

## ОЦЕНКА КАЧЕСТВА РАБОТЫ МНОГОКРИТЕРИАЛЬНОГО НЕЙРОСЕТЕВОГО СТАТИСТИЧЕСКОГО АНАЛИЗА НА ПРИМЕРЕ ПРОВЕРКИ ГИПОТЕЗЫ НЕЗАВИСИМОСТИ ДАННЫХ МАЛЫХ ВЫБОРОК

А. И. Иванов<sup>1</sup>, А. И. Ермакова<sup>2</sup>

<sup>1</sup> Пензенский научно-исследовательский электротехнический институт, Пенза, Россия

<sup>2</sup> Пензенский государственный университет, Пенза, Россия

<sup>1</sup> ivan@pnici.penza.ru, <sup>2</sup> aermakova27061992@mail.ru

**Аннотация.** *Актуальность и цели.* Целью статьи является создание критерия качества оценки эффективности нейросетевой обработки малых выборок при многокритериальной проверке гипотезы независимости. *Материалы и методы.* На сегодняшний день известно порядка 200 статистических критериев проверки гипотезы независимости малой выборки. Используя их совместно, удастся существенно повысить качество принимаемых нейросетью решений. В связи с этим возникает проблема эффективной сравнительной оценки достигнутого качества той или иной нейросетью. *Результаты и выводы.* Предложено оценивать качество нейросетевого решения по достигнутому увеличению выборки при ее анализе одним-единственным критерием Пирсона – Эдлтона – Эджурта конца 19 в. на выборке большого объема в наиболее сложной ситуации проверки гипотезы независимости.

**Ключевые слова:** проверка гипотезы независимости, коэффициент корреляции, малые выборки, нейросетевое объединение большого числа критериев

**Для цитирования:** Иванов А. И., Ермакова А. И. Оценка качества работы многокритериального нейросетевого статистического анализа на примере проверки гипотезы независимости данных малых выборок // Надежность и качество сложных систем. 2024. № 3. С. 13–18. doi: 10.21685/2307-4205-2024-3-2

## ASSESSING THE PERFORMANCE QUALITY OF MULTICRITERIA NEURAL NETWORK STATISTICAL ANALYSIS USING THE EXAMPLE OF TESTING THE HYPOTHESIS OF INDEPENDENCE OF SMALL SAMPLE DATA

A.I. Ivanov<sup>1</sup>, A.I. Ermakova<sup>2</sup>

<sup>1</sup> Penza Scientific Research Electrotechnical Institute, Penza, Russia

<sup>2</sup> Penza State University, Penza, Russia

<sup>1</sup> ivan@pnici.penza.ru, <sup>2</sup> aermakova27061992@mail.ru

**Abstract.** *Background.* The purpose of the article is to create a quality criterion for assessing the effectiveness of neural network processing of small samples in multicriteria testing of the independence hypothesis. *Materials and methods.* To date, about 200 statistical criteria for testing the hypothesis of independence of a small sample are known. Using them together can significantly improve the quality of decisions made by the neural network. In this regard, the problem of effective comparative assessment of the achieved quality by one or another neural network arises. *Results and conclusions.* It is proposed to evaluate the quality of a neural network solution by the achieved increase in the sample when analyzing it using a single Pearson-Edleton-Edgeworth criterion of the late 19th century on a large sample in the most difficult situation of testing the independence hypothesis.

**Keywords:** testing the independence hypothesis, correlation coefficient, small samples, neural network combination of a large number of criteria

**For citation:** Ivanov A.I., Ermakova A.I. Assessing the performance quality of multicriteria neural network statistical analysis using the example of testing the hypothesis of independence of small sample data. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):13–18. (In Russ.). doi: 10.21685/2307-4205-2024-3-2

**Постановка задачи**

В конце 19 в. основатели современной математической статистики Пирсон, Эдлтон, Эджуорт для оценки коэффициентов корреляции активно использовали ставшую сегодня классической формулу [1]:

$$r_1(x, y) \approx \sum_{i=1}^{16} \frac{(x_i - M(x)) \cdot (y_i - M(y))}{\sigma(x) \cdot \sigma(y)}, \tag{1}$$

где  $M(\cdot)$  – функционал вычисления математического ожидания малой выборки в 16 опытов;  $\sigma(\cdot)$  – функционал вычисления стандартного отклонения.

К сожалению, классическая формула (1) дает требуемую практикой достоверную вероятность оценок только при больших выборках. При малых выборках в 16 опытов для слабо коррелированных данных  $r \approx 0.0$ , ошибка может достигать значений  $\Delta r = \pm 0.7$ . Столь значительный интервал ошибок недопустим при анализе биометрических данных медицины, биологии, экономики, психологии, педагогики. Сегодня статистические оценки, выполненные на малых выборках, воспринимаются физико-математической общественностью как недостоверные. Последнее сильно влияет на материальные затраты, связанные с обучением следующего поколения исследователей, так как увеличивает срок сбора реальных данных в аспирантуре или адъюнктуре. Так, аспирант биолог за три года обучения в аспирантуре вполне способен вырастить и исследовать 16 кроликов, опираясь только на собственные усилия. При этом к концу своего обучения он столкнется с недоверием коллег к его статистическим оценкам в его диссертации. Ситуация меняется кардинально, если он успеет вырастить и исследовать 160 кроликов, однако при этом сильно увеличиваются материально-технические затраты и время на проведение натурального эксперимента.

Положение кардинально меняется, если отказаться от практики прошлого века, когда исследователи-одиночки старались под свою конкретную задачу создать свой наилучший статистический критерий [1], учитывающий некоторую совокупность граничных условий. В 21 в. парадигма вычислительных оценок изменилась, появилась возможность использовать большие сети искусственных нейронов, т.е. каждому из известных статистических критериев (их более 200 приведено в справочнике [1]). Каждому из известных статистических критериев можно поставить в соответствие свой искусственный нейрон, далее несколько нейронов могут быть объединены в одну нейросеть [2].

Практика показала, что нейросетевое объединение примерно 20 классических статистических критериев позволяет малую исходную выборку реальных данных в 16 опытов виртуально увеличить, добавив к ней еще примерно 20 виртуальных опытов (каждый искусственный нейрон увеличивает выборку примерно на один виртуальный опыт). Соответственно, диссертация аспиранта-биолога могла бы выглядеть более убедительной, если бы он обрабатывал данные, полученные не только от 16 реальных кроликов, а от существенно большей выборки в 36 виртуально-реальных кроликов. Именно это и дает некоторый дополнительный эффект роста доверия к многокритериальной статистической обработке данных в сравнении с однокритериальной обработкой.

Приведенный выше тезис доказан численными экспериментами для достаточно больших сетей, собранных из бинарных искусственных нейронов [3]. При этом на сегодня по реальным данным нейрофизиологов [4] у живых существ нет бинарных естественных нейронов. Все живые существа используют более сложные q-арные естественные нейроны [5].

Живые существа не используют бинарные естественные нейроны в силу того, что они способны работать, выполняя параллельный анализ только двух статистических гипотез [2, 3]. Если искусственные нейроны делать троичными ( $q = 3$ ), то они становятся способными выполнять параллельный анализ сразу трех статистических гипотез. Это приводит к существенному росту доверия для новых нейросетевых решений с трехуровневыми выходными квантователями [6, 7].

Выигрыш от первого шага замены бинарных нейронов на троичные нейроны подтвержден имитационным моделированием. Преимущества троичных нейронов над бинарными очевидны [6, 7], и именно по этой причине природа не использует бинарные нейроны в живых организмах [4, 5]. Однако и троичные естественные нейроны, видимо, используются природой достаточно редко. Предположительно дальнейшее повышение числа состояний выходного квантователя должно приводить к росту качества решений, принимаемых сетями q-арных искусственных нейронов [8, 9]. Рисунок 1 иллюстрирует этот тезис, на нем отображен 9-арный искусственный нейрон, полученный 9-арным квантованием выходных состояний функционала оценки коэффициентов корреляции (1) для малой выборки в 16 опытов.

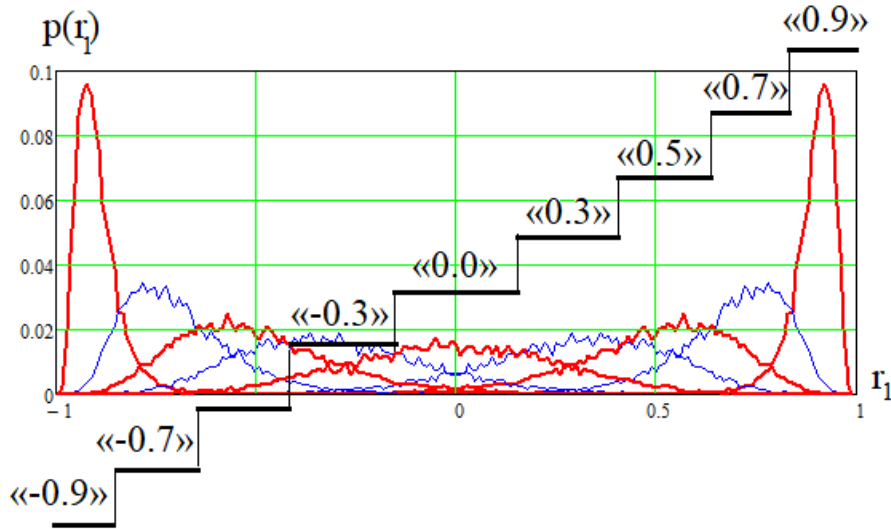


Рис. 1. Распределения плотности вероятности значений –  $p(r_1)$  коэффициентов корреляции, вычисленных по формуле (1) на малой выборке в 16 опытов

Из рисунка видно, что центр каждого стабильного участка 9-арного квантователя искусственного нейрона Пирсона – Эдлтона – Эджуорта совпадает с математическим ожиданием 9 использованных для настройки плотностей распределения. Очевидно, что столь узкие интервалы стабильности или округления данных 9-арного квантователя должны давать значительные ошибки оценок для одного искусственного нейрона. Однако использование десятков подобных искусственных нейронов [3] после усреднения их откликов должно давать достаточно точные итоговые результаты нейровычислений. В этой ситуации возникает задача сравнения данных разных нейросетевых вычислителей корреляции в рамках приведения их данных к одной шкале, соответствующей классической формуле Пирсона – Эдлтона – Эджуорта.

**Эталонная точка шкалы Пирсона – Эдлтона – Эджуорта,  
удобная для сравнения качества работы тестируемых нейросетей**

Из данных рис. 1 следует, что интервал ошибок  $\Delta r$  существенно зависит от математического ожидания контролируемого параметра –  $M(r)$ . Так, для пяти положительных значений математических ожиданий интервалы ошибок и амплитуда ошибок составят:

$$\begin{cases} M(r) = 0,0 \rightarrow -0,7 < \Delta r < +0,7 \rightarrow A(r) \approx 1,4, \\ M(r) = +0,3 \rightarrow -0,3 < \Delta r < +0,8 \rightarrow A(r) \approx 1,1, \\ M(r) = 0,5 \rightarrow +0,0 < \Delta r < +0,85 \rightarrow A(r) \approx 0,85, \\ M(r) = 0,7 \rightarrow +0,3 < \Delta r < +0,95 \rightarrow A(r) \approx 0,65, \\ M(r) = 0,9 \rightarrow +0,6 < \Delta r < +1,0 \rightarrow A(r) \approx 0,4, \end{cases} \quad (2)$$

где  $A(r) = (\max(r) - \min(r))$  – амплитуда или размах интервала ошибок  $\Delta r$ .

При любом конечном объеме выборки реальных данных оценка коэффициентов корреляции по формуле (1) хуже всех (имеет наибольшую амплитуду ошибок) при  $M(r) = 0,0$ . По мере увеличения значений коэффициентов корреляции ошибка его оценок падает (точность увеличивается примерно в три раза).

Несмотря на то, что шкала оценок коэффициентов корреляции Пирсона – Эдлтона – Эджуорта имеет максимальную амплитуду ошибок в точке  $M(r) = 0,0$ , именно эту точку следует выбирать как эталонную. Это связано с тем, что обеспечить попадание в эту точку проще всего. Для этого достаточно двукратного программного обращения к вызову программного генератора псевдослучайных чисел. Существующие программные генераторы дают независимые (некоррелированные) данные. Кроме того, эта точка стабильна и не меняется для любого объема малых выборок.

На рис. 2 представлены результаты численного эксперимента по имитации плотностей распределения коэффициентов корреляции независимых данных для выборок разного объема.

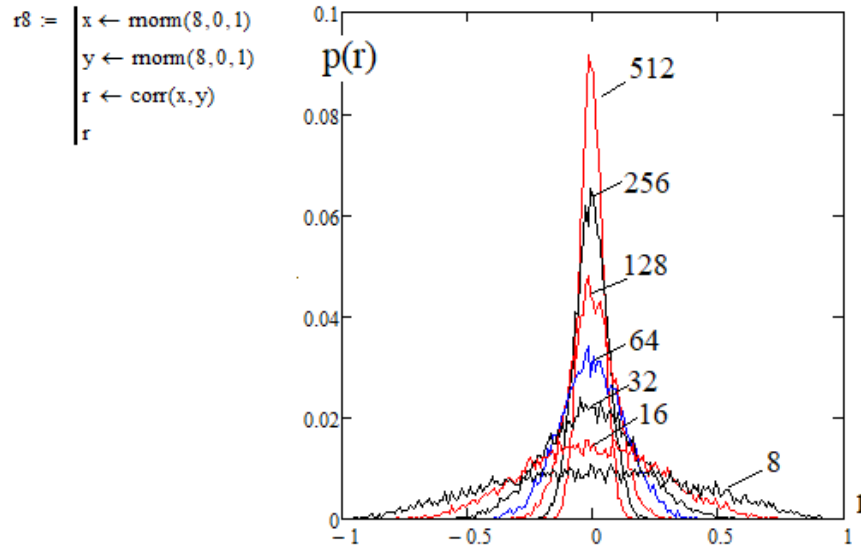


Рис. 2. Распределения значений коэффициентов корреляции для выборок разного объема в 8, 16, 32, 64, 128, 256, 512 опытов

Еще одной причиной выбора точки  $M(r) = 0,0$  при тестировании качества работы нейросетевых вычислителей является то, что именно в этой точке закон плотности распределения значений коэффициентов корреляции близок к нормальному для выборки любого объема. Во всех иных точках закон распределения оказывается асимметричным и, соответственно, уже не описывается двумя первыми статистическими моментами.

В левой части рис. 2 приведена программа на языке MathCAD, позволившая получить данные, отображенные в правой части рисунка.

Только в одной точке  $M(r) = 0,0$  мы имеем самое простое статистическое описание распределения данных, где стандартное отклонение  $\sigma(r)$  является простой функцией, описывающей как связь объема выборки с интервалом неопределенности  $\Delta r$ , так и связи объема выборки с амплитудой интервала неопределенности при доверительной вероятности 0,99:

$$\begin{cases} \Delta r \approx 6\sigma(r), \\ A(r) \approx 6\sigma(r). \end{cases} \quad (3)$$

Оба статистических параметра совпадают  $\Delta r = A(r)$  только в точке  $M(r) = 0,0$ . Формально по любому из статистических параметров системы уравнений (3) может быть найден объем эквивалентной входной выборки. Графики связи стандартного отклонения с объемом выборки приведены на рис. 3.

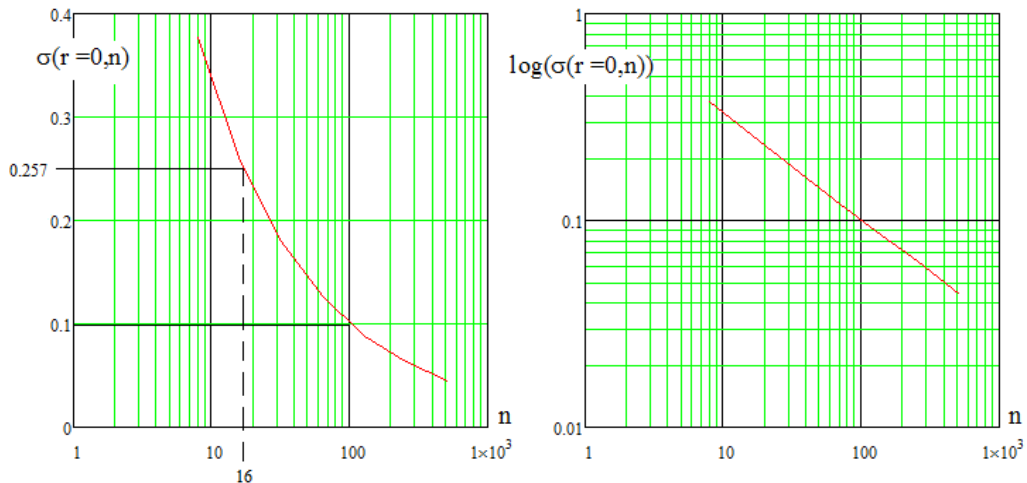


Рис. 3. Обратная оценка объема выборки по значению стандартного отклонения в двух типах систем координат (полулогарифмической системе и полностью логарифмической)

Приведенными на рис. 3 номограммами (как правой, так и левой) следует пользоваться для решения обратной задачи пересчета стандартного отклонения в размер эквивалентной выборки. В качестве примера рассмотрим ситуацию стандартного отклонения  $\sigma(r = 00, n = 16) = 0,257$ . Для этой ситуации мы имеем дело с простейшим нейросетевым анализатором с одним нейроном Пирсона – Эдлтона – Эджуорта, анализирующим малую выборку объемом в 16 опытов.

Однако, если мы наблюдаем ситуацию  $\sigma(r = 00, n = 16) = 0,1$ , то по номограмме рис. 3 мы должны сделать вывод о том, что для одного эталонного нейрона Пирсона – Эдлтона – Эджуорта малая выборка должна иметь объем  $N = 100$  вместо реальной выборки объемом  $n = 16$ . Это вполне возможно из-за добавления в сеть нескольких дополнительных нейронов. Даже если мы не знаем число добавленных дополнительных нейронов и их тип, мы должны прийти к выводу, что тестируемая нейросеть по сравнению с эталонной формулой (1) снизила вероятность ошибок в 2,57 раза, что эквивалентно увеличению объема входной выборки с 16 до 100 опытов. При этом 16 опытов являются реальными, а дополнительные 84 виртуальными. Дополнительные виртуальные опыты появляются из-за более глубокой нейросетевой обработки исходных данных.

Также следует отметить, представленные на рис. 3 зависимости аналитически описываются в разных формах. Зависимость в левой части рисунка хорошо приближается квадратичным полиномом. Зависимость в правой части удобна тем, что хорошо описывается линейно. Параллельно с зависимостями  $\sigma(r = 00, n)$  и  $\log(\sigma(r = 00, n))$  могут быть использованы аналогичные зависимости  $A(r = 00, n)$  и  $\log(A(r = 00, n))$  в соответствии с соотношением (3). Использование двух зависимостей позволяет поднять точность оценок через их усреднения. Корреляционная связь для этих двух разных статистических параметров отсутствует.

### Список литературы

1. Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников. М. : ФИЗМАТЛИТ, 2006. 816 с.
2. Иванов А. И. Искусственные математические молекулы: повышение точности статистических оценок на малых выборках (программы на языке MathCAD) : препринт. Пенза : Из-во ПГУ, 2020. 36 с.
3. Иванов А. И. Нейросетевой многокритериальный статистический анализ малых выборок. Проверка гипотезы независимости : справочник. Пенза : Изд-во ПГУ, 2022. 218 с.
4. Николлс Д., Мартин Р., Валлас Б., Фукс П. От нейрона к мозгу / пер. с англ. П. М. Балабана, А. В. Галкина, Р. А. Гиниатуллина [и др.]. М. : Едиториал УРСС, 2003. 672 с.
5. Иванов А. И. Малые выборки, нейроморфные вычисления: быстрые алгоритмы оценки энтропии Шеннона – Пирсона квадратичной сложности : справочник. Пенза : Изд-во ПГУ, 2023. 32 с.
6. Технический отчет для ТК 164. «Защищенные приложения искусственного интеллекта: модификация алгоритма автоматического обучения бинарных перцептронов по ГОСТ Р 52633.5-2011 под троичные искусственные нейроны». Пенза : Пензенский государственный университет, 2023. Инв. № 1/16-23 от 06.07.23.
7. Иванов А. И., Юнин А. П., Бояршинов М. А., Иванов А. П. Перспектива совместного использования двоичных и троичных искусственных нейронов при анализе качества «белого» шума в пространстве сверток Хэмминга, вычисленных по разным модулям // Интеллектуальные системы в производстве. 2022. Т. 20, № 3. С. 88–92. doi: 10.22213/2410-9304-2022-3-88-93
8. Иванов А. И., Иванов А. П., Савинов К. Н., Еременко Р. В. Виртуальное усиление эффекта распараллеливания вычислений при переходе от бинарных нейронов к использованию Q-арных искусственных нейронов // Надежность и качество сложных систем. 2022. № 4. С. 89–97.

### References

1. Kobzar' A.I. *Prikladnaya matematicheskaya statistika. Dlya inzhenerov i nauchnykh rabotnikov = Applied mathematical statistics. For engineers and scientists*. Moscow: FIZMATLIT, 2006:816. (In Russ.)
2. Ivanov A.I. *Iskusstvennye matematicheskie molekuly: povyshenie tochnosti statisticheskikh otsenok na malykh vyborokakh (programmy na yazyke MathCAD): preprint = Artificial mathematical molecules: improving the accuracy of statistical estimates on small samples (programs in MathCAD) : preprint*. Penza: Iz-vo PGU, 2020:36. (In Russ.)
3. Ivanov A.I. *Neyrosetevoy mnogokriterial'nyy statisticheskiy analiz malykh vyborok. Proverka gipotezy nezavisimosti: spravochnik = Neural network multicriteria statistical analysis of small samples. Testing the independence hypothesis: handbook*. Penza: Izd-vo PGU, 2022:218. (In Russ.)
4. Nikolls D., Martin R., Vallas B., Fuks P. *Ot neyrona k mozgu: per. s angl. P.M. Balabana, A.V. Galkina, R.A. Giniyatullina [i dr.] = From neuron to brain : translated from English by P.M. Balaban, A.V. Galkin, R.A. Giniyatullina [et al.]*. Moscow: Editorial URSS, 2003:672. (In Russ.)

5. Ivanov A.I. *Malye vyborki, neyromorfnye vychisleniya: bystrye algoritmy otsenki entropii Shennona – Pirsona kvadratichnoy slozhnosti: spravochnik = Small samples, neuromorphic calculations: fast algorithms for estimating Shannon–Pearson entropy of quadratic complexity: handbook*. Penza: Izd-vo PGU, 2023:32. (In Russ.)
6. *Tekhnicheskyy otchet dlya TK 164. «Zashchishchennyye prilozheniya iskusstvennogo intellekta: modifikatsiya algoritma avtomaticheskogo obucheniya binarnyykh perseptronov po GOST R 52633.5-2011 pod troichnye iskusstvennye neyrony» = Technical report for TC 164. "Protected applications of artificial intelligence: modification of the algorithm for automatic training of binary perceptrons according to GOST R 52633.5-2011 for ternary artificial neurons"*. Penza: Penzenskiy gosudarstvennyy universitet, 2023;(Inv. № 1/16-23 ot 06.07.23). (In Russ.)
7. Ivanov A.I., Yunin A.P., Boyarshinov M.A., Ivanov A.P. The prospect of joint use of binary and ternary artificial neurons in the analysis of the quality of "white" noise in the space of Hamming convolutions calculated using different modules. *Intellektual'nye sistemy v proizvodstve = Intelligent systems in production*. 2022;20(3):88–92. (In Russ.). doi: 10.22213/2410-9304-2022-3-88-93
8. Ivanov A.I., Ivanov A.P., Savinov K.N., Eremenko R.V. Virtual enhancement of the effect of parallelization of calculations during the transition from binary neurons to the use of Q-ary artificial neurons. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2022;(4):89–97. (In Russ.)

### Информация об авторах / Information about the authors

#### **Александр Иванович Иванов**

доктор технических наук, профессор,  
научный консультант,  
Пензенский научно-исследовательский  
электротехнический институт  
(Россия, г. Пенза, ул. Советская, 9)  
E-mail: ivan@pniei.penza.ru

#### **Aleksandr I. Ivanov**

Doctor of technical sciences, professor,  
scientific consultant,  
Penza Scientific Research Electrotechnical Institute  
(9 Sovetskaya street, Penza, Russia)

#### **Анна Игоревна Ермакова**

преподаватель кафедры радио- и спутниковой связи,  
Военный учебный центр,  
Пензенский государственный университет  
(Россия, г. Пенза, ул. Красная, 40)  
E-mail: aermakova27061992@mail.ru

#### **Anna I. Ermakova**

Lecturer of the sub-department of radio  
and satellite communications,  
Military Training Center,  
Penza State University  
(40 Krasnaya street, Penza, Russia)

**Авторы заявляют об отсутствии конфликта интересов /**

**The authors declare no conflicts of interests.**

**Поступила в редакцию/Received 10.06.2024**

**Поступила после рецензирования/Revised 05.07.2024**

**Принята к публикации/Accepted 30.08.2024**