтез криптографических ключей и их тестирование, а также нейросетевое связывание биометрических данных пользователя с его личным криптографическим ключом. Например, такое связывание может быть выполнено автоматическим обучением нейросети по ГОСТ Р 52633.5–2011.

Проблема, рассматриваемая в данной статье, обусловлена техническим противоречием между стоимостью аппаратной части доверенной вычислительной среды и ее производительностью.

Использование мощных процессоров со значительным потреблением позволяет решать все проблемы по шифрованию, синтезу и тестированию криптографических ключей, обучению нейронных сетей, применению нейронных сетей, защищенных криптографическими механизмами. Однако стоимость таких технических решений оказывается высокой. Снизить стоимость удается, если использовать компактные, дешевые, мало потребляющие контроллеры SIM-карт и микро-SD-карт. При этом процедуры тестирования качества криптографических ключей, созданных внутри SIM-карт и микро-SD-карт, должны быть ориентированы на низкое потребление вычислительных ресурсов.

Для нас не имеет значения, как был получен «сырой» ключ [8], в рамках этой статьи для нас имеют значение только процедуры оценки качества той или иной битовой последовательности или ее близости к «белому» шуму.

Проверить качество битовой последовательности можно несколькими тестами NIST [9] (Национального института стандартизации США), признаваемыми отечественной криптографической общественностью. Также могут быть использованы и другие тесты [10–12] проверки качества генераторов псевдослучайных последовательностей.

К сожалению, не все тесты качества «белого» шума легко реализуемы на маломощных низкопотребляющих доверенных контроллерах SIM-карт или микро-SD-карт. В частности, тест «обезьяны», случайно нажимающей клавиши компьютера, имеет экспоненциальную вычислительную сложность [13, 14]. Проработке этих вопросов была посвящена статья авторов в этом журнале «Оценка энтропии длинных кодовых слов на выходе нейросетевого преобразователя биометрии в пространствах множества сверток Хэмминга» (Интеллектуальные системы в производстве. 2019. № 2). Как результат, этот тест нецелесообразно реализовывать в доверенной вычислительной среде мобильных пользователей с малым энергопотреблением.

Простейший тест качества «белого» шума подсчетом единиц

Одним из самых «легких» в вычислительном отношении является тест, построенный на подсчете числа единиц в проверяемом коде. На рис. 1 приведена программная реализация численного эксперимента на языке MathCAD и результаты численного эксперимента.

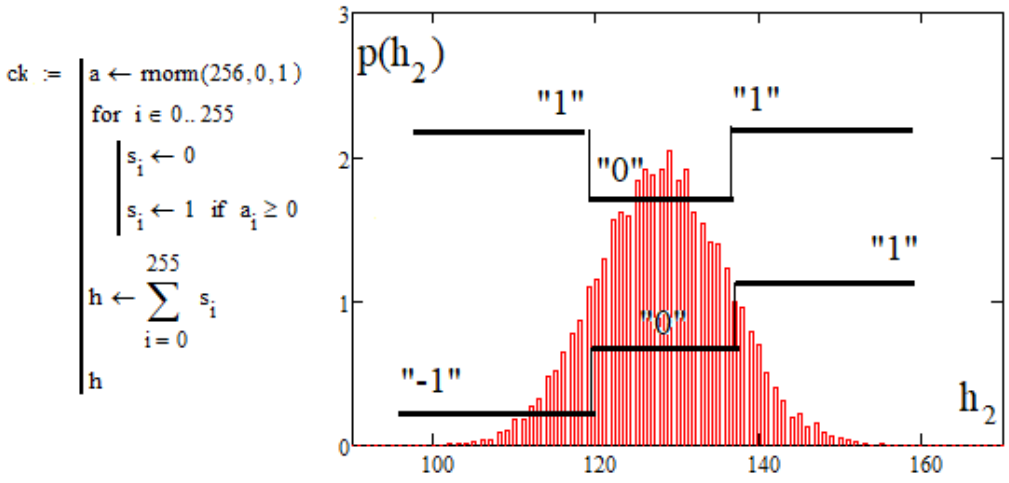

Рис. 1. Распределение расстояний Хэмминга, вычисленных по отношению к единичному коду «1111…11»

Fig. 1. Distribution of Hamming distances calculated with respect to the single code "1111…11"

Важно отметить, что тест подсчета единиц является частным случаем тестов, построенных на вычислении расстояний Хэмминга по модулю два. Его можно рассматривать как свертку Хэмминга по модулю два с кодом, состоящим только из единиц «1111…11». Совпадение выполняется только в одной точке:

$$h_2 = \sum_{i=1}^{256} ("1") \oplus ("x_i") = \sum_{i=1}^{256} ("x_i"). \qquad (1)$$

Еще одним важным моментом является то, что обычные бинарные нейроны имеют входные сумматоры. Если мы на выход сумматора подключим квантователь, то получим искусственный нейрон, работающий в пространстве расстояний Хэмминга. На рис. 1 приведены выходные состояния бинарного нейрона, выполняющего сравнения выходных данных сумматора с двумя порогами k_1=118 и k_2=138. При таких порогах рассматриваемый искусствен-

ный нейрон отбрасывает коды с большим и малым расстоянием Хэмминга с вероятностью 0,305.

Тестирование качества «белого» шума в пространстве расстояний Хэмминга

Ранее было показано, что искусственные нейроны могут быть построены для расстояний Хэмминга, вычисленных по разным модулям [15]. В том числе могут быть использованы расстояния Хэмминга, вычисленные по модулям 4, 8, 16, ..., 256, ..., 2^{16}. При этом вычислительная сложность для каждой операции обогащения данных остается линейной. Как пример, на рис. 2 приведены процедуры накопления данных, построенные на вычислении расстояний Хэмминга по модулю 256 (окно анализа по 8 бит без перекрытия с соседями).

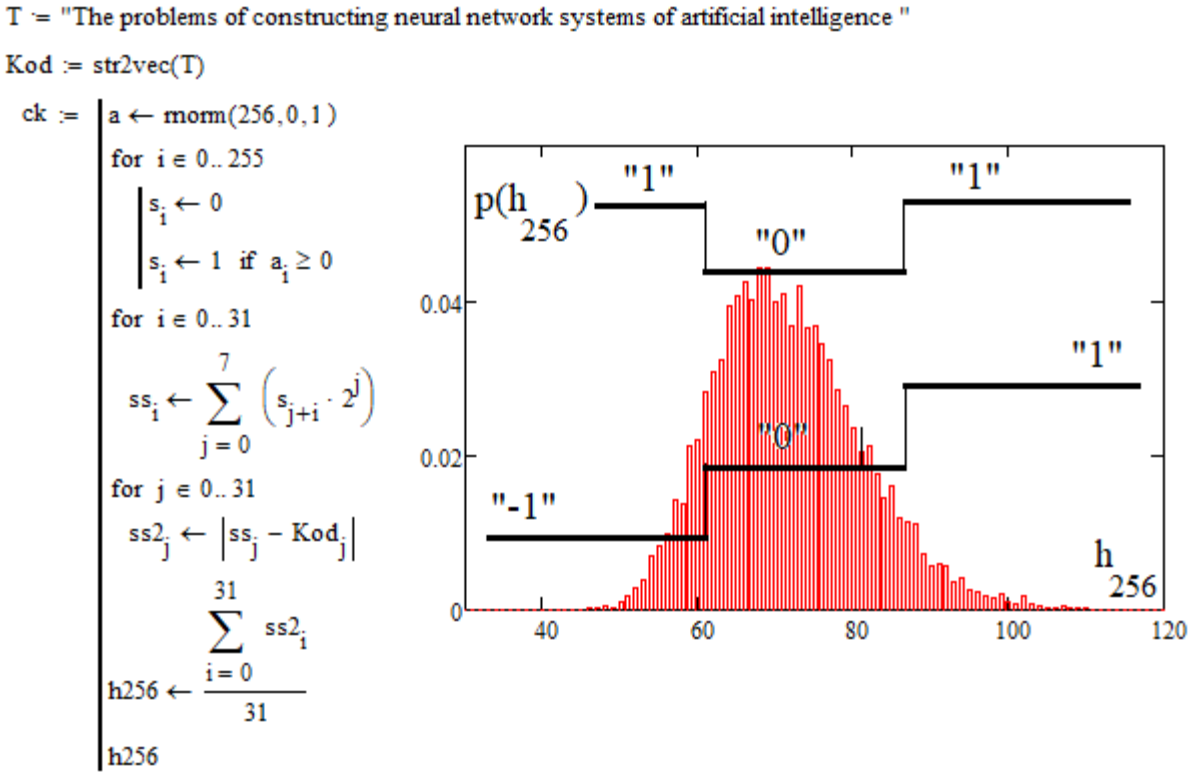

Рис. 2. Распределение расстояний Хэмминга, вычисленных по модулю 256

Fig. 2. Distribution of Hamming distances modulo 256

Выбор окна наблюдения в 8 бит обусловлен тем, что классические символьные *ASCII*-коды 8-битные. То есть эти условия наиболее близки к исходным параметрам теста случайных нажатий на клавиши обезьяны. При этом распределение расстояний Хэмминга по модулю 256 является асимметричным. Тем не менее это распределение легко использовать для вычисления порогов квантования бинарного нейрона k_1=62 и k_2=92. При таком значении порогов выходного квантователя нейрона Хэмминга отбрасывает «слабые» ключи с вероятностью 0,297.

Принципиально важным является также то, что рассмотренный простейший критерий (1) и множество критериев, построенных на вычислении расстояний Хэмминга разных модулей [13], независимы: $corr(h_2, h_{64})\approx 0{,}0$, $corr(h_2, h_{128})\approx 0{,}0$, $corr(h_2, h_{256})\approx 0{,}0$, $corr(h_2, h_{512})\approx 0{,}0$. Это происходит из-за того, что простейший критерий (1) не является полноценной сверткой Хэмминга.

Как следствие, совместное использование двух независимых критериев должно позволить снизить вероятность пропуска «слабых» кодов до значения 0,305×0,297=0,091.

Перспектива перехода от бинарных нейронов к троичным нейронам

Ранее было показано, что нейросети обладают более высокой корректирующей способностью по сравнению с классическими избыточными кодами, способными обнаруживать и исправлять ошибки. Причина состоит в том, что классические самокорректирующиеся коды строятся под некоторую статистическую гипотезу о распределении ошибок. Классические коды *c* избыточные, не способны обучаться,

учитывая конкретные распределения ошибок в разрядах реальных кодов.

При усилении возможностей искусственных нейронов выгодным оказывается переходить от бинарных нейронов к троичным нейронам или к нейронам с большим числом выходных состояний квантователей. Для того чтобы усилить корректирующие способности нейронов, в нашем случае целесообразно перейти от двоичных нейронов к троичным нейронам с теми же порогами. Для накопления данных простым суммированием единиц эта ситуация отображена на рис. 1. Соответственно, троичный нейрон дает состояния {«–1», «0», «1»}.

Для накопления данных в пространстве расстояний Хэмминга квантователь троичного нейрона имеет те же пороги, что и у бинарного квантователя. При совместном использовании двух троичных нейронов мы имеем значительную кодовую избыточность. Три состояния в бинарной системе кодируются двумя битами, соответственно, два троичных нейрона дают четыре выходных бита. Мы имеем четырехкратную кодовую избыточность и, соответственно, возможность построить код с обнаружением и исправлением ошибок. На рис. 3 представлена спектрограмма амплитуд вероятности Хэмминга для рассматриваемой ситуации.

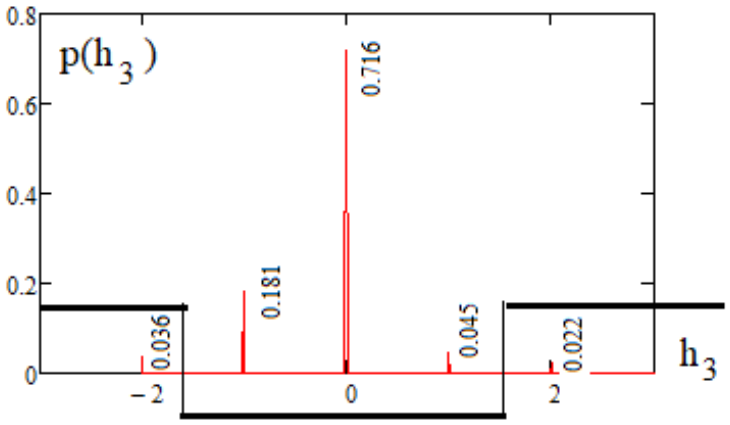

Рис. 3. Спектральные линии, корректировки кодов двумя троичными нейронами

Fig. 3. Spectral lines, code corrections by two ternary neurons

Из рис. 3 видно, что спектрограмма четырехбитного кода имеет пять значимых спектральных линий. Самокорректирующийся код учитывает как допустимые три центральные спектральные линии. Крайние спектральные линии соответствуют ситуации, когда ошибка обнаружена, но не может быть исправлена. То есть переход к троичным нейронам вместо двоичных обеспечивает снижение вероятности пропуска «слабых» случайных последовательностей до 0,058, что примерно на 57 % меньше.

Заключение

Рассмотренные в данной статье процедуры имеют линейную вычислительную сложность и, соответственно, вполне могут быть реализованы на маломощных низко потребляющих доверенных контроллерах SIM-карт или микро-SD-карт. Дополнительным положительным моментом изложенного нами подхода к тестированию качества «белого» шума является то, что программная реализация первого критерия имеет четыре строки кода, а второго критерия – в 2 раза сложнее (восемь строк кода, см. рис. 2). Параллельно со значительной экономией вычислительных ресурсов мы наблюдаем экономию постоянной памяти контроллеров SIM-карт, когда речь идет о биометрико-нейросетевой аутентификации мобильных пользователей.

Следует также отметить и нежелательный эффект наличия корреляционных связей между полноценными свертками Хэмминга, вычисленными по разным модулям $corr(h_{16}, h_{32}) \neq 0{,}0$, $corr(h_{32}, h_{64}) \neq 0{,}0$. Это снижает эффективность нейросетевых корректоров ошибок, построенных на подобных полноценных свертках Хэмминга. При этом корректирующая способность для бинарных нейронов будет всегда ниже корректирующей способности для троичных нейронов. Совместное использование корректирующих конструкций из бинарных и троичных нейронов также приводит к росту их общей корректирующей способности.

Библиографические ссылки

1. *Варновский Н. П, Шокуров А. В.* Гомоморфное шифрование // Труды Института системного программирования РАН. 2007. Т. 12. С. 27–36.

2. *Brakerski Z., Gentry C., Vaikuntanathan V.* Leveled fully homomorphic encryption without bootstrapping. Theoretical Computer Science, 2012. Pp. 309-325.

3. *Аракелов Г. Г.* Вопросы применения прикладной гомоморфной криптографии // Вопросы кибербезопасности. 2019. № 5(33). С. 70–74. DOI: 10.21681/2311-3456-2019-5-70-74.

4. *Минаков С. С.* Основные криптографические механизмы защиты данных, передаваемых в облачные сервисы и сети хранения данных // Вопросы кибербезопасности. 2020. № 3 (37). С. 66–75. DOI: 10.21681/2311-3456-2020-05-66-75.

5. *Варновский Н. П., Захаров В. А., Шокуров А. В.* К вопросу о существовании доказуемо стойких систем облачных вычислений // Вестник Московского университета. Серия 15. Вычислительная математика и кибернетика. 2016. № 2. С. 32–46.

6. *Астахова Л. В., Султанов Д. Р., Ашихмин Н. А.* Защита облачной базы персональных данных с использованием гомоморфного шифрования // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». 2016. Т. 16, № 3. С. 52–61. DOI: 10.14529/ctcr160306.

7. *Alan M. Dunn, Owen S. Hofmann, Brent Waters, Emmett Witchel.* Cloaking Malware with the Trusted Platform Module // SEC'11 Proceedings of the 20th USENIX conference on Security. – USENIX Association, 2011.

8. *Задорожный Д. И., Корнеева А. М., Фомичев В. М.* Патент RU 2628213. Способ генерации случайных двоичных последовательностей с использованием компьютера и действий пользователя. МПК G06F 7/58. Опубликовано: 15.08.2017. Бюл. № 23.

9. *Bassham L., Rukhin A., Soto J., Nechvatal J., Smid M., Leigh S., Levenson M., Vangel M., Heckert N. and Banks D.* (2010), A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, [online], https://tsapps.nist.gov/publication/ get_pdf.cfm?pub_id=906762 (Accessed June 16, 2022).

10. *Миненко А. И.* Экспериментальное исследование эффективности тестов для проверки генераторов случайных чисел // Вестник СибГУТИ. 2010. № 4. С. 36–46.

11. *Григорьев А. Ю.* Методы тестирования генераторов случайных и псевдослучайных последовательностей // Ученые записки УлГУ. Серия: Математика и информационные технологии. 2017. № 1. С. 22–28.

12. *Харин Ю. С., Ярмола А. Н., Петлицкий А. И.* Методы и алгоритмы статистического тестирования генераторов случайных и псевдослучайных последовательностей в системах информационной безопасности // Искусственный интеллект. 2006. № 3. С. 793–803.

13. *Иванов А. И., Юнин А. П.* Эмбрион искусственного интеллекта: компактная нейросетевая проверка качества случайных последовательностей, полученных из биометрических данных : препринт. Пенза : Изд-во ПГУ, 2021. 68 с. ISBN 978-5-907364-80-6.

14. *Малыгина Е. А.* Биометрико-нейросетевая аутентификация: перспективы применения сетей квадратичных нейронов с многоуровневым квантованием биометрических данных : препринт. Пенза: Изд-во ПГУ, 2020. 114 с. ISBN 978-5-907262-88- 1.

15. Оценка качества «белого» шума: реализация теста «стаи обезьян» через множество сверток Хэмминга, построенных на разных системах счисления / А. П. Юнин, А. И. Иванов, К. А. Ратников, Е. А. Кольчугина // Известия высших учебных заведений. Поволжский регион. Технические науки. 2018. № 4 (48). С. 54–64. DOI 10.21685/2072-3059-2018-4-5.

References

1. Varnovskii N.P., Shokurov A.V. [Homomorphic encryption]. *Trudy Instituta sistemnogo programmirovaniya RAN*, 2007, vol. 12, pp. 27-36 (in Russ.).

2. *Brakerski Z., Gentry C., Vaikuntanathan V.* Leveled fully homomorphic encryption without bootstrapping. Theoretical Computer Science, 2012. Pp. 309-325.

3. Arakelov G.G. [Questions of application of applied homomorphic cryptography] *Voprosy kiberbezopasnosti,* 2019, no. 5(33). pp. 70-74 (in Russ.). DOI: 10.21681/2311-3456-2019-5-70-74.

4. Minakov S.S. [The main cryptographic mechanisms for protection of data, transmitted to cloud services and storage area networks]. *Voprosy kiberbezopasnosti,* 2020, no. 3. Pp. 66-75 (in Russ.). DOI: 10.21681/2311-3456-2020-05-66-75.

5. Varnovskii N.P., Zakharov V.A., Shokurov A.V. [On the existence of provably secure cloud computing systems]. *Vestnik Moskovskogo universiteta, Seriya 15, Vychislitel'naya matematika i kibernetika*, 2016, no. 2, pp. 32-46 (in Russ.).

6. *Astakhova L.V., Sultanov D.R., Ashikhmin N.A.* [Protection of Cloud Database Containing Personal Information Using Homomorphic Encryption] Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics, 2016, vol. 16, no. 3, pp. 52-61. (in Russ.). DOI: 10.14529/ctcr160306.

7. *Alan M. Dunn, Owen S. Hofmann, Brent Waters, Emmett Witchel.* Cloaking Malware with the Trusted Platform Module // SEC'11 Proceedings of the 20th USENIX conference on Security. - USENIX Association, 2011.

8. Zadorozhnyi D. I., Korneeva A. M., Fomichev V. M. *Sposob generatsii sluchainykh dvoichnykh posledovatel'nostei s ispol'zovaniem komp'yutera i deistvii pol'zovatelya* [Method for generating random binary sequences using a computer and user actions]. Patent RF, no. 2628213, 2017.

9. *Bassham L., Rukhin A., Soto J., Nechvatal J., Smid M., Leigh S., Levenson M., Vangel M., Heckert N.*

and Banks D. (2010), A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, [online], https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=906762 (Accessed June 16, 2022).

10. Minenko A. I. [Experimental research of efficiency of tests for random number generators] *Vestnik SibGUTI*, 2010, no. 4, pp. 36-46 (in Russ.).

11. Grigor'ev A. Yu. [Methods for testing generators of random and pseudo-random sequences]. *Uchenye zapiski UlGU. Seriya: Matematika i informatsionnye tekhnologii*, 2017, no. 1, pp. 22-28 (in Russ.).

12. *Kharin Yu. S., Yarmola A. N., Petlitskii A. I.* [Methods and Algorithms for Statistical Testing of Random and Pseudo-Random Sequences Generators in Information Security Systems]. *Iskusstvennyi intellekt*, 2006, no. 3, pp. 793-803 (in Russ.).

13. Ivanov A.I., Yunin A.P. *Embrion iskusstvennogo intellekta: kompaktnaya neyrosetevaya proverka kachestva sluchaynykh posledovatel'nostey, poluchennykh iz biometricheskikh dannykh* [Embryo of artificial intelligence: compact neural network quality check of random sequences obtained from biometric data]. Penza, Izd-vo PGU, 2021, 68 p. (in Russ.).

14. Malygina E.A. *Biometriko-nejrosetevaya autentifikaciya: perspektivy primeneniya setej kvadratichnyh nejronov s mnogourovnevym kvantovaniem biometricheskih dannyh* [Biometric-Neural Network Authentication: Prospects for Using Quadratic Neuron Networks with Multilevel Quantization of Biometric Data]. Penza, Izd-vo PGU, 2020, 114 p. (in Russ.).

15. *Yunin A.P., Ivanov A.I., Ratnikov K.A., Kol'chugina E.A.* [Quality evaluation of "white" noise: implementation of the "monkeys" test through a set of Hamming convolutions constructed for different number systems]. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki*, 2018, no. 4 (48), pp. 54-64 (in Russ.). DOI 10.21685/2072-3059-2018-4-5.

* * *

The Perspective of Binary and Tertiary Artificial Neurons Mutual Application for White Noise Quality Analysis within Hamming Window Convolution Calculated by Different Modules

A. I. Ivanov, DSc in Engineering, Penza Research Institute of Electrical Engineering, Penza, Russia
A. P. Yunin, Penza Research Institute of Electrical Engineering, Penza, Russia
M. A. Boyarshinov, PhD in Engineering, Kalashnikov ISTU, Izhevsk, Russia
A. P. Ivanov, PhD in Engineering, Penza State University, Penza, Russia

The work is devoted to simplification of binary sequence quality testing, obtained either from physical noise generators, or from instable part of human biometric data. The problem originated from exponential computational complexity of a frequently used test "pressing keys by mistake". As a result, such a test can be realized only on a powerful computer.

The work shows that the simplest testing of white noise quality is performed by counting of units. It can be considered as a special case of Hamming window convolution, calculated by module two in comparison with unit code «111...11». A generalization of this test with its analogues using Hamming window convolutions calculated by modules 2 and 256 was presented. Programs based on MathCAD for performing numerical tests are given. Hamming spectral line probability amplitude distribution was presented. It was suggested to enhance verification of white noise with additional analysis using tertiary artificial neurons. The assessment of benefit from transition to mutual application of binary and tertiary artificial neurons when verifying the white noise hypothesis for one or another binary sequence was given. The calculational procedures considered in the article do not require significant computational resources. They are oriented on low power and consumption reliable microcontrollers of SIM-cards and SD-cards.

This is stipulated by linear computational complexity of all tests built on Hamming window convolution irrespectively of module values, they are calculated by. The number of output quantizer levels of one or another artificial neurons has no effect on computational complexity.

Keywords: white noise hypothesis verification, binary neuron, tertiary neuron, error checking and correction redundant code.

Получено: 29.06.22