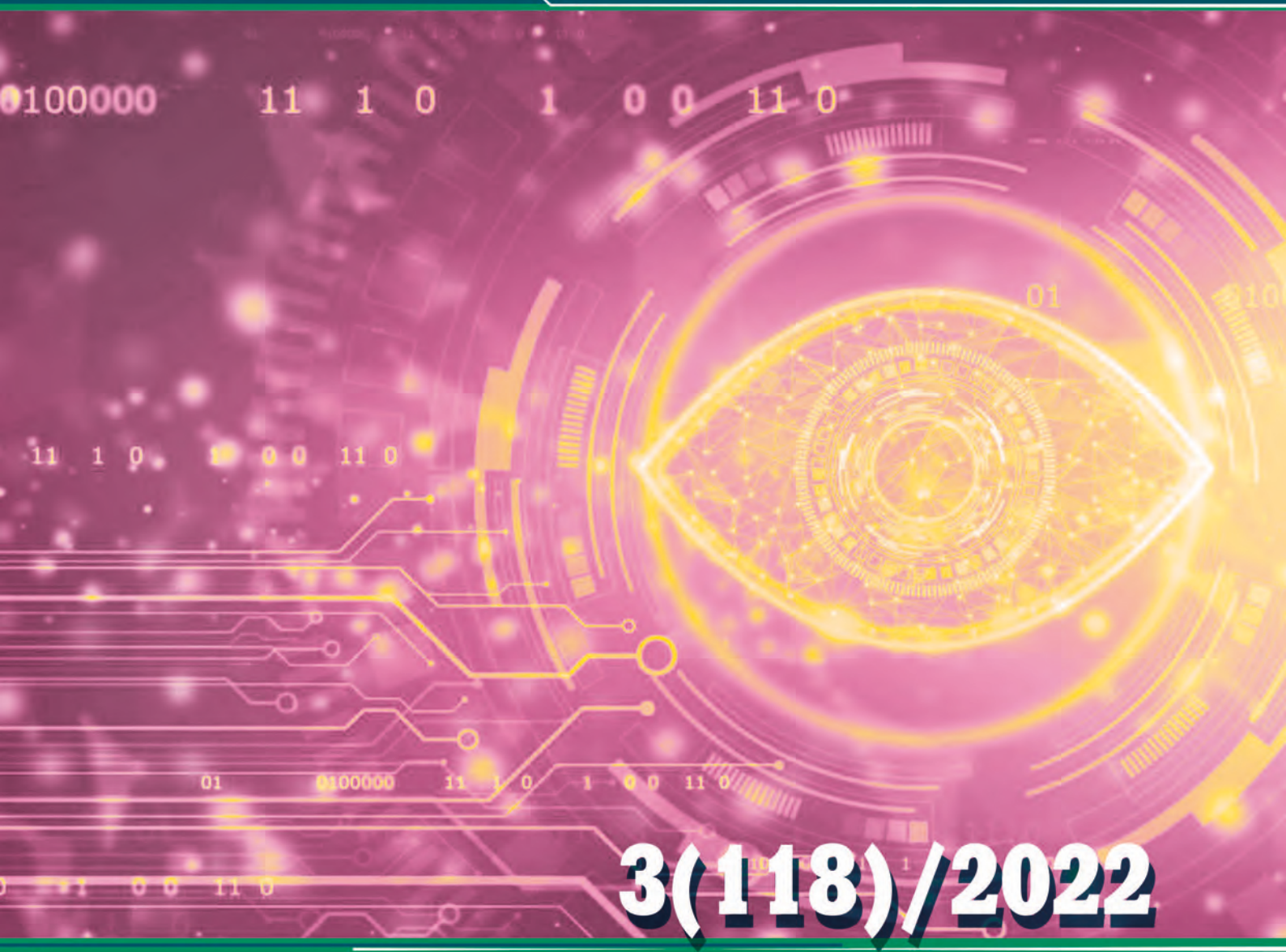
ISSN 1684-8853 (print); ISSN 2541-8610 (online)

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ



3(118)/2022

doi:10.31799/1684-8853-2022-3

ISSN 1684-8853 (print); ISSN 2541-8610 (online)

SCIENTIFIC JOURNAL

3(118)/2022

PEER REVIEWED JOURNAL

INFORMATSIONNO-UPRAVLIAIUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Contact information
The Editorial and Publishing Center, SUAI
67A, Bol'shaya Morskaia, 190000, Saint Petersburg, Russia
Website: http://i-us.ru/en, e-mail: ius.spb@gmail.com
Tel.: +7 - 812 494 70 02

The Journal was registered in the Ministry of Press, Broadcasting and Mass Media of the Russian Federation.
Registration Certificate JD № 77-12412 from April, 19, 2002.
Re-registration in the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR) due to change of the founder A. Vostrikov,
PI № FS 77-82226 from November, 23, 2021

Submitted for publication 05.05.22. Passed for printing 23.06.22. Signed for printing: 27.06.2022.
Format $60\times84_{1/8}$.Phototype SchoolBookC. Digital printing.

Layout original is made at the Editorial and Publishing Center, SUAI.
67A, Bol'shaya Morskaia, 190000, Saint Petersburg, Russia
Printed from slides at the Editorial and Publishing Center, SUAI.
67A, Bol'shaya Morskaia, 190000, Saint Petersburg, Russia

The journal is indexed in Scopus.

Free distribution.

doi:10.31799/1684-8853-2022-3 ISSN 1684-8853 (print); ISSN 2541-8610 (online)

НАУЧНЫЙ ЖУРНАЛ

3(118)/2022

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Учредитель
А. А. Востриков
Издатель
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Адрес редакции: 190000, г. Санкт-Петербург,
ул. Большая Морская, д. 67, лит. А, ГУАП, РИЦ
Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,
сайт: http://i-us.ru

Журнал входит в БД Scopus и в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук.

Сдано в набор 05.05.22. Подписано в печать 23.06.22. Дата выхода в свет: 27.06.2022.
Формат 60×84 1/8. Гарнитура SchoolBookC. Печать цифровая.
Усл. печ. л. 8,3. Уч.-изд. л. 11,2. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 305.

Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.
190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.

Отпечатано в редакционно-издательском центре ГУАП.
190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.

Распространяется бесплатно.

Журнал зарегистрирован в Министерстве РФ по делам печати, телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.
Перерегистрирован в Роскомнадзоре.
Свидетельство о регистрации ПИ № ФС77-82226 от 23 ноября 2021 г.

УДК 519.614
doi:10.31799/1684-8853-2022-3-2-8

Научные статьи
Articles

Матрицы Адамара как результат произведения Скарпи без циклического смещения блоков

Н. А. Балонин[a], *доктор техн. наук, профессор, orcid.org/0000-0001-7338-4920, korbendfs@mail.ru*
А. М. Сергеев[a], *канд. техн. наук, доцент, orcid.org/0000-0002-4788-9869*
[a]*Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ*

Введение: *ортогональные матрицы Адамара, состоящие из элементов 1 и –1 (вещественное число), существуют для порядков, кратных 4. Рассматривается произведение ортогональной матрицы Адамара на ее основу (core), получившее название произведения Скарпи и близкое по смыслу к произведению Кронекера.* ***Цель:*** *выявлением симметрий блочных матриц Адамара показать, что соблюдение их способствует произведению, обобщающему метод Скарпи на случай отсутствия конечного поля.* ***Результаты:*** *показано, что ортогональность является инвариантом рассматриваемого произведения при соблюдении двух условий: один из сомножителей вставляется в другой с учетом знака элементов второго сомножителя (произведение Кронекера), но с выборочным действием знака на элементы и, главное, с циклическим смещением основы, зависящим от места вставки. Выявлено, что таких смещений можно избежать совсем при использовании симметрий, характерных для универсальных форм матриц Адамара. Кроме того, такой прием является общим для многих разновидностей корректируемых произведений Кронекера.* ***Практическая значимость:*** *ортогональные последовательности и методы их эффективного нахождения теорией конечных полей и групп имеют непосредственное практическое значение для задач помехоустойчивого кодирования, сжатия и маскирования видеоинформации.*

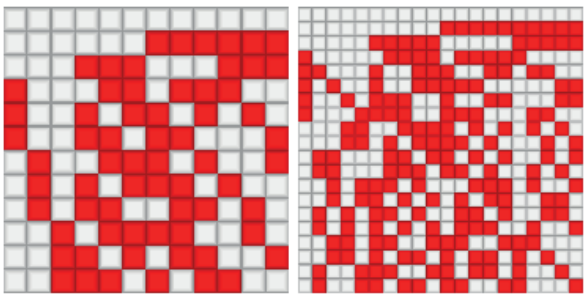
Ключевые слова *– матрицы Адамара, матрицы Мерсенна, произведение Скарпи, кососимметрические матрицы, симметрии матриц.*

Для цитирования: Балонин Н. А., Сергеев А. М. Матрицы Адамара как результат произведения Скарпи без циклического смещения блоков. *Информационно-управляющие системы*, 2022, № 3, с. 2–8. doi:10.31799/1684-8853-2022-3-2-8
For citation: Balonin N. A., Sergeev A. M. Hadamard matrices as a result of Scarpis product without cyclic shifts. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 2–8 (In Russian).

Введение

Матрицы Адамара — это матрицы порядка $n = 4t$ (помимо стартового порядка 2) с элементами 1 и –1, ортогональные в смысле $\mathbf{H}^{\mathrm{T}}\mathbf{H} = n\mathbf{I}$, где $\mathbf{I}$ — единичная матрица [1, 2]. Почти сразу вслед за вычислением Адамаром двух матриц порядков 12 и 20, отличающихся от степеней двойки, возник вопрос об их симметрировании как более экономной форме представления.

На портретах матриц Адамара (рис. 1) красное поле соответствует элементу со значением –1, а белое — единице.



■ ***Рис. 1.*** Портреты матриц Адамара порядков 12 и 20
■ ***Fig. 1.*** Hadamard matrix portraits of orders 12 and 20

Как известно, «кронекерово умножение двух матриц $\mathbf{A}$ и $\mathbf{B}$ с элементами 1 и –1 реализуется вставкой матрицы $\mathbf{B}$ по месту элементов матрицы $\mathbf{A}$ с наследованием знака замещаемого элемента» [3] в следующем виде:

$$\mathbf{A}\otimes\mathbf{B}=\begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \ldots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \ldots & a_{nn}\mathbf{B} \end{pmatrix}.$$

Результатом умножения, например, двух матриц Адамара порядков n и m будет матрица Адамара порядка nm. Именно произведение Кронекера первооснователями направления исследований Сильвестром и Адамаром использовалось для вычисления ортогональных матриц увеличенных порядков.

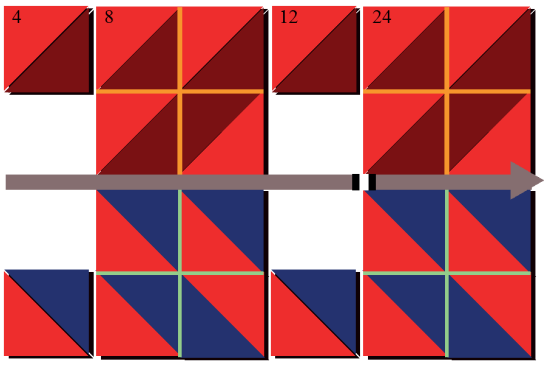
В мире таблиц, которыми являются матрицы Адамара, существуют два ярко выраженных симметричных начала: матрицы симметричные и кососимметричные (здесь и далее — с точностью до диагонали). Разновидность правила Сильвестра

$$\begin{pmatrix} \mathbf{H} & \mathbf{H} \\ \mathbf{H} & -\mathbf{H} \end{pmatrix} \text{ и } \begin{pmatrix} \mathbf{H} & \mathbf{H} \\ -\mathbf{H}^{\mathrm{T}} & \mathbf{H}^{\mathrm{T}} \end{pmatrix}$$

позволяет распространить тип симметрии на удвоенные порядки (рис. 2), так что достаточно рассмотреть порядки $n = 4 + 8t = 4, 12, 20, ...$, идущие с шагом 8, при котором четверть n — нечетное число.

Далеко не сразу было осознано [4], что оба этих типа сосуществуют вдоль всей оси допустимых порядков одновременно. Для наглядности наших рассуждений на рис. 3, *а* и *б* матрицы приведены рядом с удвоенными основами.

Так получилось, что универсальные орнаменты, в которых можно найти матрицу Адамара независимо от ее порядка, были исследованы позднее. Наиболее простая форма — двоякосимметричная — содержит переставляемые местами

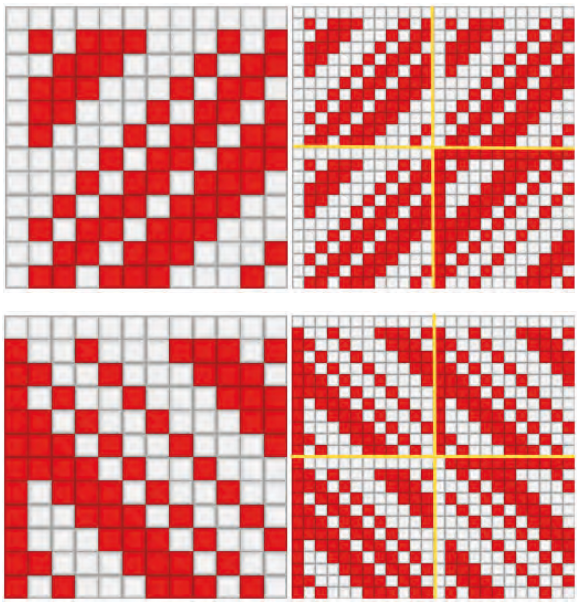


■ ***Рис. 2.*** Распространение обеих симметрий удвоением порядка

■ ***Fig. 2.*** Extending both symmetries by doubling the order

а)

б)



■ ***Рис. 3.*** Симметричные (*а*) и кососимметричные (*б*) матрицы Адамара

■ ***Fig. 3.*** Symmetric (*а*) and skew (*б*) Hadamard matrices

симметричный и кососимметричный блоки (матрицы) и парную кайму, тоже перестраиваемую. Эту форму легко найти при помощи полей Галуа [5, 6] на случай, если $n - 1$ представляет собой простое число или его степень. В некотором смысле она сопровождает простые числа и порождает легко находимые универсальные матрицы, отвечающие им.

Целью настоящей работы является представление новой реализации произведения Скарпи без смещений как основы упрощенного вычисления матриц Адамара.

Блочные конструкции матриц Адамара

Отсутствие поля делает невозможным совмещение обеих симметрий, но для симметрирования матрицы в целом это не обязательно. Косвенное свидетельство этому разработано в теории групп Ито, согласно которой есть взаимно однозначное соответствие между матрицами Адамара и дициклическими группами [6, 7]. Для нахождения симметричной или кососимметричной матрицы в универсальной форме без каймы число ее блоков — матриц порядка $v = n/4$ — увеличивается до четырех: **A**, **B**, **C**, **D**. Для матриц порядков, идущих с шагом 8: $n = 4 + 8t = 4(2t + 1)$, размер блока $v = 2t + 1$ — нечетное число.

Условие ортогональности дает квадратичное уравнение связи $w^2 + x^2 + y^2 + z^2 = n$, регламентирующее число −1 в них: $k_1 = (v - w)/2$, $k_2 = (v - x)/2$, $k_3 = (v - y)/2$, $k_4 = (v - z)/2$ [8].

Для кососимметричного (более простого) варианта матрицы $\mathbf{A} - \mathbf{I} = (\mathbf{I} - \mathbf{A})^{\mathrm{T}}$ параметр $w = 1$ зажат этим видом симметрии, так что $k_1 = (v - 1)/2$, что сводит уравнение орнамента к уравнению сферы $x^2 + y^2 + z^2 = n - 1$. Для симметричного варианта решения $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$ и $\mathbf{B} = \mathbf{C}$ (менее жестко $k_2 = k_3$, но это редко востребуемо) поменяв местами обозначения w и x, связав свободную переменную x с первой матрицей, имеем $w = y$, что сразу же дает уравнение сфероида $x^2 + 2y^2 + z^2 = n$.

В теории чисел (не матриц) разрешимостью в целых числах уравнений сферы и сфероида занимались еще Гаусс и Лиувилль. Замена $x^2 = 8\mathrm{T}_x + 1$, $y^2 = 8\mathrm{T}_y + 1$, $z^2 = 8\mathrm{T}_z + 1$ упрощает уравнения связи до линейного вида: $\mathrm{T}_x + \mathrm{T}_y + \mathrm{T}_z = t$ и $\mathrm{T}_x + 2\mathrm{T}_y + \mathrm{T}_z = t$, где t — натуральное целое, задающее номер матрицы в серии интересующих нас порядков, идущих с шагом 8. Согласно теореме Гаусса [9], любое целое число t, заданное в виде суммы, всегда разрешимо не более чем тремя треугольными числами — числами, взятыми из последовательности сумм натуральных чисел 0, 1, 3, 6, 10 и т. п. (аддитивный факториал). Лиувилль распространил это правило на второе линейное уравнение, близкое по смыслу [10].

Тем самым матрицы Адамара с их целочисленными элементами 1 и –1 представляют собой матричную иллюстрацию теорем теории чисел Гаусса и Лиувилля. Это строго доказанные факты, дающие четкие гарантии существования не только всех матриц Адамара, как это предположил Пэли [11], но и, что сильнее, всех матриц Адамара в симметричной и кососимметричной форме. Некоторая нерешительность Пэли объясняется тем, что он разбирал частный случай симметрий, зависящих от конечных полей. Но ведь поля существуют не всегда.

Историческое недоразумение состоит в том, что во второй половине XX столетия компьютеры были еще слабы, и восполнять недостатки матриц, которые не могли быть вычислены аппаратом конечных полей Пэли, решили, симметрируя не матрицы Адамара, а их блоки **A**, **B**, **C** и **D**. Эти блоки назвали именем автора этой идеи — матрицами Вильямсона [12]. Его подход представляет собой матричную иллюстрацию к другой хорошо известной теореме Лагранжа о разложимости любого целого числа на сумму не более чем четыре целых чисел.

Это направление мысли обязывает пользоваться четырьмя не симметричными, в общем, блоками в составе не симметричного массива Гетхальса — Зейделя, что уводит в сторону от наиболее простых и интересных нам форм. Основной недостаток внедрения симметричных блоков состоит в том, что пропуски порядков вследствие чрезмерной жесткости произвольно выбранного условия обнаруживаются опытным путем. Теории несовместимости с этим типом симметрии нет.

Недостаток подхода Вильямсона стал очевиден не сразу, но к концу XX века профессор Др. Джокович обнаружил первую не закрытую этим подходом матрицу порядка 140 [13], а специальное исследование уже начала XXI века [14] отметило довольно часто идущие пропуски порядков.

Кососимметричный вариант **G** матриц Адамара впервые возник в работе Дж. Себерри [15], сохранив, в силу инерции мысли, ненужную в этом подходе симметрию блоков **B**, **C** и **D**. Проблема хороших матриц **G** (Good matrices) [16] с симметричными вставками аналогична проблемам использования матриц Вильямсона.

Размер блока 35 благополучно преодолевается, но уже на порядке 100 (размер блока 25) на сфере обнаруживается точка Гаусса, не обслуживаемая такой симметрией. Проблема несовместимости лишь откладывается до матрицы порядка 188, где обнаружено решение с несимметричными свободными блоками.

Симметричная конструкция **P**, называемая массив Балонина и Себерри [17], а для краткости — Пропус, с $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$ и $\mathbf{B} = \mathbf{C}$ возникла как альтернатива довольно долго державшейся моде на кососимметричные массивы, породив ряд симметричных матриц на прежде не раскрытых порядках [18–21]. В этих работах и близких к ним блок **R** встречается записанным по другую сторону от циклической матрицы (тогда первая строка реверсируется и циклически смещается на позицию, чтобы сохранить ортогональность массива в целом):

$$\mathbf{G} = \begin{pmatrix} \mathbf{A} & \mathbf{BR} & \mathbf{CR} & \mathbf{DR} \\ -\mathbf{BR} & \mathbf{A} & \mathbf{RD} & -\mathbf{RC} \\ -\mathbf{CR} & -\mathbf{RD} & \mathbf{A} & \mathbf{RB} \\ -\mathbf{DR} & \mathbf{RC} & -\mathbf{RB} & \mathbf{A} \end{pmatrix};$$

$$\mathbf{P} = \begin{pmatrix} \mathbf{A} & \mathbf{BR} = \mathbf{CR} & \mathbf{CR} = \mathbf{BR} & \mathbf{DR} \\ \mathbf{CR} & \mathbf{RD} & -\mathbf{A} & -\mathbf{RB} \\ \mathbf{BR} & -\mathbf{A} & -\mathbf{RD} & \mathbf{RC} \\ \mathbf{DR} & -\mathbf{RC} & \mathbf{RB} & -\mathbf{A} \end{pmatrix},$$

где **R** — обратная единичная матрица, т. е. матрица с единицами вдоль второй, не главной, диагонали квадрата. Массив Гетхальса — Зейделя не пользуется упрощением вида $\mathbf{B}^{\mathrm{T}}\mathbf{R} = \mathbf{RB}$, характерным для циклических блоков.

В совокупности с теоремами Гаусса и Лиувилля образуется значительно более емкая и компактная теория симметричных и кососимметричных матриц Адамара, тесно опирающаяся на полный сет теорем теории чисел. На настоящий момент обе эти конструкции как факт общественного сознания состоялись. Немаловажно подчеркнуть, что тот же вид симметрии присущ основам (core) этих матриц, т. е. блокам, получаемым отсечением первой строки и первого столбца нормализованной матрицы Адамара (когда они состоят только из 1).

Итак, мы выяснили, что матрицы Адамара бывают двоякосимметричными, сопровождая собой простые числа $n - 1$ (или степени простых чисел) [5], или симметричного типа, и тогда они существуют на всем диапазоне порядков $4t$ в виде симметричных или кососимметричных конструкций.

Состав операций с матрицами существенно отличается в сторону простоты реализации, если матрица симметрирована. Очевидно, что если матрица не приведена к некоторой простой симметричной форме, то ее придется к ней приводить, используя аппарат конечных полей. При этом создается иллюзия, что аппарат этот в корне необходим.

Новый метод Скарпи

Так и произошло с методом [22], который возник едва ли не вместе с аппаратом матриц Адамара. Его автор У. Скарпи обнаружил, что,

помимо произведения Кронекера, когда одна матрица Адамара вставляется в другую с множителем, равным заменяемому элементу матрицы, матрицу **H** можно вставлять в ее основу **M** размера $p = n - 1$. Описание на итальянском языке практически невозможно встретить в научной литературе, а те, кто встретят его, потеряют время на изучение ненужных подробностей. Приведем его в нашей заметно более короткой редакции, структурно похожей на витражи в работе [23]:

$$\mathbf{M}\times\mathbf{H}=\begin{pmatrix} \begin{pmatrix}1 & m_{11}\mathbf{e}^{\mathrm{T}}\\ m_{11}\mathbf{e} & \mathbf{M}\end{pmatrix} & \begin{pmatrix}1 & m_{12}\mathbf{e}^{\mathrm{T}}\\ m_{12}\mathbf{e} & \mathbf{M}\end{pmatrix} & \cdots & \begin{pmatrix}1 & m_{1(n-1)}\mathbf{e}^{\mathrm{T}}\\ m_{1(n-1)}\mathbf{e} & \mathbf{M}\end{pmatrix}\\ \begin{pmatrix}1 & m_{21}\mathbf{e}^{\mathrm{T}}\\ m_{21}\mathbf{e} & \mathbf{M}\end{pmatrix} & \begin{pmatrix}1 & m_{22}\mathbf{e}^{\mathrm{T}}\\ m_{22}\mathbf{e} & \mathbf{TM}\end{pmatrix} & \cdots & \begin{pmatrix}1 & m_{2(n-1)}\mathbf{e}^{\mathrm{T}}\\ m_{2(n-1)}\mathbf{e} & \mathbf{T}^{n-2}\mathbf{M}\end{pmatrix}\\ \cdots & \cdots & \ddots & \cdots\\ \begin{pmatrix}1 & m_{(n-1)1}\mathbf{e}^{\mathrm{T}}\\ m_{(n-1)1}\mathbf{e} & \mathbf{M}\end{pmatrix} & \begin{pmatrix}1 & m_{(n-1)2}\mathbf{e}^{\mathrm{T}}\\ m_{(n-1)2}\mathbf{e} & \mathbf{T}^{n-2}\mathbf{M}\end{pmatrix} & \cdots & \begin{pmatrix}1 & m_{(n-1)(n-1)}\mathbf{e}^{\mathrm{T}}\\ m_{(n-1)(n-1)}\mathbf{e} & \mathbf{T}^{(n-2)(n-2)}\mathbf{M}\end{pmatrix} \end{pmatrix},$$

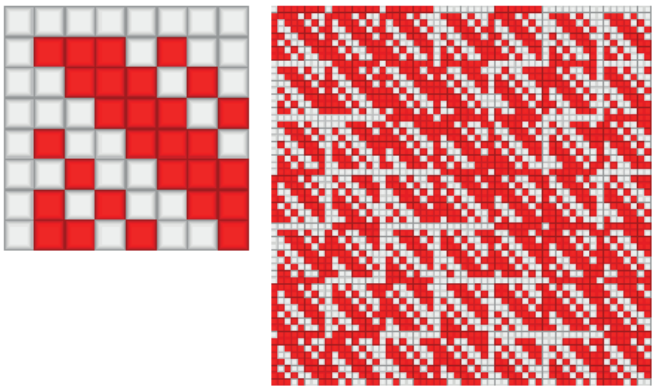
где $\mathbf{H}=\begin{pmatrix}1 & \mathbf{e}^{\mathrm{T}}\\ \mathbf{e} & \mathbf{M}\end{pmatrix}$ является нормальной формой матрицы Адамара; **M** — ее основой, элементы которой используются для модификации знака каймы; **e** — вектор единичных элементов каймы; **T** — матрица циклического смещения.

Пользуясь своим предложением, Скарпи нашел новую матрицу Адамара порядка 56 (7 × 8), не достижимую кронекеровым произведением адамаровых матриц. На рис. 4 в качестве примера приведено решение с матрицей Адамара указанного порядка.

По правилам кронекерова произведения на месте элемента матрицы **M** стоит модифицируемая знаком этого элемента матрица **H**. У Скарпи модификации подвергается только вектор каймы **e**. Этот алгоритм неработоспособен на случай сложных полей $GF(q)$, $q = p^m$, исключительно ввиду краткой формы его записи.

Недостаток, впрочем, легко поправить [24]. Согласно оригинальному алгоритму k-я строка (или столбец) матрицы **M** меняет номер k на величину $k + i \times j$ по mod q. Это общее место теории сложных полей: алгебраические операции ведутся не над тремя числами, а над элементами поля, номера которых эти числа задают. Номер итогового элемента поля служит указанием, куда именно строку (или столбец) перенести.

Теперь обратим особое внимание на то, что алгоритм Скарпи становится работоспособен с *любой* матрицей. Такая производительность излишняя, она не нужна при работе с заранее отсортированной кососимметричной матрицей **H**. Метод полей Галуа, по сути, здесь навязывается, он не нужен для реализации умножения. Оказывается, и в этом состоит смысл нашего нового предложения, секрет более быстрого и более универсального алгоритма заключается в *перестановке сомножителей* с модификацией теперь уже не каймы, а диагонали:



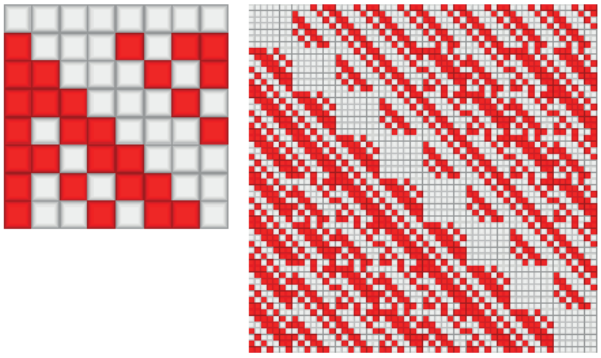
■ ***Рис. 4.*** Портреты матриц Адамара и произведения Скарпи
■ ***Fig. 4.*** Hadamard matrix and Scarpis product portraits

$$\mathbf{H}\times\mathbf{M}=\begin{pmatrix}\mathbf{J} & h_{12}\mathbf{M} & \dots & h_{1n}\mathbf{M}\\ h_{21}\mathbf{M} & \mathbf{J} & \dots & h_{2n}\mathbf{M}\\ \dots & \dots & \ddots & \dots\\ h_{n1}\mathbf{M} & h_{n2}\mathbf{M} & \dots & \mathbf{J}\end{pmatrix},$$

где **J** — матрица из единиц (рис. 5).

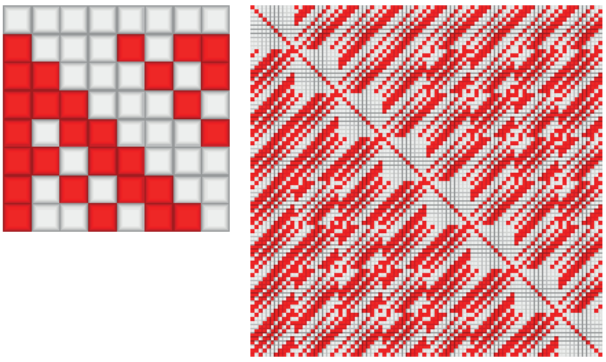
Размер вставляемой в кососимметричную матрицу **H** матрицы **M** не принципиален. Важно лишь, чтобы взаимная разница в размерах не превосходила 4. Если порядок **M** больше **H** на 3, блок коррекции **J** − 2**I** на диагонали теряет парную симметрию, симметрируется и вставляемый блок **MR** или **RM**, как это показано на рис. 6.

Вследствие небольшого рассогласования в размерах сомножителей произведение Кронекера получает небольшой легко поправимый центральными клетками дефект, порождая замечательно простую формулу произведения, дающего все матрицы Адамара независимо от простоты характерного размера сомножителей.



■ ***Рис. 5***. Портреты матриц Адамара и быстрого произведения Скарпи

■ ***Fig. 5***. Hadamard matrix and Scarpis fast product portraits



■ ***Рис. 6***. Портреты матриц Адамара и второго произведения Скарпи

■ ***Fig. 6***. Hadamard matrix and second Scarpis product portraits

В отличие от оригинала, предлагаемая реализация произведения выдерживает, например, работу с кососимметричной матрицей Себерри порядка 36 ($p = n - 1 = 35$, составное число), порождая произведение порядка $35 \times 36 = 1260$. Достижим и порядок $36 \times 37 = 1404$. Как Скарпи мог не заметить более простой формулы вставки в виде витража? Достаточно взглянуть на вид первых матриц Адамара, чтобы понять, что ни выделенных симметрий, ни нормальной и универсальных форм матриц в то время не было.

Заключение

Подход Скарпи при всех недостатках, присущих ранней стадии обнаружения ортогональных конструкций (сложность описания, потребность в арифметике конечных полей), оказал соответствующее влияние на Пэли. Порядки матриц, достижимые витражами, не совпадают с порядками матриц, которые Пэли вычислял прямым применением конечных полей. Отсюда возникло представление, что поиск множества матриц Адамара невозможно закрыть одним каким-либо комбинаторным методом, требуется бесконечное «лоскутное одеяло». Появились матрицы Пэли, матрицы Скарпи, и смысл соревновательности математиков XX века состоял в предложении все новых и новых семейств матриц Адамара.

Апелляция к научному наследию в виде теорем Гаусса и Лиувилля решает проблему существования иным путем. Конечные поля позволяют упростить вычисление ортогональных малоуровневых матриц, но эти поля существуют далеко не всегда. Модифицированное произведение Кронекера в полях для проведения операций не нуждается.

Новый изложенный выше метод позволяет находить матрицы и при отсутствии полей, сохраняя свою основную черту — вставку матрицы Адамара в свою основу или наоборот.

Финансовая поддержка

Статья подготовлена при финансовой поддержке Министерства науки и высшего образования Российской Федерации, соглашение № FSRF-2020-0004.

Литература

1. **Hadamard J.** Résolution d'une question relative aux determinants. *Bulletin des Sciences Mathématiques*, 1893, vol. 17, pp. 240–246.

2. **Jennifer S., Yamada M.** *Hadamard matrices: Constructions using number theory and linear algebra*. Wiley, 2020. 384 p.
3. **Craigen R.** *Hadamard Matrices and Designs*. CRC Handbook of Combinatorial Designs. C. J. Colbourn and J. H. Dinitz eds. CRC Press, 1996. Pp. 229–516.
4. **Colbourn C. J., Dinitz J. H.** *Handbook of Combinatorial Designs*. Second ed. Chapman and Hall/CRC, 2007. 967 p.
5. **Балонин Н. А., Сергеев М. Б.** Критские матрицы Одина и Тени, сопровождающие простые числа и их степени. *Информационно-управляющие системы*, 2022, № 1, с. 2–7. doi:10.31799/1684-8853-2022-1-2-7
6. **Балонин Н. А., Сергеев А. М., Синицына О. А.** Алгоритмы конечных полей и групп поиска ортогональных последовательностей. *Информационно-управляющие системы*. 2021, № 4, с. 2–17. doi:10.31799/1684-8853-2021-4-2-17
7. **Ito N.** On Hadamard Groups III. *Kyushu J. Math*, 1997, no. 51, pp. 369–379.
8. **Балонин Н. А., Сергеев М. Б., Себерри Дж., Синицына О. И.** Окружности на решетках и матрицы Адамара. *Информационно-управляющие системы*, 2019, № 3, с. 2–9. doi:10.31799/1684-8853-2019-3-2-9
9. **Гаусс К. Ф.** *Труды по теории чисел*/ пер. Б. Б. Демьянова; под ред. И. М. Виноградова, комментарии Б. Н. Делоне. М., АН СССР, 1959. 978 с.
10. **Liouville J.** Nouveaux théorémes concernant les nombres triangulaires. *Journal de Mathématiques Pures et Appliquées*, 1863, no. 8, pp. 73–84.
11. **Paley R. E. A. C.** On orthogonal matrices. *Journal of Mathematics and Physics*, 1933, no. 12, pp. 311–320.
12. **Williamson J.** Hadamard's determinant theorem and the sum of four squares. *Duke Math. J.*, 1944, vol. 11, pp. 65–81.
13. **Doković D. Ž.** Williamson matrices of order 4n for n = 33; 35; 39. *Discrete Math.*, 1993, vol. 115, pp. 267–271.
14. **Holzmann W. H., Kharaghani H., Tayfeh-Rezaie B.** Williamson matrices up to order 59. *Designs, Codes and Cryptography*, 2008, no. 46 (3), pp. 343–352.
15. **Seberry J.** A skew-Hadamard matrix of order 92. *Bulletin of the Australian Mathematical Society*, 1971, no. 5, pp. 203–204.
16. **Awyzio G., Seberry J.** On good matrices and skew Hadamard matrices. *Proc. Algebraic Design Theory and Hadamard Matrices*, 2015, pp. 13–28. http://dx.doi.org/10.1007/978-3-319-17729-8_2
17. **Balonin N. A., Seberry J.** Two infinite families of symmetric Hadamard matrices. *Australian Journal of Combinatorics*, 2017, vol. 69(3), pp. 349–357.
18. **Balonin N. A., Balonin Y. N., Đjoković D. Ž., Karbovskiy D. A., Sergeev M. B.** Construction of symmetric Hadamard matrices. *Информационно-управляющие системы*, 2017, № 5, с. 2–11. doi:10.15217/issn1684-8853.2017.5.2
19. **Balonin N. A., Doković D. Ž., Karbovskiy D. A.** Construction of symmetric Hadamard matrices of order 4v for v = 47, 73, 113. *Special Matrices*, 2018, vol. 6, pp. 11–22.
20. **Balonin N. A., Doković D. Ž.** Symmetric Hadamard matrices of orders 268, 412, 436 and 604. *Информационно-управляющие системы*, 2018, № 4, с. 2–8. doi:10.31799/1684-8853-2018-4-2-8
21. **Doković D. Ž.** Some new symmetric Hadamard matrices. arXiv:2101.05429v2. https://arxiv.org/abs/2101.05429.
22. **Scarpis U.** Sui determinanti di valore massimo. *Rendiconti della R. Istituto Lombardo di scienze e lettere*, 1898, no. 31, pp. 1441–1446.
23. **Вострикова А. А.** Матричные витражи и регулярные матрицы Адамара. *Информационно-управляющие системы*, 2021, № 5, с. 2–9. doi.org/10.31799/1684-8853-2021-5-2-9
24. **Doković D. Ž.** Generalization of Scarpis' theorem on Hadamard matrices. *Linear and Multilinear Algebra*, 2017, vol. 65, no. 10, pp. 1985–1987.

UDC 519.614
doi:10.31799/1684-8853-2022-3-2-8

Hadamard matrices as a result of Scarpis product without cyclic shifts

N. A. Balonin[a], Dr. Sc., Tech., Professor, orcid.org/0000-0001-7338-4920, korbendfs@mail.ru
A. M. Sergeev[a], PhD, Tech., Associate Professor, orcid.org/0000-0002-4788-9869
[a]Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation

Introduction: Orthogonal Hadamard matrices consisting of elements 1 and −1 (real number) exist for orders that are multiples of 4. The study considers the product of an orthogonal Hadamard matrix and its core, which is called the Scarpis product, and is similar in meaning to the Kronecker product. **Purpose:** To show by revealing the symmetries of the block Hadamard matrices that their observance contributes to a product that generalizes the Scarpis method to the nonexistence of a finite field. **Results:** The study demonstrates that orthogonality is an invariant of the product under discussion, subject to the two conditions: one of the multipliers is inserted into the other one, the sign of the elements of the second multiplier taken into account (the Kronecker product), but with a selective action of the sign on the elements and, most importantly, with the cyclic permutation of the core which depends on the insertion location. The paper shows that such shifts can be completely avoided by using symmetries that are characteristic of the universal forms of Hadamard matrices. In addition, this technique is common for many varieties of adjustable Kronecker products. **Practical relevance:** Orthogonal

sequences and effective methods for their finding by the theory of finite fields and groups are of direct practical importance for the problems of noiseless coding, video compression and visual masking.

Keywords — Hadamard matrices, Mersenne matrices, Scarpis product, skew-symmetric matrices, symmetric matrices.

For citation: Balonin N. A., Sergeev A. M. Hadamard matrices as a result of Scarpis product without cyclic shifts. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 2–8 (In Russian).

Financial support

The article was prepared with the financial support of the Ministry of Science and Higher Education of the Russian Federation, agreement No. FSRF-2020-0004.

References

1. Hadamard J. Résolution d'une question relative aux déterminants. *Bulletin des Sciences Mathématiques*, 1893, vol. 17, pp. 240–246 (In French).
2. Jennifer S., Yamada M. *Hadamard matrices: Constructions using number theory and linear algebra*. Wiley, 2020. 384 p.
3. Craigen R. *Hadamard matrices and designs*. In: *CRC Handbook of Combinatorial Designs*. C. J. Colbourn and J. H. Dinitz eds. CRC Press, 1996. Pp. 229–516.
4. Colbourn C. J., Dinitz J. H. *Handbook of combinatorial designs*. Second ed. Chapman and Hall/CRC, 2007. 967 p.
5. Balonin N. A., Sergeev M. B. Odin and Shadow Cretan matrices accompanying primes and their powers. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 1, pp. 2–7 (In Russian). doi:10.31799/1684-8853-2022-1-2-7
6. Balonin N. A., Sergeev A. M., Sinitsyna O. I. Finite field and group algorithms for orthogonal sequence search. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 4, pp. 2–17 (In Russian). doi:10.31799/1684-8853-2021-4-2-17
7. Ito N. On Hadamard Groups III. *Kyushu J. Math*., 1997, no. 51, pp. 369–379.
8. Balonin N. A., Sergeev M. B., Seberry J., Sinitsyna O. I. Circles on lattices and Hadamard matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 3, pp. 2–9 (In Russian). doi:10.31799/1684-8853-2019-3-2-9
9. Gauss K. F. *Trudy po teorii chisel* [Works on number theory]. Moscow, Akademiia Nauk SSSR Publ., 1959. 978 p. (In Russian).
10. Liouville J. Nouveaux théorémes concernant les nombres triangulaires. *Journal de Mathématiques Pures et Appliquées*, 1863, no. 8, pp. 73–84 (In French).
11. Paley R. E. A. C. On orthogonal matrices. *Journal of Mathematics and Physics*, 1933, no. 12, pp. 311–320.
12. Williamson J. Hadamard's determinant theorem and the sum of four squares. *Duke Math. J*., 1944, vol. 11, pp. 65–81.
13. Đoković D. Ž. Williamson matrices of order 4n for n = 33; 35; 39. *Discrete Math*., 1993, vol. 115, pp. 267–271.
14. Holzmann W. H., Kharaghani H., Tayfeh-Rezaie B. Williamson matrices up to order 59. *Designs, Codes and Cryptography*, 2008, no. 46 (3), pp. 343–352.
15. Seberry J. A skew-Hadamard matrix of order 92. *Bulletin of the Australian Mathematical Society*, 1971, vol. 5, pp. 203–204.
16. Awyzio G., Seberry J. On good matrices and skew Hadamard matrices. *Proc. Algebraic Design Theory and Hadamard Matrices*, 2015, 15 p. http://dx.doi.org/10.1007/978-3-319-17729-8_2
17. Balonin N. A., Seberry J. Two infinite families of symmetric Hadamard matrices. *Australian Journal of Combinatorics*, 2017, vol. 69(3), pp. 349–357.
18. Balonin N. A., Balonin Y. N., Đoković D. Ž., Karbovskiy D. A., Sergeev M. B. Construction of symmetric Hadamard matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 5, pp. 2–11. doi:10.15217/issn1684-8853.2017.5.2
19. Balonin N. A., Djokovic D. Z., Karbovskiy D. A. Construction of symmetric Hadamard matrices of order 4v for v = 47, 73, 113. *Special Matrices*, 2018, vol. 6, pp. 11–22.
20. Balonin N. A., Đoković D. Ž. Symmetric Hadamard matrices of orders 268, 412, 436 and 604. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 4, pp. 2–8. doi:10.31799/1684-8853-2018-4-2-8
21. Djokovic D. Z. Some new symmetric Hadamard matrices. arXiv:2101.05429v2. https://arxiv.org/abs/2101.05429.
22. Scarpis U. Sui determinanti di valore massimo. *Rendiconti della R. Istituto Lombardo di Scienze e Lettere*, 1898, no. 31, pp. 1441–1446 (In Italian).
23. Vostrikov A. A. Matrix vitrages and regular Hadamard matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 2–9. (In Russia). doi:10.31799/1684-8853-2021-5-2-9
24. Djokovic D. Z. Generalization of Scarpis' theorem on Hadamard matrices. *Linear and Multilinear Algebra*, 2017, vol. 65, no. 10, pp. 1985–1987.

УДК 004.93
doi:10.31799/1684-8853-2022-3-9-19

Научные статьи
Articles

Метод создания синтетических наборов данных для обучения нейросетевых моделей распознаванию объектов

С. Ю. Пчелинцев[а], *аспирант, orcid.org/0000-0001-9195-8318, veselyrojer@mail.ru*
М. А. Ляшков[а], *аспирант, orcid.org/0000-0002-7793-7024*
О. А. Ковалева[а,б], *доктор техн. наук, профессор, orcid.org/0000-0003-0735-6205*
[а]*Тамбовский государственный университет им. Г. Р. Державина, Интернациональная ул., 33, Тамбов, 392000, РФ*
[б]*Тамбовский государственный технический университет, Советская ул., 106, Тамбов, 392000, РФ*

Введение: *недостаток обучающих данных приводит к низкой точности распознавания визуальных образов. Одним из способов решения данной проблемы является использование реальных данных в сочетании с синтетическими.* ***Цель:*** *повышение эффективности распознавания образов системами компьютерного зрения путем использования для обучения смешанных (реальных и синтетических) данных; снижение временных затрат на подготовку данных обучающей выборки.* ***Результаты:*** *на базе предложенного метода генерации синтетических изображений построена интеллектуальная информационная система, позволяющая генерировать репрезентативные выборки большого объема, содержащие изображения, предназначенные для обучения нейронных сетей распознаванию образов. Разработано программно-алгоритмическое обеспечение генератора синтетических изображений для обучения нейросетей. Разработанный генератор имеет модульную архитектуру, что позволяет легко модифицировать, удалять или добавлять отдельные этапы в конвейер генерирования синтетических изображений. Отдельные параметры (как освещение или размытие) для генерируемых изображений можно настраивать. Идея эксперимента заключалась в сравнении точности распознавания образов для нейронной сети, обученной на различных обучающих выборках. Комбинация реальных и синтетических данных при обучении модели показала наилучшую эффективность распознавания. Искусственные обучающие выборки, в которых масштаб фоновых объектов примерно равен масштабу объекта интереса, а количество объектов интереса в кадре выше, оказались эффективнее других искусственных обучающих выборок. Изменение фокусного расстояния камеры в сцене генерации синтетических изображений не оказало влияния на эффективность распознавания.* ***Практическая значимость:*** *предложенный метод генерирования изображений позволяет создать большой набор искусственных данных для обучения нейронных сетей распознаванию образов за меньшее время, чем заняло бы создание такого же набора реальных данных.*

Ключевые слова *— нейронные сети, искусственный интеллект, машинное обучение, синтетические наборы данных, генерирование изображений.*

Для цитирования: Пчелинцев С. Ю., Ляшков М. А., Ковалева О. А. Метод создания синтетических наборов данных для обучения нейросетевых моделей распознаванию объектов. *Информационно-управляющие системы*, 2022, № 3, с. 9–19. doi:10.31799/1684-8853-2022-3-9-19

For citation: Pchelintsev S. Y., Liashkov M. A., Kovaleva O. A. Method for creating synthetic data sets for training neural network models for object recognition. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 9–19 (In Russian). doi:10.31799/1684-8853-2022-3-9-19

Введение

Способность обнаруживать объекты в сложных условиях является фундаментальной для многих задач машинного зрения и робототехники. Современные архитектуры сверточных нейронных сетей, такие как Faster R-CNN, SSD, R-FCN, Yolo9000, YoloV5 и RetinaNet, достигли очень впечатляющих результатов в области распознавания образов. Однако обучение таких моделей с миллионами параметров требует огромного количества маркированных обучающих данных для достижения конкурентных результатов. Очевидно, что создание таких массивных наборов данных стало одним из основных ограничений этих подходов: они требуют участия человека и много времени, очень дороги и подвержены ошибкам.

Обучение с использованием искусственных данных снижает нагрузку, затрачиваемую на сбор данных и их аннотацию [1]. Кроме того, оно решает некоторые проблемы формирования обучающей выборки [2]. Теоретически можно генерировать бесконечное количество обучающих изображений с большими вариациями, где разметка осуществляется автоматически. Процесс генерирования данных называют аугментацией [3]. Обучение с искусственными образцами позволяет точно контролировать рендеринг изображений и, следовательно, различные свойства набора данных [4].

Существует понятие области, или домена (domain), — характеристики способа сбора данных для обучения моделей искусственного интеллекта. Так, в частности, изображения для

обучения распознаванию образов могут быть получены с камеры либо программно генерироваться, как в предлагаемом методе, и относиться, таким образом, к разным областям.

Очевидно, что изображения с камеры и сгенерированные изображения могут и, по существу, должны отличаться. Поэтому модели, обученные на данных, собранных в одной области, обычно имеют низкую точность в других областях. Такое явление называется доменным сдвигом (domain shift), или доменным разрывом (domain gap). Для решения этой проблемы можно повышать реалистичность обучающих данных, смешивать искусственные и реальные данные, использовать архитектуры с предварительно обученными экстракторами признаков или применять трансферное обучение [5].

Предлагаемое в данной работе решение использует рандомизацию доменов (domain randomization) [6]. Суть данного подхода заключается в том, что генерируются заведомо нереалистичные данные, и, таким образом, реальные данные можно рассматривать как частный случай сгенерированных искусственных данных. Совсем недавно эта концепция была расширена за счет добавления реальных фоновых изображений, смешанных со случайными сценами, и дополнительно улучшена за счет фотореалистичного рендеринга [7]. Несмотря на то, что такой подход дал впечатляющие результаты, основным его недостатком по-прежнему остается зависимость от реальных данных. Распространенным подходом к повышению эффективности обнаружения также является расширение реального обучающего набора данных путем добавления искусственных данных. И хотя эти методы демонстрируют значительное улучшение по сравнению с использованием только реальных данных, они по-прежнему требуют как минимум реальных фоновых изображений для конкретной предметной области.

Существует также подход композиции изображения для создания искусственных изображений путем комбинирования вырезанных объектов из разных изображений [8]. Преимущество заключается в использовании данных из одной и той же области, поскольку вырезанные объекты являются копиями реальных изображений и близко соответствуют характеристикам реального мира. Основное ограничение этих подходов состоит в том, что они требуют выполнения громоздкого процесса захвата изображений объектов со всех возможных точек обзора и их маскировки. В частности, эти методы не позволят создавать изображения из разных ракурсов или разных условий освещения, если набор для обучения объекта фиксирован. Это явное ограничение.

Другие направления работы используют фотореалистичный рендеринг и реалистичные композиции сцены для преодоления разрыва в предметной области путем синтеза изображений, максимально приближенных к реальному миру [9]. Хотя эти методы показали многообещающие результаты, они сталкиваются с множеством проблем. Во-первых, создание фотореалистичных обучающих изображений требует сложных механизмов рендеринга, а также значительных вычислительных ресурсов. Во-вторых, реалистичная композиция сцены сама по себе является нетривиальной задачей. В-третьих, современные движки рендеринга, применяемые для создания искусственных сцен, в значительной степени используют преимущества системы человеческого восприятия, чтобы обмануть человеческий глаз. Однако эти уловки не обязательно работают в нейронных сетях, и, следовательно, необходимы дополнительные усилия для преодоления доменного разрыва.

Существуют исследования, в которых были использованы генеративные состязательные сети для дальнейшего преодоления доменного разрыва [10, 11]. Однако такие подходы значительно усложняют работу, поскольку их сложно разработать и обучить. Насколько нам известно, они еще не применялись для задач обнаружения.

Другое направление работ использует адаптацию предметной области или переносное обучение, чтобы преодолеть разрыв между искусственной и реальной предметной областью [12, 13]. Это может быть достигнуто путем объединения двух предикторов, по одному для каждого домена, или путем объединения данных из двух доменов. Адаптация предметной области и переносное (transferred) обучение имеют применения, выходящие далеко за рамки переноса искусственных данных в реальные. Тем не менее они требуют значительного количества реальных данных.

Концепция рандомизации доменов для преодоления доменных разрывов предполагает использование нереалистичных текстур для рендеринга искусственных сцен, чтобы обучить детектор объектов, который обобщается на реальный мир. Другое направление работ [14] объединяет рандомизацию домена и рендеринг фотореалистичного изображения. В нем генерируют два типа данных: во-первых, искусственные изображения со случайными отвлекающими факторами и вариациями, которые кажутся неестественными с реальными фотографиями в качестве фона, и, во-вторых, фотореалистичные визуализации случайно сгенерированных сцен с использованием физического движка для обеспечения физической правдоподобности. Комбинация этих двух типов данных дает значительное улучшение по сравнению с одним источником данных и расши-

ряет области применения сети. Также возможно использование структурированной рандомизации домена, в которой сеть может учитывать контекст. В контексте структурированных сред, таких как уличные сцены, это дает самые современные результаты, но неприменимо к таким сценариям, как выбор предмета из коробки, где нет четких пространственных отношений между расположением различных объектов.

Целью исследования является повышение эффективности распознавания образов системами компьютерного зрения путем использования для обучения смешанных (реальных и синтетических) данных, а также снижение временных затрат на подготовку данных обучающей выборки. Задачами исследования являются:

— разработка метода генерирования синтетических изображений, предназначенных для обучения нейросетевой модели распознаванию образов;

— сравнение эффективности обучения нейронных сетей с применением реальных, синтетических и смешанных данных.

Реализация предлагаемого метода генерирования данных велась в среде Unity 3D с использованием пакета Unity Perception.

Предлагаемый метод генерации искусственных обучающих данных

Метод генерации фона разработан в соответствии с тремя принципами: максимизировать фоновый беспорядок, минимизировать риск отображения в сети одного и того же фонового изображения дважды и создать фоновые изображения из элементов, подобных по масштабу объектам переднего плана. Проведенные эксперименты показывают, что эти принципы помогают создавать обучающие данные, которые позволяют сетям запоминать форму и внешний вид объектов, сводя к минимуму шансы научиться отличать искусственные объекты переднего плана от фоновых объектов просто по различным свойствам, таким как, например, различные размеры объекта или распределение шума.

Суть предлагаемого метода заключается в создании трехмерной сцены в виртуальной среде. На сцену случайным образом добавляются различного рода трехмерные объекты, а также источники освещения. Их параметры изменяются случайным образом в заданных интервалах. Кроме того, добавляются шум и размытие, а также могут меняться внутренние параметры камеры. После всех этих манипуляций осуществляются захват кадра и его сохранение. Потом происходит очистка сцены, и процесс генерирования кадра начинается заново, пока не будет достигнуто целевое количество кадров. На рис. 1 представлен предлагаемый алгоритм генерирования искусственных обучающих данных.

Перед запуском алгоритма программы требуется подготовить трехмерные модели объектов интереса, а также объектов фонового слоя и слоя помех. В качестве 3D-моделей объектов интереса используются модели объектов, поиск которых будет осуществляться системой распознавания образов, обученной на генерируемом наборе данных. Так, для системы, распознающей дорожные знаки, необходимо подготовить 3D-модели распознаваемых знаков. В слоях фоновом и окклюзии (помех) используются одни и те же модели, но масштаб может различаться. Это множество моделей не пересекается со множеством моделей объектов интереса и, по существу, может содержать модели объектов и из других предметных областей (это допустимо в соответствии с используемой концепцией рандомизации домена). Более того, в качестве таких моделей могут выступать

■ ***Рис. 1.*** Алгоритм генерирования изображений
■ ***Fig. 1.*** The algorithm of images generation

даже геометрические примитивы (кубы, шары и т. п.), но все они должны быть текстурированы, наложение текстур на эти объекты осуществляется на этапах создания соответствующих слоев. Но текстуры объектов интереса не изменяются.

Каждая обучающая выборка создается путем смешивания трех слоев изображения: искусственного фонового слоя, слоя объектов переднего плана, построенного в соответствии со стратегией учебной программы, и, наконец, последнего слоя, содержащего преграды.

Фоновый слой создается из текстурированных 3D-моделей M_{bg}, которые не пересекаются с набором объектов переднего плана M_{fg}:

$$M_{bg} \cap M_{fg} = \varnothing. \quad (1)$$

Все трехмерные фоновые модели изначально уменьшены и масштабированы таким образом, чтобы они вписывались в единичную сферу. При создании фона происходит последовательный выбор области на заднем плане, где не был визуализирован другой объект, и визуализация случайного фонового объекта в этой области. Каждый фоновый объект визуализируется в произвольной позе, и процесс повторяется до тех пор, пока весь фон не будет покрыт искусственными фоновыми объектами. Ключом к созданию фона является размер проецируемых фоновых объектов, который определяется по размеру объекта переднего плана. Поэтому мы генерируем рандомизированное изотропное масштабирование S, которое применяем к нашим унифицированным 3D-моделям перед их рендерингом. Мы используем масштабирование для создания объектов таким образом, чтобы размер их проекций на плоскость изображения соответствовал размеру среднего объекта переднего плана. Конкретнее, мы вычисляем диапазон масштабирования $S = [S_{\min}, S_{\max}]$, представляющий масштабы, которые могут применяться к объектам так, что они появляются в пределах [0,9; 1,5] размера, соответствующего среднему размеру объекта переднего плана. Затем для каждого фонового изображения мы создаем случайное подмножество $S_{bg} \subset S$, чтобы гарантировать, что мы создаем не только фоновые изображения с объектами, равномерно распределенными по всем размерам, но также и изображения, в основном, с большими или маленькими объектами. Значение S_{bg} изотропного масштабирования теперь выбирается случайным образом из S, так что размеры фоновых объектов в изображении распределяются равномерно.

Для каждой фоновой сцены дополнительно конвертируем текстуру каждого объекта в пространство HSV. Значение цветового тона H вычисляется по формуле

$$H = \begin{cases} 0^\circ, \Delta = 0 \\ 60^\circ \times \left\{ \dfrac{G-B}{\Delta} \bmod 6 \right\}, C_{\max} = R \\ 60^\circ \times \left\{ \dfrac{B-R}{\Delta} + 2 \right\}, C_{\max} = G \\ 60^\circ \times \left\{ \dfrac{R-G}{\Delta} + 4 \right\}, C_{\max} = B \end{cases}, \quad (2)$$

где R, G, $B \in [0, 1]$ — насыщенности красного, желтого, синего цветов;

$$C_{\max} = \max(R, G, B); \quad (3)$$

$$C_{\min} = \min(R, G, B); \quad (4)$$

$$\Delta = C_{\max} - C_{\min}. \quad (5)$$

Значение насыщенности S вычисляется по формуле

$$S = \begin{cases} 0, C_{\max} = 0 \\ \dfrac{\Delta}{C_{\max}}, C_{\max} \neq 0 \end{cases}. \quad (6)$$

Значение яркости V вычисляется по формуле

$$V = C_{\max}. \quad (7)$$

Затем мы случайным образом изменяем значение оттенка и, наконец, конвертируем оттенок обратно в RGB, чтобы разнообразить фон и обеспечить хорошее распределение цветов фона:

$$(R', G', B') = \begin{cases} (C+m, X+m, m), 0^\circ \le H < 60^\circ \\ (X+m, C+m, m), 60^\circ \le H < 120^\circ \\ (m, C+m, X+m), 120^\circ \le H < 180^\circ \\ (m, X+m, C+m), 180^\circ \le H < 240^\circ \\ (X+m, m, C+m), 240^\circ \le H < 300^\circ \\ (C+m, m, X+m), 300^\circ \le H < 360^\circ \end{cases}, \quad (8)$$

где R', G', $B' \in [0, 1]$ — новые значения насыщенностей красного, желтого, синего;

$$C = V \times S; \quad (9)$$

$$X = C \times \left(1 - \left| \frac{H}{60^\circ} \bmod 2 - 1 \right| \right); \quad (10)$$

$$m = B - C. \quad (11)$$

Этап генерирования объектов переднего плана (рис. 2) является ключевым в работе данного алгоритма.

Для успешного распознавания требуется, чтобы каждый объект интереса присутствовал в обучающей выборке достаточное количество раз и в различных положениях. Для равномерного рас-

■ ***Рис. 2.*** Алгоритм генерирования объектов переднего плана

■ ***Fig. 2.*** The algorithm of foreground objects generation

пределения поз объекта была придумана стратегия, именуемая обучающим планом. Для каждого объекта рекурсивно генерируются 20 вращений таким образом, чтобы множество всех вращений объекта представляло одну из 20 граней выпуклого правильного икосаэдра. Таким образом, каждая вершина представляет собой отдельный вид объекта, определяемый двумя вращениями вне плоскости. Кроме того, мы выбираем расстояние, на котором визуализируем объект переднего плана обратно пропорционально его проецируемому размеру, чтобы гарантировать приблизительное линейное изменение пиксельного покрытия объекта между последовательными уровнями масштабирования. Начинаем с ближайшего к камере расстояния и постепенно переходим к самому дальнему. В результате каждый объект изначально кажется самым большим на изображении, поэтому его легче изучить для сети. В последующих итерациях, удаляясь от камеры, объекты становятся меньше и сложнее для распознавания. Для каждого масштаба перебираем все возможные вращения вне плоскости, а для каждого вращения вне плоскости перебираем все повороты в плоскости. Когда у нас есть масштаб, вращение вне плоскости и в плоскости, перебираем все объекты и визуализируем каждый из них с заданной позой в случайном месте с использованием равномерного распределения. После обработки всех объектов, всех вращений в плоскости и вне плоскости переходим к следующему уровню масштабирования.

Для рендеринга мы разрешаем обрезку объектов переднего плана по границам изображения до 50 %. Кроме того, допускаем перекрытие между каждой парой объектов переднего плана до 30 %. Для каждого объекта случайным образом пытаемся разместить его $n = 100$ раз на сцене переднего плана. Если он не может быть помещен в сцену из-за нарушения ограничений обрезки или перекрытия, прекращаем обработку текущей сцены переднего плана и начинаем со следующей. Для последующей сцены переднего плана начинаем с того места, где остановились в последней сцене.

Также создается слой окклюзии, где случайным объектам из множества объектов, используемых в фоновом слое, разрешается перекрывать объекты переднего плана. Это делается путем определения ограничивающего прямоугольника каждого визуализированного объекта переднего плана и визуализации случайно выбранного закрывающего объекта в однородном случайном месте внутри этого ограничивающего прямоугольника. Затеняющий объект масштабируется случайным образом так, что его проекция покрывает определенный процент соответствующего объекта переднего плана (в диапазоне от 10 до 30 % объекта переднего плана). Поза и цвет закрывающего объекта рандомизируются так же, как и для фоновых объектов.

Имея фон, передний план и слой окклюзии, объединяем все три слоя в одно комбинированное изображение: слой окклюзии визуализируется поверх слоя переднего плана, а результат визуализируется поверх фонового слоя.

Далее добавляем случайные света со случайными искажениями оттенка света, а также с изменяемым направлением источников света.

Наконец, добавляем белый шум и размываем изображение с помощью размытия Гаусса, в котором случайным образом выбираются размер ядра r и стандартное отклонение σ:

$$G(r) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-r^2/(2\sigma^2)}, \qquad (12)$$

где $G(r)$ — функция Гаусса.

Таким образом, фон, передний план и закрывающие части имеют одни и те же свойства изображения, что противоречит подходам, в которых смешиваются реальные изображения и искусственные визуализации. Это делает невозможным для сети отличать передний план от фона только по атрибутам, специфичным для их домена.

Поскольку конечная цель использования сгенерированных данных — поиск экземпляров объекта, требуется обеспечить геометрическую корректность рендеринга наших объектов. Для этого настраиваются внутренние параметры ка-

меры — фокусное расстояние и главная точка. Допускаются небольшие случайные изменения этих параметров.

Эксперименты

В ходе экспериментов фокусируемся на обнаружении и распознавании дорожных знаков, соответствующих ГОСТу [15, 16].

Современные модели обнаружения объектов состоят из экстрактора признаков, который нацелен на проецирование изображений из необработанного пиксельного пространства в многоканальное пространство признаков, и нескольких весов, которые решают различные аспекты проблем обнаружения, такие как сужение ограничительной рамки и классификация. В настоящей работе мы используем популярную архитектуру Faster R-CNN с экстрактором функций InceptionResNet [17]. Веса экстрактора признаков были предварительно обучены на наборе данных ImageNet [18]. Используется общедоступная реализация GoogleFaster R-CNN с открытым исходным кодом [19].

В качестве успешного распознавания в ходе экспериментов засчитывалось 50 %-е пересечение площади рамок прогнозируемого положения объекта с его истинным положением. В качестве метрик для оценки эффективности распознавания в наших экспериментах используются усредненная средняя точность (mean average precision, mAP) и усредненный средний отзыв (mean average recall, mAR).

$$mAP = \frac{1}{K}\sum_{i=1}^{K} AP_i. \qquad (13)$$

Здесь K — количество классов распознаваемых объектов;

$$AP = \int_0^1 p(r)\mathrm{d}r, \qquad (14)$$

где r — это значение отзыва, процента истинно положительных результатов, обнаруженных среди всех истинных фактов; $p(r)$ — соответствующая точность, процент правильных положительных прогнозов.

$$mAR = \frac{1}{K}\sum_{i=1}^{K} AR_i. \qquad (15)$$

Здесь

$$AR = 2\int_{0,5}^1 R(o)\mathrm{d}o, \qquad (16)$$

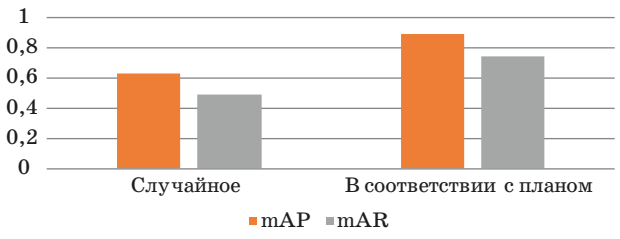
где o — это IoU, уровень перекрытия между реальным и предсказанным системой положением ограничивающей рамки объекта; $R(o)$ — соответствующий отзыв.

В следующих экспериментах мы подчеркиваем преимущества нашей стратегии обучения по учебному плану и исследуем влияние относительного масштаба фоновых объектов по отношению к объектам переднего плана, влияние количества объектов переднего плана, визуализируемых на изображение, влияние композиции фона и, наконец, эффекты случайных цветов и размытия. Модели обучаются с использованием распределенного асинхронного стохастического градиентного спуска.

Данные генерируются в соответствии с учебным планом, который гарантирует, что все модели представлены одинаково в плане позы и в условиях с возрастающей сложностью. В этом эксперименте мы сравниваем две модели Faster R-CNN, инициализированные с одинаковыми весами, первая из которых обучается с использованием полной выборки случайных поз, а другая — в соответствии с нашей стратегией учебного плана. Очевидные преимущества нашего подхода по сравнению со стратегией простой случайной выборки демонстрирует рис. 3.

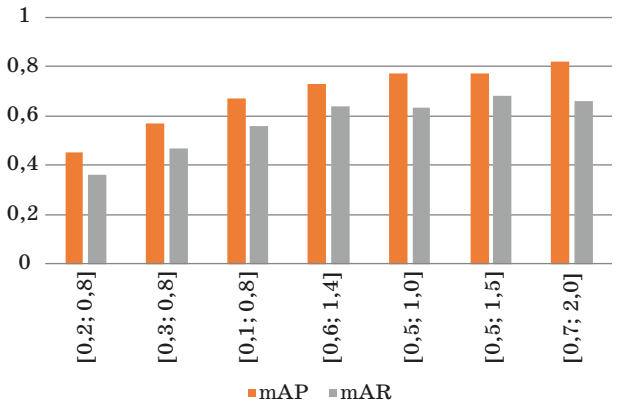
В следующих экспериментах мы анализируем влияние изменения относительного диапазона масштабов фоновых объектов по отношению к объектам переднего плана. На рис. 4 показано, что наилучшие результаты могут быть получены для диапазона, в котором фоновые объекты имеют такой же или больший размер, чем объекты переднего плана. Использование меньших диапазонов масштабирования дает фоновые изображения, которые больше похожи на текстуры, что упрощает сети распознавание объектов переднего плана.

В следующем эксперименте мы изучаем влияние количества объектов переднего плана, отображаемых в обучающих изображениях. Видно (рис. 5), что большее количество объектов переднего плана дает лучшую производительность. Мы устанавливаем только верхний предел количества объектов переднего плана, нарисованных на одном изображении, поэтому среднее количе-



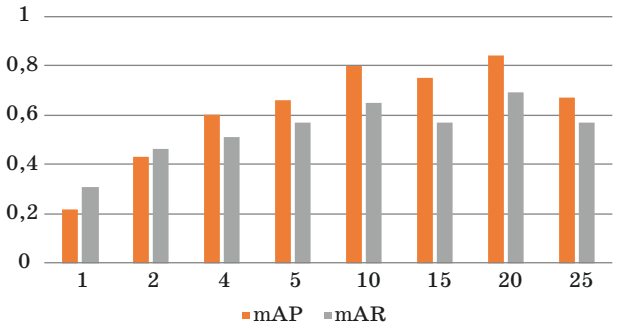
■ ***Рис. 3.*** Сравнение стратегий плана обучения и случайных поз

■ ***Fig. 3.*** Comparison of training plan and random pose strategies



■ ***Рис. 4.*** Сравнение распознаваний с разными масштабами фоновых объектов

■ ***Fig. 4.*** Comparison of recognitions with different background objects scale



■ ***Рис. 5.*** Сравнение распознаваний с разным количеством объектов интереса в кадре

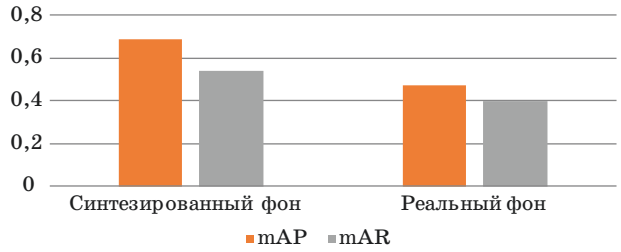
■ ***Fig. 5.*** Comparison of recognitions with different number of objects of interest in one frame

ство объектов обычно ниже. В частности, на начальных этапах изучения учебного плана можно уместить в среднем только 8–9 объектов на одном изображении.

В следующем эксперименте сравнивалось обучение на синтетических данных, в которых фоновый слой был сгенерирован путем добавления множества мелких текстурированных объектов (как и предполагает предложенный алгоритм), с обучением на данных, в которых фон представлял собой реальное изображение, растянутое на весь кадр. Обучение на наборе данных, в котором фоновые слои сгенерированы, показывает лучшие результаты (рис. 6).

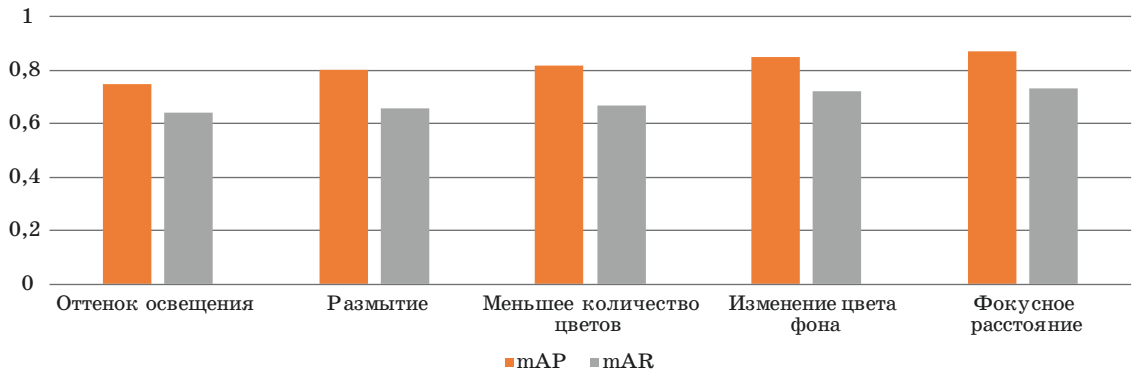
В серии экспериментов (рис. 7) мы исследовали влияние отдельных шагов в конвейере генерации изображений. Было обнаружено, что наибольшее влияние оказывают размытие и случайный оттенок источника света. Наименее важным оказалось изменение фокусного расстояния камеры.

В следующей серии экспериментов мы сравниваем временные затраты на создание реальных и синтетических наборов данных. Весь сбор реальных данных осуществлялся с помощью камеры смартфона. Было отобрано 1200 снимков, вошедших в итоговый набор для обучения распознаванию на реальных данных. Разрешение каждого снимка составляет 1280 × 720 пикселей. На всех этих изображениях содержатся случайные подмножества объектов, подлежащих распознаванию. Фон разный на всех изображениях, освеще-



■ ***Рис. 6.*** Сравнение обучения с полностью искусственным фоном и обучения с реальным фоном на данных в обучающей выборке

■ ***Fig. 6.*** Comparison of training with a completely artificial background and training with a real background on the data in the training sample



■ ***Рис. 7.*** Влияние характеристик изображения на распознавание

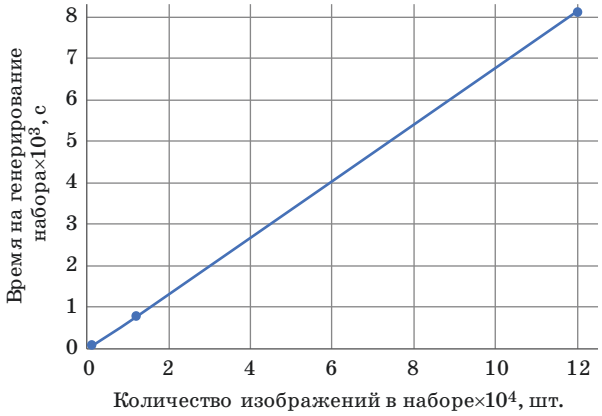
■ ***Fig. 7.*** Influence of image characteristics on recognition

ние тоже отличается, сами фотографии сделаны с различных ракурсов. Это нужно, чтобы соблюсти равномерность данных в тестовой выборке для лучших результатов распознавания.

Разметка данных на изображениях велась вручную с использованием программы VGG Image Annotator. Результат разметки корректировался сторонним наблюдателем, что позволило исправить возникшие в ходе разметки ошибки. Количество времени, затраченного на получение реальных изображений, составило около 10 ч, для маркировки обучающего набора потребовалось примерно 170 ч, а еще 5 ч было потрачено на исправления. Стоит отметить, что для добавления дополнительных реальных данных в набор всегда требуются дополнительные действия по их сбору и разметке, кроме того, будет необходимо создать снимки, сочетающие новые и старые объекты в одном кадре.

Подготовка синтетических данных заняла в общей сложности 5 ч. За это время созданы путем сканирования 3D-модели дорожных знаков — объектов интереса, частично представленных на рис. 8, а также были загружены из открытых источников 3D-модели и текстуры объектов для фонового и помехового слоев. Добавление дополнительных моделей — процесс единоразовый: от пользователя требуется просто добавить модель в проект и запустить генерирование новых данных.

В рамках следующей серии экспериментов измерялось время, затрачиваемое на генерирование набора данных определенного размера. Соотношение количества сгенерированных изображений в наборе ко времени их генерирования представлено на рис. 9. Проводить дополнительную разметку для сгенерированных данных не требуется, поскольку она осуществлялась автоматически на этапе генерирования средствами пакета Unity Perception. Таким образом, с учетом времени на подготовку данных получение набора

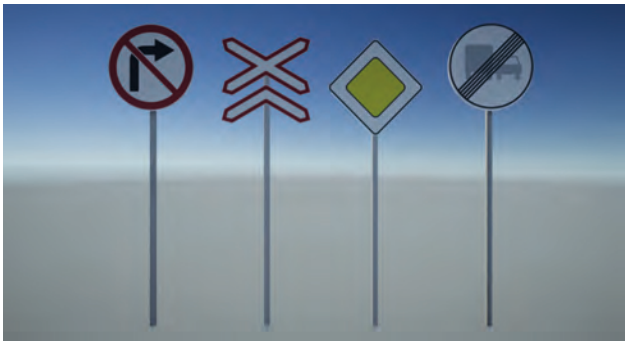


■ ***Рис. 9.*** Зависимость времени генерирования от размера набора данных

■ ***Fig. 9.*** Dependence of generation time on data set size

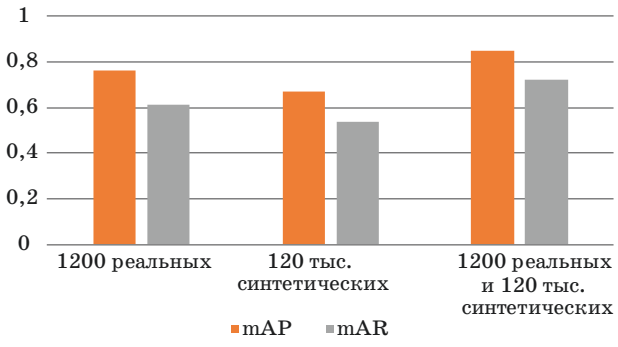
из 1200 реальных изображений заняло 185 ч, что в 37 раз больше, чем время на получение синтетического набора данных аналогичного размера. Российский набор дорожных знаков содержит более 100 тыс. изображений [16]. Генерирование схожего количества изображений, как видно из рис. 9, заняло 2 ч 15 мин (без учета времени, затраченного на подготовку к генерированию). Если и этих данных окажется недостаточно для обучения модели, то можно сгенерировать дополнительные.

В следующем эксперименте мы сравниваем эффективность распознавания при обучении на реальных, синтетических и смешанных данных. Были заняты все 1200 собранных реальных изображений, а также все 120 тыс. изображений, сгенерированных с использованием нашего алгоритма. Как видно по рис. 10, обучение на полностью синтетических данных позволяет распознавать объекты, однако распознавание даже на в стократ меньшем наборе реальных данных может быть эффективнее. Вместе с тем объеди-



■ ***Рис. 8.*** Часть используемых моделей дорожных знаков на одной сцене

■ ***Fig. 8.*** Some of the used road signs models on the same scene



■ ***Рис. 10.*** Сравнение подходов с обучением на реальных данных, на синтетических и на смешанных данных

■ ***Fig. 10.*** Comparison of approaches with training on real data, on synthetic data and on mixed data

■ Характеристики наборов данных
■ Characteristics of datasets

Данные	Количество изображений, шт.	mAP	mAR	Время на формирование
Реальные	1200	0,76	0,61	185 ч
Синтетические	120 000	0,67	0,54	7 ч 15 мин
Смешанные	132 000	0,85	0,72	192 ч 15 мин

нение реальных и синтетических данных в одну обучающую выборку повышает эффективность распознавания.

Данные по времени генерирования наборов данных, их размер и эффективность распознавания с точки зрения характеристик mAP и mAR приведены в итоговой таблице.

Заключение

В работе представлен собственный алгоритм создания искусственных данных для обучения нейронных сетей распознаванию образов. Был использован большой набор трехмерных фоновых моделей, которые плотно визуализированы в частично рандомизированном режиме для создания фоновых изображений. Это позволило создавать локально реалистичные искажения фона, которые делают обученные модели устойчивыми к изменениям окружающей среды. Поверх этих фоновых изображений были визуализированы трехмерные модели интересующих нас объектов. Во время обучения процесс генерации данных следует стратегии обучения, которая гарантирует, что все объекты переднего плана представлены в сгенерированной выборке в равной степени в случайном порядке во всех возможных позах с возрастающей сложностью распознавания, с учетом добавления случайного освещения, размытия и шума. Разработанный подход не требует сложных композиций сцены, создания сложных фотореалистичных изображений или реальных фоновых изображений для обеспечения необходимого фонового беспорядка и хорошо масштабируется для больших наборов исходных данных.

По результатам исследования была разработана «Интеллектуальная система генерирования изображений, предназначенных для обучения нейросетевых моделей распознавания визуальных образов» [20].

Экспериментально доказаны преимущества разработанной стратегии аргументации по сравнению со случайной генерацией поз. В ходе экспериментов установлено, что сгенерированные изображения в идеале должны состоять только из искусственного контента и что все фоновое изображение должно быть заполнено фоновым беспорядком. Проведенные эксперименты также позволили выделить влияние различных факторов, таких как шум или масштаб объектов, на эффективность распознавания.

Выполнено сравнение обучения распознаванию на реальном, на сгенерированном с помощью предложенного решения, а также на смешанном наборах данных. Наилучшую эффективность распознавания показал смешанный набор данных. Предлагаемый метод генерации позволяет создавать большие объемы синтезированных изображений с автоматической разметкой в существенно меньшие сроки, чем заняло бы создание такого же набора из реальных изображений. При этом реальные данные все так же остаются более предпочтительными для обучения моделей. Однако, поскольку их сбор и обработка в достаточном количестве занимают большое количество времени, комбинация синтетических и реальных данных позволяет повысить эффективность обучения модели распознаванию объектов и получить выигрыш по времени формирования обучающих данных.

Литература

1. **Беляева О. В., Перминов А. И., Козлов И. С.** Использование синтетических данных для тонкой настройки моделей сегментации документов. *Тр. ИСП РАН*, 2020, т. 32, № 4, с. 189–202. doi:10.15514/ISPRAS-2020-32(4)-14
2. **Парасич А. В., Парасич В. А., Парасич И. В.** Формирование обучающей выборки в задачах машинного обучения. Обзор. *Информационно-управляющие системы*, 2021, № 4, с. 61–70. doi:10.31799/1684-8853-2021-4-61-70
3. **Konushin A. S., Faizov B. V., Shakhuro V. I.** Road images augmentation with synthetic traffic signs using neural networks. *Computer Optics*, 2021, no. 5, pp. 736–748. doi:10.18287/2412-6179-CO-859
4. **Пчелинцев С. Ю., Ковалева О. А., Суслин А. А.** Использование синтетических данных для обучения нейронных сетей. *Наука. Технология. Производство — 2021: материалы Всерос. науч.-техн. конф.*, Салават, 19–23 апреля 2021 г., Уфа, 2021, с. 8–10.
5. **Каляшов Е. В., Савельева А. А., Тарлыков А. В.** Сегментация реальных объектов с использованием нейронной сети, обученной на синтетических данных. *Актуальные проблемы инфотелекоммуникаций в науке и образовании: VIII Междунар. науч.-техн. и науч.-метод. конф.; сб. науч. ст. в 4 т.*,

Санкт-Петербург, 27–28 февраля 2019 г., СПб., 2019, с. 472–476.

6. **Tobin J., Fong R., Ray A., Schneider J., Zaremba W., Abbeel P.** Domain randomization for transferring deep neural networks from simulation to the real world. *IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS)*, Vancouver, 2017, pp. 23–30. doi:10.1109/IROS37595.2017
7. **Prakash A., Boochoon S., Brophy M., Acuna D., Cameracci E., State G., Shapira O., Birchfield S.** Structured domain randomization: Bridging the reality gap by context aware synthetic data. *Intern. Conf. on Robotics and Automation (ICRA)*, Montreal, 2019, pp. 7249–7255. doi:10.1109/ICRA39644.2019
8. **Dwibedi D., Misra I., Hebert M.** Cut, paste and learn: surprisingly easy synthesis for instance detection. *IEEE Intern. Conf. on Computer Vision (ICCV)*, Venice, 2017, pp. 1310–1319. doi:10.1109/ICCV.2017.146
9. **Richter S. R., Vineet V., Roth S., Koltun V.** Playing for data: Ground truth from computer games. *European Conf. on Computer Vision*, Amsterdam, 2016, pp. 102–118. doi:10.1007/978-3-319-46475-6_7
10. **Bousmalis K., Silberman N., Dohan D., Erhan D., Krishnan D.** Playing for data: Unsupervised pixel-level domain adaptation with generative adversarial networks. *Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 2017, pp. 95–104. doi:10.1109/CVPR.2017.18
11. **Chen B.-C., Kae A.** Playing for data: Toward realistic image compositing with adversarial learning. *Conf. on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 2019, pp. 8407–8416. doi:10.1109/CVPR.2019.00861
12. **Inoue T., Chaudhury S., De Magistris G., Dasgupta S.** Transfer learning from synthetic to real images using variational autoencoders for precise position detection. *Intern. Conf. on Image Processing (ICIP)*, Athens, 2018, pp. 2725–2729. doi:10.1109/ICIP.2018.8451064
13. **Yao T., Pan Y., Ngo C.-W., Li H., Mei T.** Semi-supervised domain adaptation with subspace learning for visual recognition. *Conf. on Computer Vision and Pattern Recognition (CVPR)*, Boston, 2015, pp. 2142–2150. doi:10.1109/CVPR31182.2015
14. **Tremblay J., To T., Sundaralingam B., Xiang Y., Fox D., Birchfield S.** Deep object pose estimation for semantic robotic grasping of household objects. *Conf. on Robot Learning (CoRL)*, Zurich, 2018, pp. 306–316.
15. ГОСТ Р 52289-2019. *Технические средства организации дорожного движения. Правила применения дорожных знаков, разметки, светофоров, дорожных ограждений и направляющих устройств*. М., Стандартинформ, 2020. 134 с.
16. **Шахуро В. И., Конушин А. С.** Российская база изображений автодорожных знаков. *Компьютерная оптика*, 2016, т. 40, № 2, с. 294–300. doi:10.18287/2412-6179-2016-40-2-294-300
17. **Ren S., He K., Girshick R., Sun J.** Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 6, pp. 1137–1149. doi:10.1109/TPAMI.2016.2577031
18. **Krizhevsky A., Sutskever I., Hinton G.** ImageNet classification with deep convolutional neural networks. *Neural Information Processing Systems*, 2012, no. 25, pp. 1097–1105. doi:10.1145/3065386
19. *ImageNet*. https://www.image-net.org (дата обращения: 05.01.2022).
20. Свид. о рег. прогр. для ЭВМ RU 2021666818. *Интеллектуальная система генерирования изображений, предназначенных для обучения нейросетевых моделей распознавания визуальных образов*, С. Ю. Пчелинцев, О. А. Ковалева, С. В. Ковалев. № 2021666314; заявл. 20.10.21; опубл. 20.10.21, Бюл. № 10.

UDC 004.93
doi:10.31799/1684-8853-2022-3-9-19

Method for creating synthetic data sets for training neural network models for object recognition

S. Y. Pchelintsev[a], Post-Graduate Student, orcid.org/0000-0001-9195-8318, veselyrojer@mail.ru
M. A. Liashkov[a], Post-Graduate Student, orcid.org/0000-0002-7793-7024
O. A. Kovaleva[a,b], Dr. Sc., Tech., Professor, orcid.org/0000-0003-0735-6205
[a]Derzhavin Tambov State University, 33, Internatsionalnaya St., 392000, Tambov, Russian Federation
[b]Tambov State Technical University, 106, Sovetskaya St., 392000, Tambov, Russian Federation

Introduction: The lack of training data leads to low accuracy of visual pattern recognition. One way to solve this problem is to use real data in combination with synthetic data. **Purpose:** To improve the performance of pattern recognition systems in computer vision by mixing real and synthetic data for training, and to reduce the time needed for preparing training data. **Results:** We have built an intelligent information system on the basis of the proposed method which allows the generation of synthetic images. The system allows to generate large and representative samples of images for pattern recognition neural network training. We have also developed software for the synthetic image generator for neural network training. The generator has a modular architecture, which makes it easy to modify, remove or add individual stages to the synthetic image generation pipeline. One can adjust individual parameters (like lighting or blurring) for generated images. The experiment was aimed to compare the accuracy of pattern recognition for a neural network trained on different training samples. The combination of real and synthetic data in model training showed the best recognition performance. Artificially generated training samples, in which the scale of background objects is approximately equal to

the scale of the object of interest, and the number of objects of interest in the frame is higher, turned out to be more efficient than other artificially constructed training samples. Changing focal length of the camera in the synthetic image generation scene had no effect on the recognition performance. **Practical relevance:** The proposed image generation method allows to create a large set of artificially constructed data for training neural networks in pattern recognition in less time than it would take to create the same set of real data.

Keywords — neural networks, artificial intelligence, machine learning, synthetic data sets, image generation.

For citation: Pchelintsev S. Y., Liashkov M. A., Kovaleva O. A. Method for creating synthetic data sets for training neural network models for object recognition. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 9–19 (In Russian). doi:10.31799/1684-8853-2022-3-9-19

References

1. Belyaeva O. V., Perminov A. I., Kozlov I. S. Synthetic data usage for document segmentation models fine-tuning. *Proc. of ISP RAS*, 2020, vol. 32, no. 4, pp. 189–202 (In Russian). doi:10.15514/ISPRAS-2020-32(4)-14
2. Parasich A. V., Parasich V. A., Parasich I. V. Training set formation in machine learning tasks. Survey. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 4, pp. 61–70 (In Russian). doi:10.31799/1684-8853-2021-4-61-70
3. Konushin A. S., Faizov B. V., Shakhuro V. I. Road images augmentation with synthetic traffic signs using neural networks. *Computer Optics*, 2021, no. 5, pp. 736–748. doi:10.18287/2412-6179-CO-859
4. Pchelintsev S. Y., Kovaleva O. A., Suslin A. A. Using synthetic data for training neural networks. *Materialy Vseros. nauch.-tekhn. konf. "Nauka. Tekhnologiya. Proizvodstvo — 2021"* [Proc. Vseros. Sci.-Tech. Conf. "Science. Technology. Production — 2021"]. Ufa, 2021, pp. 8–10 (In Russian).
5. Kalyashov E. V., Savel'eva A. A., Tarlykov A. V. Segmentation of real objects with using neural network trained on real data. *Sbornik statej VIII Mezhdunarodnoj nauchno-tekhnicheskoj i nauchno-metodicheskoj konferencii "Aktual'nye problemy infotelekommunikacij v nauke i obrazovanii"* [Proc. VIII Intern. Scien.-Tech. and Scien.-Method. Conf. "Actual problems of infotelecommunication in science and education"]. Saint-Petersburg, 2019, pp. 472–476 (In Russian).
6. Tobin J., Fong R., Ray A., Schneider J., Zaremba W., Abbeel P. Domain randomization for transferring deep neural networks from simulation to the real world. *IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS)*, Vancouver, 2017, pp. 23–30. doi:10.1109/IROS37595.2017
7. Prakash A., Boochoon S., Brophy M., Acuna D., Cameracci E., State G., Shapira O., Birchfield S. Structured domain randomization: Bridging the reality gap by context aware synthetic data. *Intern. Conf. on Robotics and Automation (ICRA)*, Montreal, 2019, pp. 7249–7255. doi:10.1109/ICRA39644.2019
8. Dwibedi D., Misra I., Hebert M. Cut, paste and learn: surprisingly easy synthesis for instance detection. *IEEE Intern. Conf. on Computer Vision (ICCV)*, Venice, 2017, pp. 1310–1319. doi:10.1109/ICCV.2017.146
9. Richter S. R., Vineet V., Roth S., Koltun V. Playing for data: Ground truth from computer games. *European Conf. on Computer Vision*, Amsterdam, 2016, pp. 102–118. doi:10.1007/978-3-319-46475-6_7
10. Bousmalis K., Silberman N., Dohan D., Erhan D., Krishnan D. Playing for data: Unsupervised pixel-level domain adaptation with generative adversarial networks. *Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 2017, pp. 95–104. doi:10.1109/CVPR.2017.18
11. Chen B.-C., Kae A. Playing for data: Toward realistic image compositing with adversarial learning. *Conf. on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 2019, pp. 8407–8416. doi:10.1109/CVPR.2019.00861
12. Inoue T., Chaudhury S., De Magistris G., Dasgupta S. Transfer learning from synthetic to real images using variational autoencoders for precise position detection. *Intern. Conf. on Image Processing (ICIP)*, Athens, 2018, pp. 2725–2729. doi:10.1109/ICIP.2018.8451064
13. Yao T., Pan Y., Ngo C.-W., Li H., Mei T. Semi-supervised domain adaptation with subspace learning for visual recognition. *Conf. on Computer Vision and Pattern Recognition (CVPR)*, Boston, 2015, pp. 2142–2150. doi:10.1109/CVPR31182.2015
14. Tremblay J., To T., Sundaralingam B., Xiang Y., Fox D., Birchfield S. Deep object pose estimation for semantic robotic grasping of household objects. *Conf. on Robot Learning (CoRL)*, Zurich, 2018, pp. 306–316.
15. State Standard 52289-2019. *Traffic control devices. Rules of application of traffic signs, markings, traffic lights, guardrails and delineators*. Moscow, Standartinform Publ., 2020. 134 p. (In Russian).
16. Shakhuro V. I., Konushin A. S. Russian traffic sign images dataset. *Computer Optics*, 2016, no. 2, pp. 294–300 (In Russian). doi:10.18287/2412-6179-2016-40-2-294-300
17. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 6, pp. 1137–1149. doi:10.1109/TPAMI.2016.2577031
18. Krizhevsky A., Sutskever I., Hinton G. ImageNet classification with deep convolutional neural networks. *Neural Information Processing Systems*, 2012, no. 25, pp. 1097–1105. doi:10.1145/3065386
19. *ImageNet*. Available at: https://www.image-net.org (accessed 5 January 2022).
20. Pchelintsev S. Y., et al. *Intellektual'naya sistema generirovaniya izobrazhenij, prednaznachennyh dlya obucheniya nejrosetevyh modelej raspoznavaniya vizual'nyh obrazov* [Intellectual system of generation images aimed at training neural network models for visual images recognition]. Computer program registration certificate Russia, no. 2021666818, 2021.

УДК 621.396
doi:10.31799/1684-8853-2022-3-20-30

Научные статьи
Articles

Адаптивное применение моделей машинного обучения на отдельных сегментах выборки в задачах регрессии и классификации

И. С. Лебедев[a], *доктор техн. наук, профессор, orcid.org/0000-0001-6753-2181, isl_box@mail.ru*
[a]*Санкт-Петербургский Федеральный исследовательский центр РАН, 14-я линия В.О., 39, Санкт-Петербург, 199178, РФ*

Введение: *достижение заданных качественных показателей в решениях, связанных с машинным обучением, зависит не только от эффективности алгоритмов, но и от свойств данных. Одним из направлений развития моделей классификации и регрессии является уточнение локальных свойств информации.* ***Цель:*** *повышение показателей качества при решении задач классификации и регрессии на основе адаптивного выбора различных моделей машинного обучения на отдельных локальных сегментах выборки данных.* ***Результаты:*** *предложен метод, использующий комбинирование различных моделей и алгоритмов машинного обучения на отдельных подвыборках в задачах регрессии и классификации. Метод основывается на вычислении качественных показателей и выборе лучших моделей на локальных сегментах выборки. Выявление изменений данных и временных последовательностей дает возможность сформировать выборки, где данные имеют различные свойства (например, дисперсия, выборочная доля, размах данных и т. д.). Рассмотрено сегментирование на основе алгоритма поиска точек смены тренда временного ряда и аналитической информации. На примере реальных данных датасета приведены экспериментальные значения функции потерь для предлагаемого метода у различных классификаторов на отдельных сегментах и всей выборке.* ***Практическая значимость:*** *результаты могут быть использованы в задачах классификации и регрессии при разработке моделей и методов машинного обучения. Предложенный метод позволяет повысить показатели качества классификации и регрессии за счет назначения моделей, имеющих лучшие показатели на отдельных сегментах.*

Ключевые слова *— машинное обучение, сегментирование множества данных, временные последовательности, изменяющиеся свойства данных.*

Для цитирования: Лебедев И. С. Адаптивное применение моделей машинного обучения на отдельных сегментах выборки в задачах регрессии и классификации. *Информационно-управляющие системы*, 2022, № 3, с. 20–30. doi:10.31799/1684-8853-2022-3-20-30
For citation: Lebedev I. S. Adaptive application of machine learning models on separate segments of a data sample in regression and classification problems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 20–30 (In Russian). doi:10.31799/1684-8853-2022-3-20-30

Введение

Применение технологий искусственного интеллекта в различных областях дает возможность добиваться результатов, сопоставимых с деятельностью человека. Современные подходы, используемые в методах машинного обучения, направлены на автоматическое создание моделей на основе выборок, где обучение происходит непосредственно на данных.

В задачах классификации и регрессии оценивается взаимосвязь между входными и выходными переменными, используются алгоритмы оптимизации, которые минимизируют ошибку аппроксимации. Качество обучения модели зависит от свойств совокупности объектов наблюдений, на которых она обучалась [1, 2]. Основная проблема заключается в том, что ошибка аппроксимации сходится к разумным значениям только с большим объемом данных, которые часто трудно получить, анализировать и интерпретировать. Одновременно с этим возникают ситуации, когда собранные значения показателей системы с течением времени могут менять свои свойства. Изменения распределений, частоты событий, дисбаланс классов приводят к тому, что применение моделей и алгоритмов машинного обучения без учета динамически изменяющихся свойств может существенно влиять на результат, увеличивая ошибки [3].

Обработка информационных потоков

Большинство задач регрессии и предсказания связаны с анализом информационных потоков. Обработка последовательностей и временных рядов имеет определенные особенности. В простейших случаях поступающие данные образуют размеченную выборку, над которой реализуются различные методы обучения [4, 5]. Такие подходы хорошо отражены в классических работах по искусственному интеллекту, в течение длительного времени они подвергались всесторонней оценке и имеют проработанную технологию внедрения и использования. Однако в случае изменения данных и их свойств возникает необходимость поддерживать заданные качественные показатели

алгоритмов, что может являться трудоемкой задачей.

Огромное количество решений задач классификации, регрессии, предсказания поведения системы в динамике использует представление информационных потоков временной последовательностью [6]. Появление моделей и методологий ARMA, ARIMA и других позволило существенно повысить точность прогнозов временного ряда. Однако их построение требует знаний о природе последовательности. Возникает необходимость ее перенастройки при появлении новых данных. Требуются постоянная оценка и подбор различных параметров для достижения заданной точности.

Современная парадигма машинного обучения состоит в том, что модели учатся непосредственно на данных, автоматически вычисляя и оценивая возникающие в выборках закономерности [7]. Поэтому для достижения качественных показателей приходится особое внимание уделять данным.

В работах [8, 9] акцентируется внимание на ряде проблемных вопросов формирования кортежей признаков, создания паттернов поведения, которые подаются на вход классифицирующих алгоритмов. Такие подходы изначально используют статическое представление информации, что не всегда оправдано, особенно для информации, поступающей от реальных систем. Одновременно с этим необходимо решать вопросы длины последовательности; определять характеристики алгоритмов, на основе которых будет производиться разделение выборки; оценивать влияние изменения распределений, частоты событий, дисбаланса классов [10–12].

При обработке данных имеют место вопросы влияния производительности и скорости алгоритмов анализа на качество результатов. В работе [13] предложена оригинальная каскадная модель, элементами которой являются классифицирующие алгоритмы. Однако в ней возрастает сложность агрегации результатов. В ряде других исследований [14–18] отмечается, что особую важность приобретает выделение наиболее информативных признаков, которые вносят основной вклад в задачах классификации и регрессии. Снижение размерности признакового пространства, например методом главных компонент, подсчетом информативности на основе энтропии, частотными методами и т. д., не всегда возможно, в частности, когда имеется одномерный временной ряд [19, 20]. Сокращение размерности признакового пространства позволяет повысить скорость обработки, но с течением времени в случае возникновения эффекта «дрейфа концепта» свойства признаков могут поменяться, что приведет к устареванию классифицирующей модели [8, 21, 22].

В связи с этим возникает необходимость разработать методы и алгоритмы, ориентированные на повышение качественных показателей моделей в условиях изменения свойств данных.

В статье предлагается решение, направленное на повышение показателей качества обработки выборки данных. Рассматривается задача адаптивного применения моделей на отдельных сегментах.

Описание предлагаемого метода

Одним из путей повышения качества классификации является использование моделей, которые основаны на уточненной локальной информации [23–25]. В большинстве задач обучающая выборка рассматривается как единое множество. Однако составляющие ее кортежи данных могут быть получены под воздействием различных факторов [26]. Например, появление отдельных управляющих команд вызывает рост количества служебных сообщений в сетевом трафике. Смена сезонов года, увеличение продолжительности дня отражаются на потребляемых мощностях в электросетях. В реальных системах возникают ситуации, когда проявляются воздействия, изменяющие их состояния. Часть таких факторов можно определить заранее, другая часть возникает случайно и не поддается прогнозированию. Однако в любом случае цена ошибки может быть очень высока. Предполагая наличие факторов, под влиянием которых происходит изменение значений целевых переменных, можно идентифицировать кортежи, полученные в момент воздействия.

В связи с этим в ряде случаев возникает определенная возможность осуществить сегментирование выборки с учетом информации о действующих факторах, оказывающих влияние на свойства данных:

$x \in X$ — значения выборки данных X;

$\{a_1, ..., a_n\} \in A$ — множество моделей, используемых методами машинного обучения для решения практических задач классификации или регрессии;

$\{v_1, ..., v_k\} \in V$ множество воздействующих факторов, которые изменяют диапазоны значений целевых переменных. Часть из них поддаются аналитике и связаны, например, с цикличностью процессов. Влияние других можно определить исходя из анализа изменения свойств данных.

Формализацию воздействующих факторов можно осуществить с помощью функции принадлежности.

$I(v) : X \to M$, $M = \{1, 2, ..., m\}$ — индикаторная функция, разбивающая выборку данных X на

множество сегментов X^1, ..., X^m, в которых под влиянием фактора $v \in V$ изменялись диапазоны значений целевых переменных.

Функция разделяет последовательность значений на отдельные сегменты

$$(x_1^1, \ldots, x_{n1}^1) \in X^1, (x_1^2, \ldots, x_{n2}^2) \in X^2, \ldots,$$
$$(x_1^m, \ldots, x_{nm}^m) \in X^m,$$

где $\{X^1, \ldots, X^m\} \in X$ — множество сегментов выборки данных X. Временная последовательность делится на m отдельных сегментов. Получается разбиение, где данные можно рассматривать как сегменты временных последовательностей. Каждый сегмент $X^i \in X$ обладает своими свойствами (частотой объектов наблюдений, плотностью вероятности распределения данных и т. д.). В ходе разбиения могут возникать сегменты со сходными свойствами. В целях экономии вычислительных ресурсов определяется функция $f(X^i)$, вычисляющая свойства последовательности X^i. В случае если значение лежит внутри диапазона $H_0 \le \frac{f(X^i)}{f(X^j)} \le H_1$, где H_0 и H_1 — пороговые значения, то возможно формирование подвыборок объединением сегментов $X^i = X^i \cup X^j$.

Каждая модель машинного обучения $a_j(x)$ в зависимости от базовых алгоритмов и свойств данных подвыборки имеет свои качественные показатели, которые можно вычислить в процессе обучения.

$Q(a_j(x), X^i)$ — функционал качества модели $a_j(x)$ для подвыборки X^i.

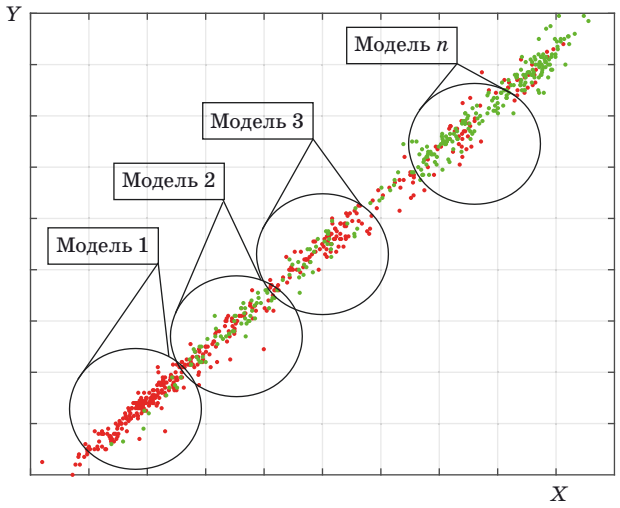
Тогда возникает необходимость выбора модели $a(x)$, обладающей лучшими качественными показателями на подвыборке данных:

$$a(x) = \underset{a_j(x) \in A, X^i \in X}{\arg\max} Q(a_j(x), X^i). \quad (1)$$

В статье рассматривается применение различных моделей машинного обучения на отдельных сегментах выборки данных.

Для каждой модели используются свои словари признаков, которые отличаются друг от друга. В статистических методах могут определяться значения дисперсии, выборочной доли, размаха данных и т. д. При обработке временных рядов, например методом скользящих средних, необходимо вычислять последовательность средних, ширину окна.

Предлагаемый метод иллюстрирует рис. 1. Область данных делится на отдельные сегменты. В зависимости от свойств данных на каждый сегмент назначается своя модель. Выбор модели и ее назначение определяются на основе значений функционала качества.



■ ***Рис. 1.*** Посегментное использование моделей
■ ***Fig. 1.*** Models processing in local segments

Таким образом, в основе предлагаемого метода лежит сегментация выборки, в результате которой необходимо выявить основные свойства входящих в нее данных и исходя из этого назначить наиболее подходящую из заранее предопределенных моделей.

Метод адаптивного применения моделей на отдельных сегментах выборки

Одним из проблемных вопросов адаптации моделей машинного обучения является отсутствие эффективных методов предобработки информации, направленных на вычисление и анализ свойств, позволяющих в режиме реального времени разделять поступающие последовательности на сегменты. Комплекс таких методов должен не только решать обычные задачи фильтрации, удаления шумов и выбросов, но и предоставлять информацию о свойствах данных для выбора и определения наиболее подходящих моделей. В целях решения обозначенных проблемных вопросов применяются постоянно обучающиеся модели.

Пример последовательности шагов постоянно обучающейся модели показан на рис. 2. Модель является двухуровневой. На первом уровне происходит обработка постоянного информационного потока, на втором действуют процедуры, обеспечивающие реализацию «механизма» обучения. Особенностью представленного решения является сегментирование обучающей выборки.

Для начального запуска процессов необходимо иметь предварительную информацию о значениях x_1, ..., x_n информационной последовательности. Они входят в первоначальное обучающее

■ ***Рис. 2.*** Пример последовательности шагов постоянно обучающейся модели

■ ***Fig. 2.*** A sequence steps example of constantly learning model

множество. На верхнем уровне модели выборка анализируется в целях определения отдельных сегментов, где свойства данных различаются. Возможно ее разделение как на основе заданной заранее системы правил, так и с помощью алгоритмов, выполняющих в автоматическом режиме поиск характерных точек, где изменяются свойства поступающих рядов.

В первом случае происходит изучение последовательности. На основе анализа совокупности данных выделяются тренды, периоды, сегменты, кластеры, обладающие отличающимися характеристиками. Эффективность такого подхода к разбиению определяется полнотой знаний о воздействующих факторах, под влиянием которых меняются диапазоны значений целевых переменных, частот событий, распределения вероятностей. В результате получается статичная система, настройка которой при изменении свойств может быть сложной.

Во втором случае разделение объектов наблюдения можно осуществить с помощью моделей, методов, алгоритмов, вычисляющих точки разладки, смену концепции. С помощью алгоритма автоматически определяются границы сегментов. Однако априорная информация о моделях смены концепции или разладки временного ряда может быть ограничена или отсутствовать. Недостатки моделей связаны с необходимостью повысить эффективность процедур детектирования изменения свойств временных последовательностей. Требуется постоянно отслеживать текущие настройки и базовые параметры при появлении новых данных.

Цель сегментирования состоит в том, чтобы обнаружить ситуации трансформации свойств последовательностей данных. Это осуществляется поиском момента θ, где происходит изменение характеристик наблюдаемого процесса:

$$x_t^i = \begin{cases} x_t^i, & 0 < t < \theta_i \\ x_t^{i+1}, & t \geq \theta_i \end{cases}.$$

В результате исходная выборка делится на несколько частей X^1, ..., X^m. Их свойства анализируются, и если имеется совпадение, где заранее определенные параметры одинаковы, то можно уменьшить количество рассматриваемых сегментов.

Подвыборки X^1, ..., X^m поступают на вход моделей a_1, a_2, ..., a_m. Происходит их обучение и анализ достигаемых качественных показателей. На каждом сегменте X^i для каждой модели $a_j(x)$ определяется функционал качества $Q(a_j(x), X^i)$. На основе его значений возможно ранжировать модели $\{a_1, ..., a_n\} \in A$ и осуществлять выбор имеющих наиболее высокие качественные показатели для каждого сегмента. В качестве условия выбора рассматривается выражение (1).

Процедуры сегментирования и определения свойств последовательности данных выполняются при обработке поступающего потока. Анализ свойств сегментов, выявленных при обработке информационного потока, и сопоставление их со свойствами подвыборок, полученных из обучающей выборки, позволяют назначить одну из заранее обученных моделей $\{a_1, ..., a_n\} \in A$ на текущий сегмент.

На последнем этапе выбранная модель $a_j(x)$ используется для решения задач обработки потока. Полученные результаты сравниваются с имеющимися, производится их анализ. Сопоставление

полученных моделью и реальных значений позволяет принять решение о формировании данных для уточнения алгоритма, которые впоследствии добавляются в обучающую выборку.

Таким образом, возможна реализация постоянно обучающейся модели, где процессы обучения и обработки информационных потоков могут осуществляться параллельно. В случае использования сложных моделей классификации или регрессии заранее предобученные модели позволяют уменьшить временные затраты на обучение при изменении свойств данных.

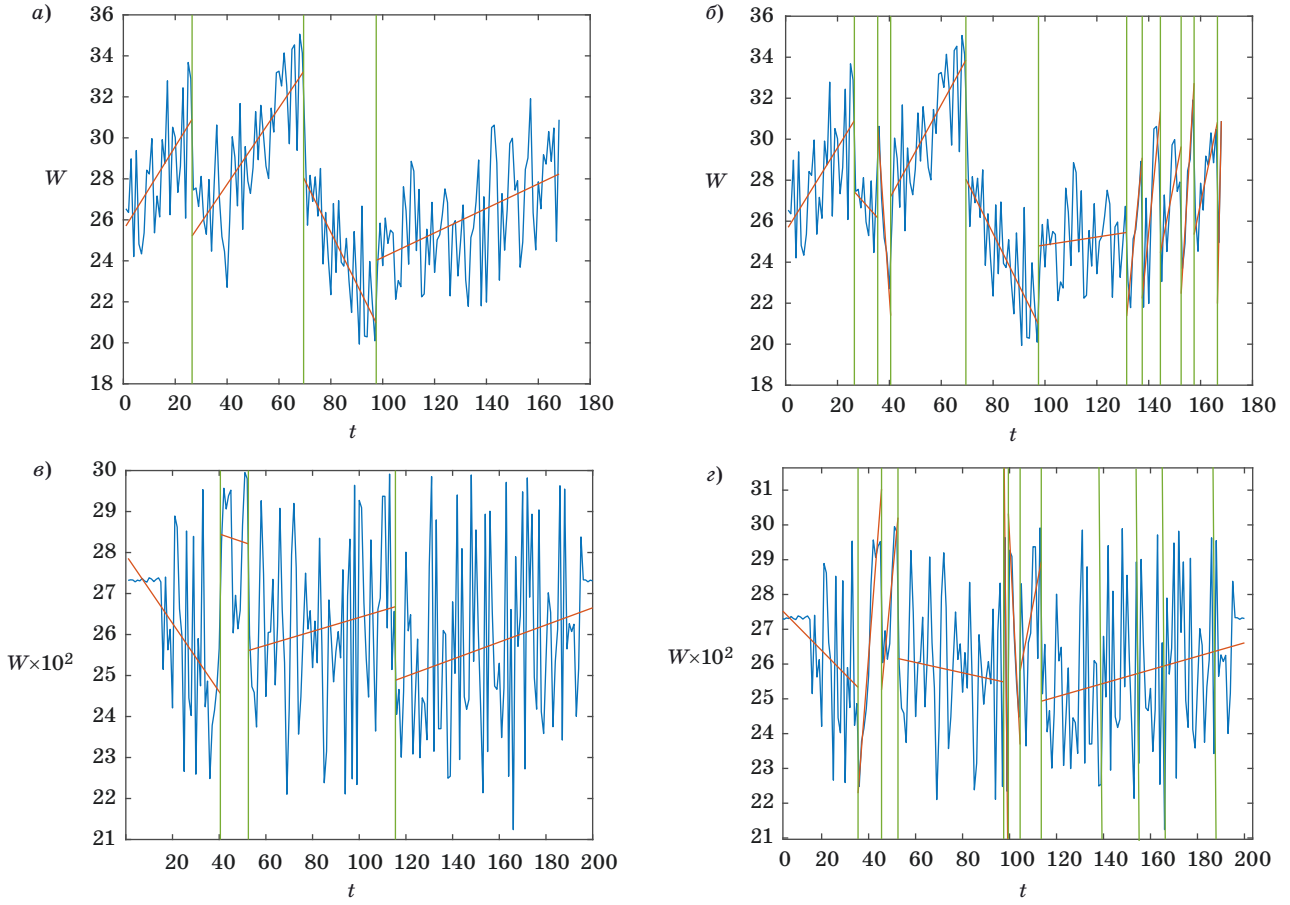
Эксперимент

Анализ применения предложенного подхода осуществлялся на данных, содержащих информацию о почасовой генерации электроэнергии солнечными и ветровыми электростанциями.

Целью эксперимента являлся анализ влияния размеров и способов получения сегментов данных на достигаемые качественные показатели в задачах регрессии по сравнению с целой выборкой. В первом случае данные датасета были разбиты на четыре части по кварталам и на двенадцать частей по месяцам согласно информации календаря. Во втором случае разбиение производилось с помощью алгоритма поиска точек смены направления тренда [27, 28]. Параметры алгоритма были подобраны таким образом, чтобы осуществить разбиение автоматическим способом также на четыре и двенадцать частей.

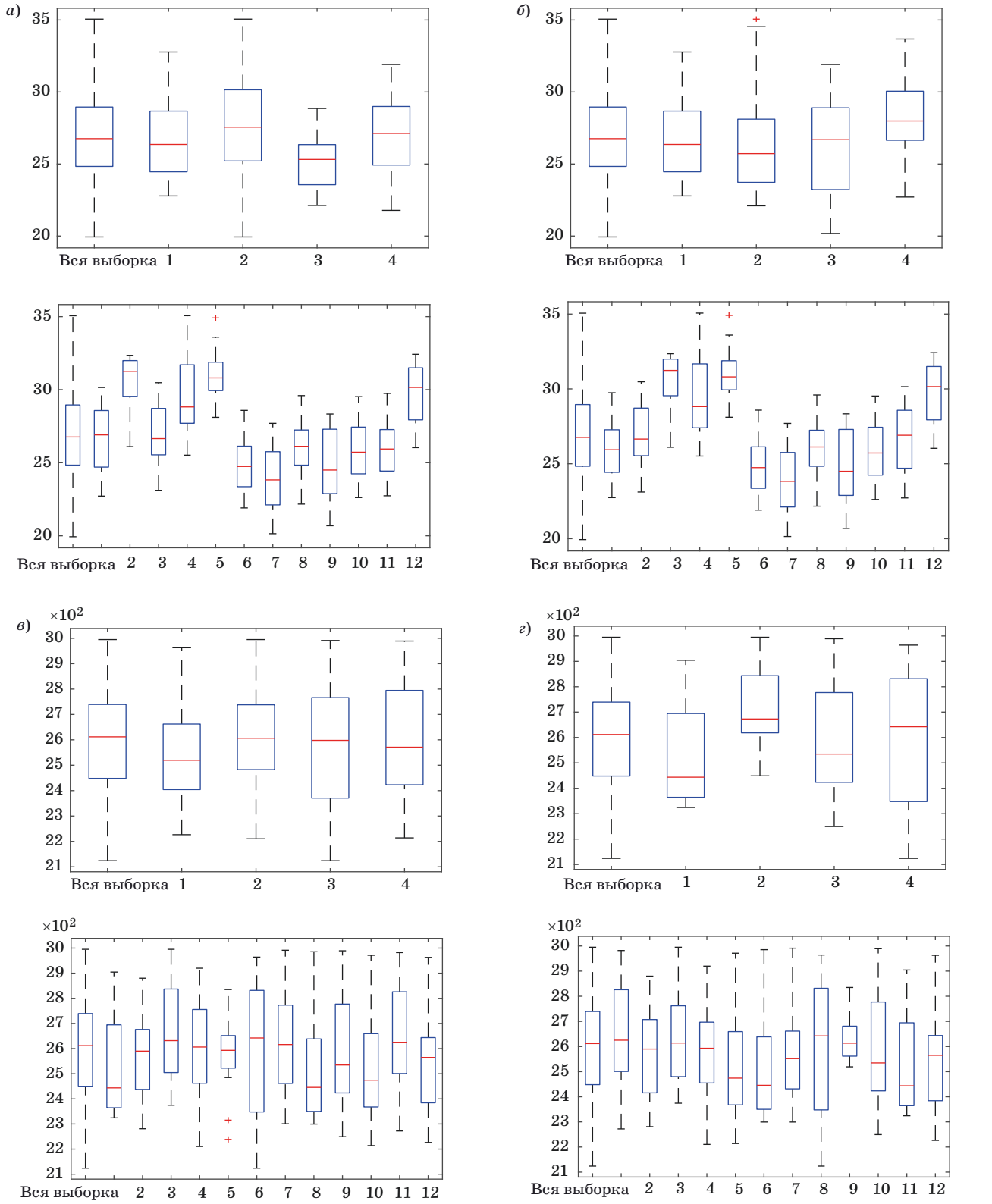
Получены сегменты временных последовательностей значений для солнечной и ветровой генерации энергии (рис. 3, *а–г*), в которых в целях определения свойств проводился статистический анализ данных.

Диаграммы результатов обработки значений целевых значений генерации электроэнергии солнечными батареями и ветряными установками (рис. 4, *а–г*) отражают медианное значение, первый и второй квартили, разброс. На диаграммах виден больший разброс значений для всей выборки в представленных множествах по сравнению с сегментами по месяцам или кварталам и



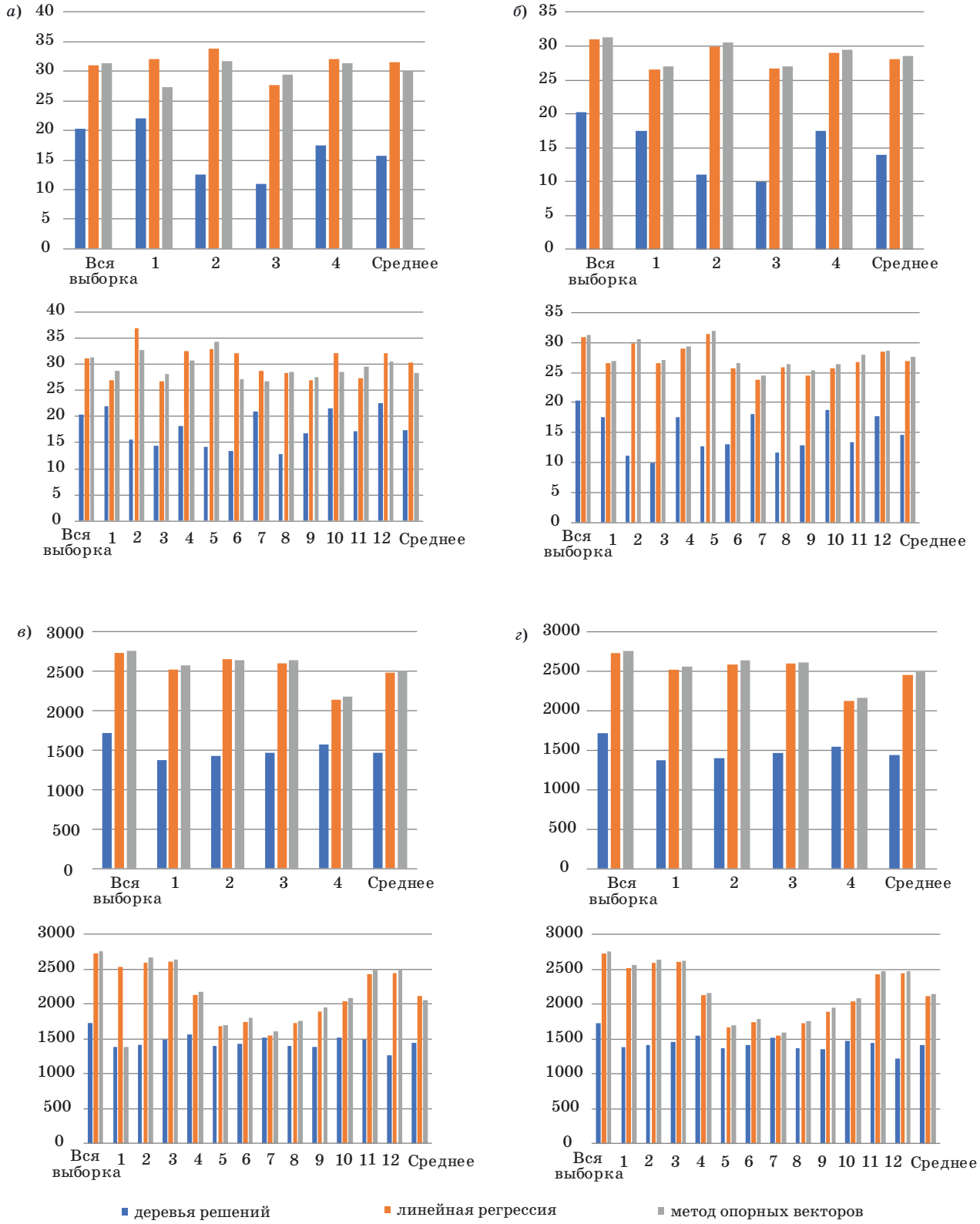
■ ***Рис. 3.*** Сегментирование датасета временной последовательности генерации солнечной (*а*, *б*) и ветровой (*в*, *г*) энергии

■ ***Fig. 3.*** Segmentation dataset of solar energy generation (*а*, *б*) and wind energy (*в*, *г*) time sequences values



■ ***Рис. 4.*** Свойства выборки данных при сегментации: по календарной информации для временной последовательности генерации солнечной (*а*) и ветровой (*в*) энергии; на основе алгоритма для временной последовательности генерации солнечной (*б*) и ветровой (*г*) энергии

■ ***Fig. 4.*** Data sampling properties: segmentation by calendar information for the time sequence of solar energy generation (*а*), wind energy (*в*); segmentation based on the algorithm for the time sequence of solar energy generation (*б*), wind energy (*г*)



■ ***Рис. 5.*** Функции потерь RMSE регрессионных моделей предсказания генерации электроэнергии при сегментации: по календарной информации для временной последовательности генерации солнечной (*а*) и ветровой (*в*) энергии; на основе алгоритма для временной последовательности генерации солнечной (*б*) и ветровой (*г*) энергии

■ ***Fig. 5.*** Loss functions RMSE of regression models in predicting electricity generatio: segmentation by calendar information for the time sequence of solar energy generation (*a*), wind energy (*в*); segmentation based on the algorithm for the time sequence of solar energy generation (*б*), wind energy (*г*)

сегментами, выявленными автоматическим способом. Применительно к рассматриваемому датасету диаграммы демонстрируют, что, несмотря на возможные «выбросы» данных, при сегментировании выборки диапазон между крайними значениями уменьшается по сравнению со всей выборкой в целом (что иллюстрирует визуальное сравнение сегментов с левым элементом «Вся выборка»). Применение сегментирования в ряде случаев уменьшает размах данных, частично борется с выбросами.

Для оценки влияния разделения на сегменты выборки на качество результатов машинного обучения были выбраны различные модели: линейной регрессии, деревьев решений и метод опорных векторов.

Данные представлялись одномерными временными рядами. На практике реальны более сложные модели. Рассматривалась возможность повышения качества за счет адаптивного выбора моделей. Выбор алгоритма определялся низкой вычислительной сложностью. На каждую модель подавались вся выборка полностью и данные из разделенных сегментов.

В качестве меры оценки алгоритма регрессии была выбрана функция потерь RMSE — классическая регрессионная метрика с одним выходом, которая вычисляет абсолютную разницу между прогнозируемыми и фактическими выходными данными:

$$L_{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}, \qquad (2)$$

где $y_i = a(x_i)$ — результат предсказания выбранного алгоритма; $\hat{y}_i$ — фактическое значение целевой переменной.

Функции потерь регрессионных моделей для различных сегментов данных согласно выражению (2) показаны на рис. 5, *а–г*. На гистограммах видно, что в результате разделения временной последовательности удается получить значения функции потерь меньше как в среднем, так и на большинстве отдельных сегментов.

При сравнении выбранных моделей лучшие результаты показывает алгоритм деревьев решений. При делении выборки на четыре части этот классификатор имеет значения функции потерь меньше, чем у других алгоритмов, во всех сегментах. Однако при делении на 12 сегментов на отдельных подвыборках его опередили метод опорных векторов и линейная регрессия. Анализируя свойства данных, можно назначать разные модели на отдельные сегменты.

Сегментация данных дает возможность уменьшить функцию потерь для разных областей выборки. Алгоритм поиска точек смены направления тренда позволяет выделить отдельные сегменты с меньшим размахом данных, что определяет более низкие значения функции потерь в среднем в регрессионной задаче.

Выделение сегментов последовательностей информационного потока данных и оценка их свойств позволяют осуществлять поиск и выбор моделей машинного обучения, обладающих лучшими характеристиками. Произведя оценку гистограмм рис. 5, сравнив значения функции потерь классификаторов на отдельных сегментах с левым столбцом «Вся выборка», видим, что на отдельных сегментах алгоритмы имеют лучшие результаты, чем при обработке всей выборки целиком. Результаты показывают, что применение предложенного метода, где каждому сегменту выборки данных назначается модель, имеющая на нем лучшие показатели качества, дает возможность уменьшить значения функции потерь RMSE от 8 до 18 % по сравнению с обработкой выборки целиком.

Предварительное обучение на выборках со сходными свойствами может сократить время на подготовку модели. Анализ результатов, полученных моделью, и реальных значений последовательности возможно использовать для формирования обучающих данных в целях уточнения модели. В дальнейшем осуществимо построение иерархий, когда модель верхнего уровня применяется для назначения наиболее эффективной модели нижнего уровня на отдельный сегмент.

Заключение

Одним из направлений, связанных с увеличением качественных показателей моделей классификации, является повышение качества данных, поступающих на вход алгоритмов. Для этих целей предложен метод, использующий адаптивное применение моделей машинного обучения на отдельных сегментах выборки. Сегментация, в определенных случаях, позволяет уменьшить разброс данных, выбросов и использовать изменение диапазонов значений переменных для повышения качества моделей.

Применение предложенного метода, основанного на разделении данных и выбора моделей с лучшими качественными показателями, помогает уменьшить значения функции потерь по сравнению с обработкой выборки целиком. Разделение последовательностей дает возможность бороться с выбросами и шумами и формировать компактно локализованные подмножества в пространстве объектов.

Свойства данных, на которых обучаются и тестируются регрессионные модели, влияют на их эффективность. Анализ информации об из-

менении диапазонов значений, балансов событий используется для формирования обучающих выборок в целях локального повышения качественных показателей моделей.

Новизна предлагаемого метода заключается в том, что с помощью правил или алгоритмов выборка разделяется на отдельные сегменты, каждый из которых обладает своими свойствами. Предварительное обучение на них алгоритмов способствует при изменении свойств потоков данных выбирать и назначать модели, обладающие лучшими качественными показателями.

На пути дальнейшего развития метода возможна его адаптация для задач прогнозирования и проактивного управления, направленного на оценку развития ситуации в динамике. Разбиение временных рядов на отдельные сегменты, анализ и сопоставление свойств последовательностей могут служить информацией для определения состояний. В результате можно выявить последовательности сегментов и определить переходы состояний. Обработка последовательностей, выделение сегментов предоставляют информацию, на основе которой можно строить графы и матрицы переходов. Анализ переходов состояний делает возможным в текущий дискретный момент времени определение наиболее вероятных переходов из текущего состояния в последующие.

Литература

1. **Di Franco G., Santurro M.** Machine learning, artificial neural networks and social research. *Qual Quant*, 2021, no. 5, pp. 1007–1025. https://doi.org/10.1007/s11135-020-01037-y
2. **Bonikowski B., Di Maggio P.** Varieties of American popular nationalism. *American Sociological Review*, 2016, no. 81(5), pp. 949–980.
3. **Sabar N. R., Ayob M., Kendall G., Qu R.** A dynamic multiarmed bandit-gene expression programming hyper-heuristic for combinatorial optimization problems. *IEEE Transactions on Cybernetics*, 2014, no. 45(2), pp. 217–228.
4. **Park J., Kim S.** Machine learning-based activity pattern classification using personal PM2.5 exposure information. *International Journal of Environmental Research and Public Health*, 2020, no. 17(18), pp. 65–73. https://doi.org/10.3390/ijerph17186573
5. **Maletzke A., dos Reis D., Batista G.** Combining instance selection and self-training to improve data stream quantification. *Journal of the Brazilian Computer Society*, 2018, vol. 24, no. 12, pp. 123–141. doi:10.1186/s13173-018-0076-0
6. **Jordan M. I., Mitchell T. M.** Machine learning: Trends, perspectives, and prospects. *Science*, 2015, no. 349(6245), pp. 255–260.
7. **Fanaee T. H., Gama J.** Event labeling combining ensemble detectors and background knowledge. *Progress in Artificial Intelligence*, 2014, no. 2, pp. 113–127. https://doi.org/10.1007/s13748-013-0040-3
8. **Bishop C. M., Nasser M. N.** Pattern recognition and machine learning. *Journal of Electronic Imaging*, 2007, no. 16, pp. 49–69.
9. **Sukhoparov M. E., Semenov V. V., Salakhutdinova K. I., Boitsova E. P., Lebedev I. S.** The state identification of industry 4.0 mechatronic elements based on behavioral patterns. *Internet of Things, Smart Spaces, and Next Generation Networks and Systems: Proc. of 20th Intern. Conf., NEW2AN 2020, and 13th Conf., ruSMART 2020*, Saint-Petersburg, Russia, August 26–28, 2020, Part I, Aug 2020, pp. 126–134. https://doi.org/10.1007/978-3-030-65726-0_12
10. **Oikarinen E., Tiittanen H., Henelius A.** Detecting virtual concept drift of regressors without ground truth values. *Data Mining and Knowledge Discovery*, 2021, vol. 35, iss. 3, pp. 821–859. doi:10.1007/s10618-021-00739-7
11. **Lu J., Liu A., Dong F., Gu F., Gama J., Zhang G.** Learning under concept drift: a review. *IEEE Transactions on Knowledge and Data Engineering*, 2019, no. 31(12), pp. 2346–2363.
12. **Wang L. Y., Park C., Yeon K., Choi H.** Tracking concept drift using a constrained penalized regression combiner. *Computational Statistics & Data Analysis*, 2017, no. 108, pp. 52–69.
13. **Lei P., Todorovic S.** Temporal deformable residual networks for action segmentation in videos. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2018, pp. 6742–6751.
14. **Khan S., Yairi T.** A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing*, 2018, no. 107, pp. 241–265. doi:10.1016/j.ymssp.2017.11.024
15. **Zhou Z.-H., Feng J.** Deep forest. *National Science Review*, 2019, vol. 6, no. 1, pp. 74–86. doi:10.1093/nsr/nwy108
16. **Salehi H., Burgueño R.** Emerging artificial intelligence methods in structural engineering. *Engineering Structures*, 2018, no. 171, pp. 170–189. doi:10.1016/j.engstruct.2018.05.084
17. **Zissis D., Lekkas D.** Addressing cloud computing security issues. *Future Generation Computer Systems*, 2012, vol. 28, no. 3, pp. 583–592.
18. **Liu J., Li Y., Song S., Xing J., Lan C., Zeng W.** Multi-modality multi-task recurrent neural network for online action detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, no. 29(9), pp. 2667–2682.
19. **Татарникова Т. М., Богданов П. Ю.** Обнаружение атак в сетях интернета вещей методами машинного обучения. *Информационно-управляющие систе-*

мы, 2021, № 6, с. 42–52. doi:10.31799/1684-8853-2021-6-42-52
20. **Зегжда Д. П., Калинин М. О., Крундышев В. М., Лаврова Д. С., Москвин Д. А., Павленко Е. Ю.** Применение алгоритмов биоинформатики для обнаружения мутирующих кибератак. *Информатика и автоматизация*, 2021, № 4(20), с. 820–844.
21. **Chao Y. W., Vijayanarasimhan S., Seybold B.** Rethinking the faster r-cnn architecture for temporal action localization. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2018, pp. 1130–1139.
22. **Nguyen P., Liu T., Prasad G.** Weakly supervised action localization by sparse temporal pooling network. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2018, pp. 6752–6761.
23. **Atsuya O., Genki Y.** Computational mechanics enhanced by deep learning. *Computer Methods in Applied Mechanics and Engineering*, 2017, vol. 327, pp. 327–351. https://doi.org/10.1016/j.cma.2017.08.040
24. **Takacs A., Toledano-Ayala M., Dominguez-Gonzalez A., Pastrana-Palma A., Velazquez D. T., Ramos J. M., Rivas-Araiza A. E.** Descriptor generation and optimization for a specific out-door environment. *IEEE Access*, 2020, vol. 8, pp. 2169–2176. doi:10.1109/ACCESS.2020.2975474
25. **Wong J. C., Lian H., Cheong S. A.** Detecting macroeconomic phases in the Dow Jones Industrial Average time series. *Physica A: Statistical Mechanics and its Applications*, 2009, no. 388 (21), pp. 4635–4645.
26. **Лебедев И. С.** Сегментирование множества данных с учетом информации воздействующих факторов. *Информационно-управляющие системы*, 2021, № 3, с. 29–38. doi:10.31799/1684-8853-2021-3-29-38
27. **Killick R., Paul F., Eckley I. A.** Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 2012, vol. 107, no. 500, pp. 1590–1598.
28. **Lavielle M.** Using penalized contrasts for the change-point problem. *Signal Processing*, 2005, vol. 85, pp. 1501–1510.

UDC 621.396
doi:10.31799/1684-8853-2022-3-20-30

Adaptive application of machine learning models on separate segments of a data sample in regression and classification problems

I. S. Lebedev[a], Dr. Sc., Tech., Professor, orcid.org/0000-0001-6753-2181, isl_box@mail.ru
[a]St. Petersburg Federal Research Center of the RAS, 39, 14th Line, 199178, Saint-Petersburg, Russian Federation

Introduction: Achievement of specified qualitative indicators in machine learning solutions depends not only on the efficiency of algorithms, but also on data properties. One of the lines for the development of classification and regression models is the specification of local properties of data. **Purpose:** To improve the qualitative predictors when solving classification and regression problems based on the adaptive selection of various machine learning models on separate local segments of data sample. **Results:** We propose a method that uses a combination of different models and machine learning algorithms on subsamples in regression and classification problems. The method is based on the calculation of qualitative predictors and the selection of the best models on the local segments of data sample. The finding of transformations of data and time series allows to create sample sets, with the data having different properties (for example, variance, sampling fraction, data range, etc.). We consider the data segmentation based on the change point detection algorithm in time series trends and on analytical information. On the example of the real dataset, we show the experimental values of the loss function for the proposed method with different classifiers on separate segments and on the whole sample. **Practical relevance:** The results can be used in classification and regression problems for the development of machine learning models and methods. The proposed method allows to improve classification and regression qualitative predictors by assigning models that have the best performance on separate segments.

Keywords — machine learning, data segmentation, time series, data transformations.

For citation: Lebedev I. S. Adaptive application of machine learning models on separate segments of a data sample in regression and classification problems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 20–30 (In Russian). doi:10.31799/1684-8853-2022-3-20-30

Reference

1. Di Franco G., Santurro M. Machine learning, artificial neural networks and social research. *Qual Quant*, 2021, no. 5, pp. 1007–1025. https://doi.org/10.1007/s11135-020-01037-y
2. Bonikowski B., Di Maggio P. Varieties of American popular nationalism. *American Sociological Review*, 2016, no. 81(5), pp. 949–980.
3. Sabar N. R., Ayob M., Kendall G., Qu R. A dynamic multi-armed bandit-gene expression programming hyper-heuristic for combinatorial optimization problems. *IEEE Transactions on Cybernetics*, 2014, no. 45(2), pp. 217–228.
4. Park J., Kim S. Machine learning-based activity pattern classification using personal PM2.5 exposure information. *International Journal of Environmental Research and Public Health*, 2020, no. 17(18), pp. 65–73. https://doi.org/10.3390/ijerph17186573
5. Maletzke A., dos Reis D., Batista G. Combining instance selection and self-training to improve data stream quantification. *Journal of the Brazilian Computer Society*, 2018, vol. 24, no. 12, pp. 123–141. doi:10.1186/s13173-018-0076-0
6. Jordan M. I., Mitchell T. M. Machine learning: Trends, perspectives, and prospects. *Science*, 2015, no. 349(6245), pp. 255–260.
7. Fanaee T. H., Gama J. Event labeling combining ensemble detectors and background knowledge. *Progress in Artificial Intelligence*, 2014, no. 2, pp. 113–127. https://doi.org/10.1007/s13748-013-0040-3

8. Bishop C. M., Nasser M. N. Pattern recognition and machine learning. *Journal of Electronic Imaging*, 2007, no. 16, pp. 49–69.
9. Sukhoparov M. E., Semenov V. V., Salakhutdinova K. I., Boitsova E. P., Lebedev I. S. The state identification of industry 4.0 mechatronic elements based on behavioral patterns. *Internet of Things, Smart Spaces, and Next Generation Networks and Systems: Proc. of 20th Intern. Conf., NEW2AN 2020, and 13th Conf., ruSMART 2020*, Saint-Petersburg, Russia, August 26–28, 2020, Part I, Aug 2020, pp. 126–134. https://doi.org/10.1007/978-3-030-65726-0_12
10. Oikarinen E., Tiittanen H., Henelius A. Detecting virtual concept drift of regressors without ground truth values. *Data Mining and Knowledge Discovery*, 2021, vol. 35, iss. 3, pp. 821–859. doi:10.1007/s10618-021-00739-7
11. Lu J., Liu A., Dong F., Gu F., Gama J., Zhang G. Learning under concept drift: a review. *IEEE Transactions on Knowledge and Data Engineering*, 2019, no. 31(12), pp. 2346–2363.
12. Wang L. Y., Park C., Yeon K., Choi H. Tracking concept drift using a constrained penalized regression combiner. *Computational Statistics & Data Analysis*, 2017, no. 108, pp. 52–69.
13. Lei P., Todorovic S. Temporal deformable residual networks for action segmentation in videos. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2018, pp. 6742–6751.
14. Khan S., Yairi T. A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing*, 2018, no. 107, pp. 241–265. doi:10.1016/j.ymssp.2017.11.024
15. Zhou Z.-H., Feng J. Deep forest. *National Science Review*, 2019, vol. 6, no. 1, pp. 74–86. doi:10.1093/nsr/nwy108
16. Salehi H., Burgueño R. Emerging artificial intelligence methods in structural engineering. *Engineering Structures*, 2018, no. 171, pp. 170–189. doi:10.1016/j.engstruct.2018.05.084
17. Zissis D., Lekkas D. Addressing cloud computing security issues. *Future Generation Computer Systems*, 2012, vol. 28, no. 3, pp. 583–592.
18. Liu J., Li Y., Song S., Xing J., Lan C., Zeng W. Multi-modality multi-task recurrent neural network for online action detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, no. 29(9), pp. 2667–2682.
19. Tatarnikova T. M., Bogdanov P. Yu. Intrusion detection in internet of things networks based on machine learning methods. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 6, pp. 42–52 (In Russian). doi:10.31799/1684-8853-2021-6-42-52
20. Zegzhda D., Kalinin M., Kundyshev V., Lavrova D., Moskvin D., Pavlenko E. Application of bioinformatics algorithms for polymorphic cyberattacks detection. *Informatics and Automation (SPIIRAS Proc.)*, 2021, no. 4(20), pp. 820–844 (In Russian).
21. Chao Y. W., Vijayanarasimhan S., Seybold B. Rethinking the faster r-cnn architecture for temporal action localization. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2018, pp. 1130–1139.
22. Nguyen P., Liu T., Prasad G. Weakly supervised action localization by sparse temporal pooling network. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2018, pp. 6752–6761.
23. Atsuya O., Genki Y. Computational mechanics enhanced by deep learning. *Computer Methods in Applied Mechanics and Engineering*, 2017, vol. 327, pp. 327–351. https://doi.org/10.1016/j.cma.2017.08.040
24. Takacs A., Toledano-Ayala M., Dominguez-Gonzalez A., Pastrana-Palma A., Velazquez D. T., Ramos J. M., Rivas-Araiza A. E. Descriptor generation and optimization for a specific out-door environment. *IEEE Access*, 2020, vol. 8, pp. 2169–2176. doi:10.1109/ACCESS.2020.2975474
25. Wong J. C., Lian H., Cheong S. A. Detecting macroeconomic phases in the Dow Jones Industrial Average time series. *Physica A: Statistical Mechanics and its Applications*, 2009, no. 388 (21), pp. 4635–4645.
26. Lebedev I. S. Dataset segmentation considering the information about impact factors. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 29–38 (In Russian). doi:10.31799/1684-8853-2021-3-29-38
27. Killick R., Paul F., Eckley I. A. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 2012, vol. 107, no. 500, pp. 1590–1598.
28. Lavielle M. Using penalized contrasts for the change-point problem. *Signal Processing*, 2005, vol. 85, pp. 1501–1510.

УВАЖАЕМЫЕ АВТОРЫ!

Научные базы данных, включая Scopus и Web of Science, обрабатывают данные автоматически. С одной стороны, это ускоряет процесс обработки данных, с другой — различия в транслитерации ФИО, неточные данные о месте работы, области научного знания и т. д. приводят к тому, что в базах оказывается несколько авторских страниц для одного и того же человека. В результате для всех по отдельности считаются индексы цитирования, что снижает рейтинг ученого.

Для идентификации авторов в сетях Thomson Reuters проводит регистрацию с присвоением уникального индекса (ID) для каждого из авторов научных публикаций.

Процедура получения ID бесплатна и очень проста, есть возможность провести регистрацию на 12 языках, включая русский (чтобы выбрать язык, кликните на зеленое поле вверху справа на стартовой странице): https://orcid.org

УДК 519.718
doi:10.31799/1684-8853-2022-3-31-44

Научные статьи
Articles

Оценивание эффективности процесса функционирования системы обеспечения информационной безопасности на основе теории стохастической индикации

А. М. Сухов[a], *канд. техн. наук, докторант, orcid.org/0000-0003-2233-811X, 19am87@mail.ru*
[a]*Краснодарское высшее военное училище им. генерала армии С. М. Штеменко, Красина ул., 4, Краснодар, 350065, РФ*

Введение: *постоянный рост деструктивных воздействий, направленных на критические информационные системы в условиях несовершенства методов и средств обнаружения и реагирования на компьютерные атаки, вызывает необходимость разработки научно-методического аппарата своевременного предупреждения систем обеспечения информационной безопасности о возможной реализации сценариев деструктивных воздействий. Одним из эффективных путей решения данной проблемы является использование методов теории стохастической индикации.* ***Цель:*** *разработка инструмента для оценивания эффективности процесса функционирования системы обеспечения информационной безопасности.* ***Результаты:*** *описаны детерминированная, случайная и неопределенная составляющие процесса функционирования системы обеспечения информационной безопасности. Построены константные и функциональные индикаторы, раскрыты их отличительные особенности. Построены стохастические супериндикаторы для решения задачи оценивания эффективности рассматриваемого процесса. На основе теории эффективности целенаправленных процессов и целеустремленных систем описаны особенности построения стохастических индикаторов различного ранга.* ***Практическая значимость:*** *благодаря разработанным стохастическим временным индикаторам оцениваются вероятностно-временные характеристики деструктивного воздействия с учетом интервалов и моментов времени процесса его реализации, что позволяет своевременно информировать систему о возможном выполнении сценария деструктивного воздействия на элементы критической информационной инфраструктуры.*

Ключевые слова *— система обеспечения информационной безопасности, атомарное событие информационной безопасности, деструктивное воздействие, эффективность, качество, теория стохастической индикации.*

Для цитирования: Сухов А. М. Оценивание эффективности процесса функционирования системы обеспечения информационной безопасности на основе теории стохастической индикации. *Информационно-управляющие системы*, 2022, № 3, с. 31–44. doi:10.31799/1684-8853-2022-3-31-44

For citation: Sukhov A. M. Evaluating the effectiveness of the information security system process based on the theory of stochastic indicators. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 31–44 (In Russian). doi:10.31799/1684-8853-2022-3-31-44

Введение

Проблемы оценивания эффективности информационных систем начинаются с неоднозначности исходных определений. В отечественной и зарубежной литературе [1–6] даны различные по своей сути определения эффективности. Так, Г. Б. Петуховым был введен показатель эффективности, рассматриваемый как вероятностная мера соответствия характеристик случайных эффектов целенаправленного процесса требуемым (директивно заданным) значениям этих характеристик [7].

Нередко эффективность функционирования систем оценивалась методами, тесно связанными с функционированием современного рынка [8–10]. Исследования операционной системы приведены в работах [11–14].

В отличие от зарубежных подходов, оценивание различных операционных свойств систем в рамках российской школы исследования реализуется количественно.

Под эффективностью будем понимать комплексное операционное свойство целенаправленного процесса применения системы, характеризующее его приспособленность к достижению цели проводимой операции [7, 15, 16]. К наиболее полной характеристике степени достижения цели операции, которую проводит система обеспечения информационной безопасности (СОИБ), относится показатель пригодности (на основе семейства временных индикаторов). Им может служить вероятность $P_{\text{д.ц}} = P\left(\tau_{\text{н}} < \hat{\tau} \le \tau_{\text{д}}\right) = F_{\hat{\tau}}\left(\tau_{\text{д}}\right) - F_{\hat{\tau}}\left(\tau_{\text{н}}\right)$, где $\tau_{\text{н}}$ — минимально необходимые технологические затраты операционного времени для выполнения задачи с требуемым качеством; $\tau_{\text{д}}$ — директивное операционное время выполнения задачи.

При исследовании эффективности операции (целенаправленного процесса функционирования) наиболее типичной является ситуация, когда основные характеристики системы и параметры условий ее применения подвержены воздействию случайных факторов. Высокая стоимость СОИБ, сложность и масштабность решаемых задач, а также большие потери из-за ошибок в процессе их производства и испытаний стимулируют исследования эффективности использования систем данного рода.

Понятие стохастического супериндикатора

Современный уровень зависимости общества от информационных технологий обусловливает появление новых типов деструктивных воздействий (ДВ) и построения надежной СОИБ единого информационного пространства (ЕИП) [17–19]. В работе под деструктивным воздействием понимается целенаправленное, скоординированное воздействие либо на информационный ресурс, либо на информационную систему или на средства получения, передачи, обработки, хранения и воспроизведения информации в ее составе с целью вызвать заданные структурные и (или) функциональные изменения. Большинство, а в ряде случаев все средства защиты информации (СЗИ), входящие в состав СОИБ ЕИП, используют в своем арсенале датчики обнаружения (ДО), распознавания (ДР) и предупреждения (ДП) [20, 21]. Для краткости излагаемого материала ограничимся приведенными типами датчиков, которые в зависимости от реализуемой ими функции направлены на обнаружение $f_{\text{о}}(r)$, распознавание $f_{\text{р}}(r)$ и предупреждение $f_{\text{п}}(r)$ ДВ (рис. 1) и построены на основе индикаторов обнаружения $I_{\text{о}}$, распознавания $I_{\text{р}}$ и предупреждения $I_{\text{п}}$. Допустим, что СОИБ проводит операцию распознавания ДВ типа «Отказ в обслуживании» в ЕИП, тогда ДВ необходимо представить в виде множества J с различными r-ми типами возможных ДВ, направленных на снижение требуемого уровня защищенности ЕИП, где $J=\left\{j_1^r,\ j_2^r,\ j_3^r,\ \ldots,\ j_m^r\right\},\ r=\overline{1(1)R},\ m=\overline{1(1)M}$.

Отсюда следует, что СОИБ распознает r-й тип из множества J ДВ, если индикатор $I_{\text{р}}$ распознавания, по значению которого датчик просигнализирует СОИБ, примет вид

$$I_J=I_{\text{р}}(r)\overset{\text{def}}{=}\begin{cases}1,\ r\in J;\\ 0,\ r\notin J.\end{cases} \tag{1}$$

Множеству J r-х типов ДВ соответствует его индикатор, и, наоборот, каждая функция, принимающая лишь одно из двух значений {0, 1}, может интерпретироваться как индикатор некоторого множества и может быть задана линейным выражением

$$I_J(r)\times f(r)=\begin{cases}f(r),\ r\in J;\\ 0,\ r\notin J.\end{cases} \tag{2}$$

Пусть теперь D — пересечение, а B — объединение двух подмножеств j_1^r и j_2^r множества J r-х типов ДВ, т. е. $D=j_1^r\cap j_2^r$, $B=j_1^r\cup j_2^r$. Очевидно, что тогда

$$I_D=\inf\left\{I_{j_1^r},\ I_{j_2^r}\right\}=\min\left\{I_{j_1^r},\ I_{j_2^r}\right\}; \tag{3}$$

$$I_B=\sup\left\{I_{j_1^r},\ I_{j_2^r}\right\}=\max\left\{I_{j_1^r},\ I_{j_2^r}\right\}, \tag{4}$$

т. е. значение индикатора I_D или I_B множества D или B равно соответственно наименьшему или наибольшему из значений индикаторов $I_{j_1^r}$ и $I_{j_2^r}$. Поэтому для обозначения наименьшего или наибольшего из значений двух функций $f(r)$ и $g(r)$ используем теоретико-множественные обозначения:

$$\inf\left\{f(r),\ g(r)\right\}\overset{\text{def}}{=}f\cap g(r); \tag{5}$$

$$\sup\left\{f(r),\ g(r)\right\}\overset{\text{def}}{=}f\cup g(r). \tag{6}$$

Любой воздействующий на ЕИП r-й тип ДВ протекает в течение определенного интервала $\tau_{\text{р}}$ времени его реализации. Для уточнения структуры r-го типа ДВ необходимо произвести декомпозицию j_m^r ДВ на атомарные события информационной безопасности (АСИБ) $C=\left\{c_1^r,\ c_2^r,\ c_3^r,\ \ldots,\ c_l^r\right\},\ l=\overline{1(1)L}$ и построить индикатор реализации j_m^r ДВ. Впоследствии уточним из C АСИБ реализованный сценарий

$$j_m^r=\left\{\text{АСИБ}_1,\ \text{АСИБ}_2,\ \text{АСИБ}_3,\ \ldots,\ \text{АСИБ}_l\right\}=\left\{c_1^r,\ c_2^r,\ c_3^r,\ \ldots,\ c_l^r\right\}$$

ДВ с учетом $\tau_{\text{р}}$, равного $\tau_{\text{р}}=\tau_1+\tau'_{(1,2)}+\tau_2+\tau'_{(2,3)}+\tau_3+\tau'_{(3,k+1)}+\ldots+\tau_{k+1}$, где $\tau_{\text{р}}$ осуществляется только на интервале $\left[t'_1,\ t''_{k_l}\right)$, что наглядно представлено на рис. 2.

Для описания индикатора $I_{j_2^r}$ процесса реализации сценария $j_2^r=\left(c_1^r,\ c_2^r\right)$ ДВ, представленного на рис. 3, будем использовать кусочно-единичные («селектирующие») функции:

– «селектор луча»

$$\Delta\left(j_2^r\right)\overset{\text{def}}{=}\begin{cases}1,\ j_m^r\ge 0\\ 0,\ j_m^r<0\end{cases}; \tag{7}$$

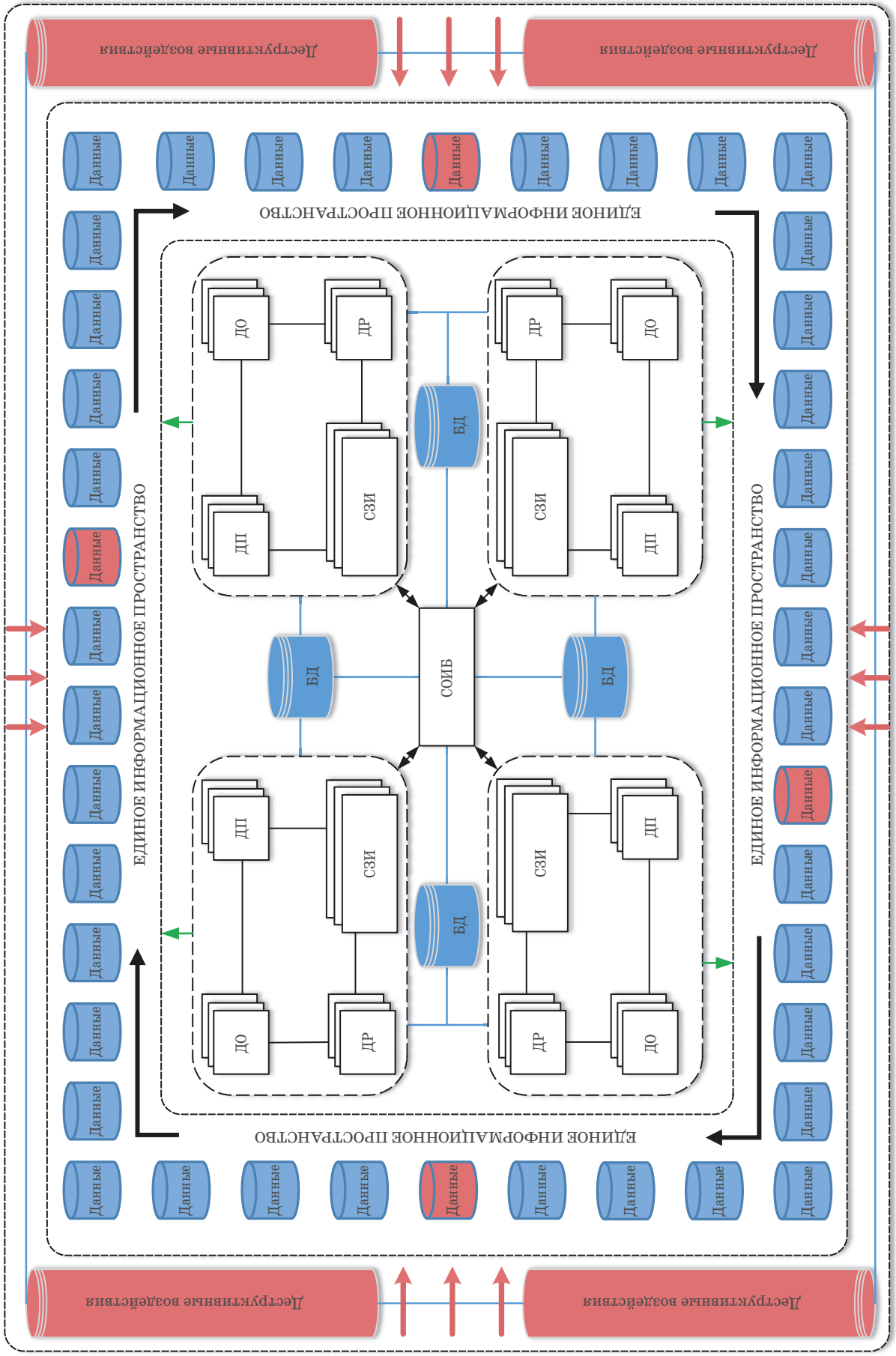
– «селектор интервала»

$$\prod\left(j_2^r;\ c_1^r,\ c_2^r\right)\overset{\text{def}}{=}\Delta\left(j_2^r-c_1^r\right)-\Delta\left(j_2^r-c_2^r\right)\triangleq\triangleq\Delta\left(j_2^r-c_1^r\right)\times\Delta\left(c_2^r-j_2^r\right); \tag{8}$$

– «селектор точки»

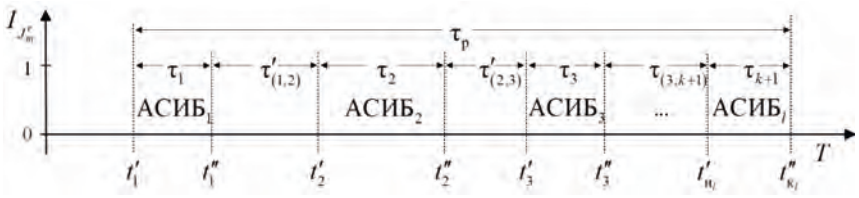
$$\varepsilon\left(j_2^r;\ a\right)\overset{\text{def}}{=}\Delta\left(j_2^r-a\right)\times\Delta\left(a-j_2^r\right). \tag{9}$$

Нетрудно видеть, что реализация C АСИБ, входящих в структуру сценария j_2^r ДВ, осущест-

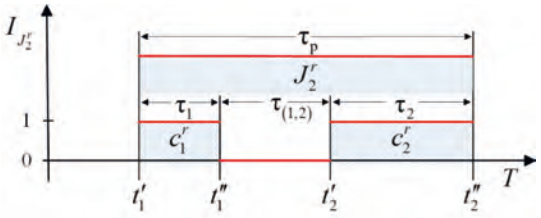


■ ■ *Рис. 1*. Структура процесса функционирования СОИБ в ЕИП

■ ■ *Fig. 1*. The structure of the process of functioning of the information security system in a single information space



■ ***Рис. 2.*** Индикатор реализации j_m^r ДВ
■ ***Fig. 2.*** Indicator of the implementation j_m^r of destructive impact



■ ***Рис. 3.*** Индикатор реализации $I_{j_2^r}$ ДВ
■ ***Fig. 3.*** Indicator of the implementation $I_{j_2^r}$ of destructive impact

вляется только на интервале $[t_1', t_2'']$, т. е. «селектор луча» («единичная функция Хевисайда») — суть индикатор полубесконечного интервала $[0, \infty)$:

$$\Delta\left(j_2^r\right) = I_J\left(j_2^r\right). \tag{10}$$

По аналогии с (10)

$$I'\left(j_2^r; c_1^r, c_2^r\right) = I_D\left(j_2^r\right), C = c_1^r \cap c_2^r,$$
$$c_1^r = [t_1', t_1''], c_2^r = [t_2', t_2''], \tag{11}$$

т. е. «селектор интервала» («единичный прямоугольный импульс») — индикатор интервала $[t_1', t_2'']$.

Таким образом, «селектор точки» («функция эквивалентности») — индикатор одноточечного множества $\{a\}$, который запишем в виде

$$\varepsilon\left(j_2^r; a\right) = I_{\{a\}}\left(j_2^r\right), \{a\} = a. \tag{12}$$

Если воспользоваться обозначением (5), то

$$\prod\left(j_2^r; c_1^r, c_2^r\right) = f \cap g\left(j_2^r\right), \tag{13}$$

где $f\left(j_2^r\right) = \Delta\left(j_2^r - c_1^r\right)$; $g\left(j_2^r\right) = \Delta\left(c_2^r - j_2^r\right)$.

И, соответственно:

$$\varepsilon\left(j_2^r; a\right) = f\left(j_2^r\right) \cap g\left(j_2^r\right), \tag{14}$$

где $f\left(j_2^r\right) = \Delta\left(j_2^r - c_1^r\right)$; $g\left(j_2^r\right) = \Delta\left(c_2^r - j_2^r\right)$.

Установлено, что кусочно-единичные («селектирующие») функции являются одним из немногих инструментов, позволяющих с высокой степенью детализации строить индикаторы $I_{j_2^r}$, показывающие множество $J = \left\{j_1^r, j_2^r, j_3^r, \ldots, j_m^r\right\}$ реализующихся сценариев ДВ r-х типов с учетом как самих интервалов времени $[t', t_k'']$ реализации j_m^r сценариев ДВ r-х типов, так и интервалов времени $[t_1', t_1'']$, $[t_2', t_2'']$, …, $\left[t_{k_l}', t_{k_{l+1}}''\right)$ реализации множества $C = \left\{c_1^r, c_2^r, c_3^r, \ldots, c_l^r\right\}$ АСИБ.

В современных условиях функционирования СОИБ наибольший для исследователей интерес представляют сценарии j_m^r ДВ r-х типов с вероятностной структурой. Предположим, что ДВ состоит из $\hat{C}$ АСИБ, $\hat{C} \subseteq X$ является случайным, так как заранее неизвестна структура ДВ (где X — универсальное множество, H — множество логических возможностей по обнаружению, распознаванию и предупреждению деструктивных возможностей злоумышленника, C — множество АСИБ). Введем допущение, что число АСИБ, составляющих сценарий j_m^r ДВ, является случайным. Временной интервал, на котором реализуются все АСИБ, входящие в состав сценария j_m^r ДВ, примем равным 1, достоверное событие (обнаружение реализации сценария j_m^r ДВ в ЕИП СОИБ) примем равным 1. Тогда индикатор обнаружения всего множества C АСИБ, входящих в состав сценария j_m^r ДВ, $I_{\hat{J}}$ будет представлять собой случайную величину $\hat{\omega}_J$ со следующими свойствами:

$$I_{\hat{J}} \overset{\text{def}}{=} \hat{\omega}_j = \begin{cases} 1, \text{ если } J \text{ произойдет;} \\ 0, \text{если } J \text{ не произойдет} \\ \left(\text{произойдет} \neg J\right). \end{cases} \tag{15}$$

В отличие от индикаторов $I_{\hat{J}}$ множеств сценариев j_m^r ДВ r-го типа, индикаторы $\hat{\omega}_J$ обнаружения на неопределенном интервале $\left(\hat{t}_{k_l}', \hat{t}_{k_{l+1}}''\right)$ времени их реализации носят случайный характер и называются стохастическими. В работе [7] авторы оперируют не случайными событиями, а предлагают переходить к случайным величинам, для которых в теории вероятностей разработан более гибкий и универсальный математический

аппарат. Таким образом, плотность распределения $\varphi_{\hat{\omega}_J}(\omega)$ и функция распределения $F_{\hat{\omega}_J}(\omega)$ стохастического индикатора $\hat{\omega}_J$ обнаружения сценариев j_m^r ДВ будут описываться следующими выражениями:

$$\varphi_{\hat{\omega}_J}(\omega) = q\delta(\omega) + p\delta(\omega - 1); \quad (16)$$

$$F_{\hat{\omega}_J}(\omega) = q\Delta(\omega) + p\Delta(\omega - 1), \quad (17)$$

где $p = P(\hat{J})$; $q = 1 - p = P(\neg \hat{J})$.

В свете вышеизложенного и с учетом (16), (17) из соотношения (15) следует

$$P(\hat{J}) = p = P(\hat{\omega}_J = 1) = M[\hat{\omega}_J] = \overline{\omega}_J. \quad (18)$$

Раскроем равенство (18), для чего воспользуемся содержательной трактовкой понятия случайного события $\hat{J}$. Допустим, что СОИБ выполняет функцию $f_o(r)$ обнаружения возможно реализующихся сценариев $\hat{J}$ ДВ, тогда под целенаправленным процессом ее функционирования стоит рассматривать операцию обнаружения признака C АСИБ, по которому СОИБ будет реализовывать функцию $f_p(r)$ распознавания сценария $\hat{j}_m^r$ ДВ r-го типа, тогда под $\hat{J}$ понимается исход операции обнаружения сценария $\hat{j}_m^r$ ДВ r-го типа, состоящей в реализации условий B', B'', где B' — условия функционирования СОИБ, а B'' — условия применения СОИБ, при которых СОИБ хочет обнаружить сценарии j_m^r ДВ r-го типа. Тогда случайное событие $\hat{J}$ есть не что иное, как исход операции обнаружения сценария j_m^r ДВ r-го типа, проходящей при воздействии на B', B'' не поддающихся учету случайных факторов, т. е. в условиях $\hat{B}'$, $\hat{B}''$ некоторой неопределенности, приводящей к тому, что обнаружение сценария j_m^r ДВ r-го типа происходит не при каждой реализации условий B', B''. Следовательно, связь между обнаружением признака C АСИБ, принадлежащего одному из сценариев j_m^r ДВ r-го типа, и предопределяющими $\hat{B}'$, $\hat{B}''$ условиями носит случайный характер. Количественное оценивание проводимой СОИБ в ЕИП операции обнаружения сценария j_m^r ДВ r-го типа будем выполнять при помощи вероятности $P(\hat{J})$ обнаружения случайного $\hat{C}$ АСИБ, которая характеризует степень объективной возможности обнаружения сценария $\hat{j}_m^r$ ДВ r-го типа в условиях $\hat{B}'$, $\hat{B}''$.

Проведем аналогию высказывания относительно обнаружения сценария j_m^r ДВ r-го типа и условий B' функционирования СОИБ и B'' применения, в которой оно истинно. Действительно, при решении конкретных прикладных задач информационной безопасности описание любого из исследуемых событий дается в форме некоторого высказывания (предположения, гипотезы). Истинность высказывания обнаружения сценария j_m^r ДВ r-го типа адекватна достоверности его обнаружения СОИБ в процессе выполнения функции $f_o(r)$. Аналогом множества $J = \{j_1^r, j_2^r, j_3^r, \ldots, j_m^r\}$ ДВ r-го типа, нарушающих установленную администратором информационной безопасности политику, является множество логических возможностей $H = \{h_1^r, h_2^r, h_3^r, \ldots, h_w^r\}$ (при которых высказывание J истинно), называемое множеством истинности высказывания J. Основываясь на описанном выше, получаем возможность вести содержательное описание функции $f_o(r)$ обнаружения вероятно реализующихся сценариев $\hat{J}$ ДВ СОИБ в терминах алгебры высказываний, что применительно к возложенным функциям по обнаружению $f_o(r)$, распознаванию $f_p(r)$ и предупреждению $f_п(r)$ ДВ в ЕИП на СОИБ обладает большей наглядностью, чем описание на языке алгебры событий, т. е.

$$I_J(r) = \begin{cases} 1, & r \in J,\ J \subseteq X; \\ 0, & r \notin J,\ J \subseteq X. \end{cases} \quad (19)$$

Необходимо отметить, что истинность и ложность высказывания по обнаружению сценария j_m^r ДВ r-го типа эквивалентны соответственно достоверности и невозможности обнаружения сценария j_m^r ДВ r-го типа, а стохастичность ситуации $\hat{B}'$, $\hat{B}''$, в которой высказывание истинно, эквивалентна неопределенности условий $\hat{B}'$, $\hat{B}''$, определяющих случайный эксперимент (единичную операцию обнаружения сценария j_m^r ДВ r-го типа при неопределенных условиях B' функционирования СОИБ и B'' применения соответственно). При такой трактовке очевидно, что каждая из алгебр (высказываний и событий) является булевой и изоморфна алгебре их индикаторов $I_J(r)$. Следовательно, эти алгебры изоморфны между собой и неопределенная ситуация на языке любой из них адекватна и применима для построения индикаторов обнаружения I_o, распознавания I_p и предупреждения $I_п$, на основе которых построены ДО, ДР и ДП (см. рис. 1).

Количественные характеристики неопределенности стохастической ситуации при исследовании процессов функционирования СОИБ

В целях определения количественных характеристик необходимо описать детерминированную, случайную и неопределенную составляющие процесса функционирования СОИБ.

Детерминированными подразумеваем процессы, вызванные действием полностью известных

условий B', B''. Такие процессы практически не встречаются в современных условиях (только лишь на этапах проектирования СОИБ, в моделях, не учитывающих J_m ДВ злоумышленника).

Случайные процессы (ошибки операторов, запуск вредоносного программного обеспечения, превышенное количество обращений легитимных пользователей и др.) возникают при воздействии не поддающихся учету $\hat{B}''$ условий применения, на известные причины B' условий функционирования СОИБ, делающих причины $\hat{B} = B' \cup \hat{B}''$, и основные свойства процесса случайными.

В связи с недостаточностью (не в полной мере учтены возможности возмущающей среды) или отсутствием наблюдений (число опытов таких систем на практике невелико, так как реализуемые злоумышленниками J ДВ в современных условиях носят уникальный, единичный характер), необходимых для определения вероятностных свойств исследуемого процесса обнаружения сценариев j_m^r ДВ r-го типа, неопределенным становится и сам процесс обнаружения сценариев $\hat{j}_m^r$ ДВ r-го типа СОИБ в условиях $\hat{B}'$, $\hat{B}''$.

Не вдаваясь в сравнительный анализ понятий «случайность» и «неопределенность», а также связанных с ними понятий «объективная» и «субъективная» вероятности, отметим, что в принципе различия между ними чисто условны. С одной стороны, при наблюдении реальных процессов функционирования СОИБ их случайность и неопределенность проявляются одинаково — как невозможность точного прогнозирования момента $t_{\text{обн}}$ обнаружения C АСИБ, входящих в состав сценария j_m^r ДВ r-го типа, на интервале $(t_{\text{н}}, t_{\text{к}})$ времени; с другой стороны, «объективные» вероятностные характеристики случайных процессов не могут быть полностью свободными от «субъективных» взглядов их исследователей. Даже задаваемые экспертами «субъективные» вероятности в значительной мере являются «объективными», поскольку основаны на опыте изучения экспертами случаев применения СОИБ.

Таким образом, степень случайности вышеупомянутого высказывания охарактеризуем его вероятностью $P\left(\hat{J}\right) = p$. В случае обнаружения сценария j_m^r ДВ r-го типа СОИБ высказывание J окажется истинным, и его вероятность станет равной 1, а если СОИБ не справится с обнаружением сценария j_m^r ДВ r-го типа, высказывание J окажется ложным, и его вероятность, соответственно, будет равна 0. Поскольку априори неизвестно, истинным или ложным окажется высказывание $\hat{J}$, следовательно, неизвестно, какое из значений (1 или 0) примет его вероятность, которая, таким образом, является случайной величиной, подчиненной закону распределения Бернулли с параметром p. Именно так и распределен стохастический индикатор $\hat{\omega}_J$ обнаружения сценария j_m^r ДВ r-го типа на неопределенном интервале $\left(\hat{t}'_{k_l}, \hat{t}''_{k_{l+1}}\right)$ времени его реализации [см. (16), (17)], и, следовательно, он приобретает смысл условной вероятности высказывания $\hat{J}$ относительно ситуации B. В этом случае неопределенность ситуации $\hat{B}$ характеризуется набором {0, 1} возможных значений стохастического индикатора $\hat{\omega}_J$, а случайность высказывания $\hat{J}$ характеризуется его вероятностью p.

Формализация стохастической ситуации

Количественные характеристики при исследовании процессов функционирования СОИБ относительно ситуации B в значительной мере носят качественный характер, затрудняющий применение математических методов исследования. Для получения возможности количественного анализа таких ситуаций необходима их формализация, т. е. построение адекватных им математических моделей.

Предположим, что СОИБ производит наблюдение за реализацией сценария j_m^r ДВ r-го типа, который состоит из C АСИБ, где $C = \left\{c_1^r, c_2^r\right\}$. Интервал времени реализации сценария j_m^r ДВ r-го типа обозначим $\tau_{\text{р.с}}$. С учетом $C = \left\{c_1^r, c_2^r\right\}$, $\tau_{\text{р.с}} = \tau_1 + \tau_{\text{о}} + \tau_2$, где $\tau_1 = \left\langle t'_1, t''_1 \right\rangle$, а $\tau_2 = \left\langle t'_2, t''_2 \right\rangle$. Временной интервал между реализацией АСИБ обозначим $\tau_{\text{о}}$: $\tau_{\text{о}} = \left\langle t''_1, t'_2 \right\rangle$. Иллюстрирует сказанное рис. 4.

С учетом вышеизложенного пусть:

U — действительная прямая;

$U^2 \overset{\text{def}}{=} U \times U$ — действительная плоскость;

c_1^r, c_2^r — константы;

τ_1, τ_2 — переменные;

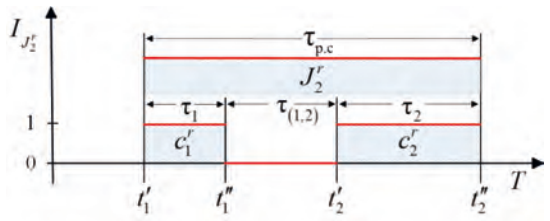
$<$, $>$, $\le$, $\ge$ — отношения порядка.

Тогда:

$c_1^r < c_2^r$, $c_1^r > c_2^r$, $c_1^r \le c_2^r$, $c_1^r \ge c_2^r$ — высказывания;

$\tau_2 < c_1^r$, $\tau_2 > c_1^r$, $\tau_2 \le c_2^r$, $\tau_2 \ge c_2^r$ — одноместные предикаты;

$\tau_2 < \tau_1$, $\tau_2 > \tau_1$, $\tau_2 \le \tau_1$, $\tau_2 \ge \tau_1$ — двухместные предикаты.



■ ***Рис. 4.*** Индикатор I_J сценария j_m^r ДВ r-го типа

■ ***Fig. 4.*** The indicator I_J of the j_m^r scenario is of destructive impact of the r-th type

Константы c_1^r, c_2^r фиксируют ситуацию реализации C АСИБ сценария j_m^r ДВ r-го типа на интервале $\langle t_1', t_2'' \rangle$ (см. рис. 4), в которой высказывание $c_1^r < c_2^r$ либо истинно (если $c_1^r < c_2^r$), либо ложно (если $c_1^r \geq c_2^r$). Практически всегда переменная $\hat{\tau}_2$ является случайной величиной, тогда для определения вероятности высказывания $\left(\hat{\tau}_2 < c_1^r\right)$ достаточно знать закон распределения $F_{\hat{\tau}_2}(\tau_2)$ случайной величины $\hat{\tau}_2$:

$$p = P\left(\hat{J}\right) = P\left(\hat{\tau}_2 < c_1^r\right) = \int_{-\infty}^{a} \mathrm{d}F_{\hat{\tau}_2}(\tau_2) = F_{\hat{\tau}_2}\left(c_1^r\right). \quad (20)$$

Когда $I_J(\hat{\tau}_2) = \hat{\omega}_J$ является стохастическим индикатором множества $J = \left(-\infty, c_1^r\right)$, то из выражений (18), (20) следует, что $P(\hat{\omega}_J = 1) = F_{\hat{\tau}_2}\left(c_1^r\right)$, и выражения (16), (17) примут вид

$$\varphi_{\hat{\omega}_J}(\omega) = R_{\hat{\tau}_2}\left(c_1^r\right) \times \delta(\omega) + F_{\hat{\tau}_2}\left(c_1^r\right) \times \delta(\omega - 1); \quad (21)$$

$$F_{\hat{\omega}_J}(\omega) = R_{\hat{\tau}_2}\left(c_1^r\right) \times \Delta(\omega) + F_{\hat{\tau}_2}\left(c_1^r\right) \times \Delta(\omega - 1). \quad (22)$$

Найдем числовые характеристики индикатора $\hat{\omega}_J$, который в дальнейшем будем называть константным. Начальный момент 1-го порядка распределения индикатора $\hat{\omega}_J$ определяется соотношением

$$v_k\left[\hat{\omega}_J\right] \overset{\text{def}}{=} M\left[\hat{\omega}_J^k\right] = \overline{\omega_J^k} = \int_{-\infty}^{\infty} \omega^k \mathrm{d}F_{\hat{\omega}_J}(\omega) = F_{\hat{\tau}_2}\left(c_1^r\right), \left[k = 1(1)K\right], \quad (23)$$

и, следовательно:

$$M_{\hat{\omega}_J} = \overline{\omega_J^1} = F_{\hat{\tau}_2}\left(c_1^r\right); \quad (24)$$

$$D_{\hat{\omega}_J} = \overline{\breve{\omega}_J^2} = \overline{\omega_J^2} - \overline{\omega}_J^2 = F_{\hat{\tau}_2}\left(c_1^r\right) - F_{\hat{\tau}_2}^2\left(c_1^r\right) = F_{\hat{\tau}_2}\left(c_1^r\right) \times R_{\hat{\tau}_2}\left(c_1^r\right). \quad (25)$$

Таким образом, как видно из равенства (24), вероятность случайного события $\hat{J}$ равна математическому ожиданию его индикатора $\hat{\omega}_J$.

Поскольку возможными значениями стохастического индикатора служат возможные степени достоверности случайного события $\hat{J}$, т. е. степени истинности неопределенного высказывания $\hat{\tau}_2 < c_1^r$, являющиеся значениями его апостериорной вероятности, то в рассматриваемом случае дисперсия $D_{\hat{\omega}_J}$ характеризует степень неопределенности предиката $\left(\hat{\tau}_2 < c_1^r\right) \simeq (\hat{\omega}_J = 1)$. При этом, как нетрудно понять, максимальная неопределенность будет иметь место при $c_1^r = \mathrm{Me}_{\hat{\tau}_2}$.

Рассмотрим неопределенный предикат $\hat{\tau}_2 < \tau_1$ и $\hat{\tau}_2 < \hat{\tau}_1$.

1. Переменная $\hat{\tau}_2$ случайна, тогда

$$p = p(\tau_1) = P\left(\hat{J}_{\tau_1}\right) = P(\hat{\tau}_2 < \tau_1) = \int_{-\infty}^{\tau_1} \mathrm{d}F_{\hat{\tau}_2}(\tau_2) = F_{\hat{\tau}_2}(\tau_1) = P\left[\hat{\omega}_J(\tau_1) = 1\right], \quad (26)$$

где $\hat{\omega}_J(\tau_1)$ — стохастический индикатор множества $J_{\tau_1} = (-\infty, \tau_1)$.

Из выражения (26) видно, что в данном случае $\hat{\omega}_{J_{\tau_1}} = \hat{\omega}_J(\tau_1) = \Delta(\tau_1 - \hat{\tau}_2)$ (рис. 5), т. е. формально индикатор случайного события $\hat{J}_{\tau_1}$ представляет собой случайную функцию, законы распределения которой имеют следующие выражения:

$$\varphi_{\hat{\omega}_{J_{\tau_1}}}(\omega;\, \tau_1) = R_{\hat{\tau}_2}(\tau_1) \times \delta(\omega) + F_{\hat{\tau}_2}(\tau_1) \times \delta(\omega - 1); \quad (27)$$

$$F_{\hat{\omega}_{J_{\tau_1}}}(\omega;\, \tau_1) = R_{\hat{\tau}_2}(\tau_1) \times \Delta(\omega) + F_{\hat{\tau}_2}(\tau_1) \times \Delta(\omega - 1). \quad (28)$$

Стохастический индикатор $\hat{\omega}_A(\tau_1)$ называется функциональным [7] с характеристиками

$$M\left[\hat{\omega}_J^k(\tau_1)\right] = \overline{\omega_J^k}(\tau_1) = \int_{-\infty}^{\infty} \omega^k \mathrm{d}F_{\hat{\omega}_{J_{\tau_1}}}(\omega;\, \tau_1) = F_{\hat{\tau}_2}(\tau_1); \quad (29)$$

$$M\left[\hat{\omega}_J(\tau_1)\right] = \overline{\omega}_J(\tau_1) = F_{\hat{\tau}_2}(\tau_1); \quad (30)$$

$$D\left[\hat{\omega}_J(\tau_1)\right] = \overline{\breve{\omega}_J^2} = \overline{\omega_J^2}(\tau_1) - \overline{\omega}_J^2(\tau_1) = F_{\hat{\tau}_2}(\tau_1) R_{\hat{\tau}_2}(\tau_1). \quad (31)$$

2. Переменная $\hat{\tau}_1$ также случайна, тогда предикат $\hat{\tau}_2 < \hat{\tau}_1$ становится дважды неопределенным, и вероятность случайного события $\hat{J}_{\hat{\tau}_1}$, при известной плотности распределения $\varphi_{\hat{\tau}_1, \hat{\tau}_2}(\tau_1, \tau_2)$, может быть представлена в виде

$$P\left(\hat{J}_{\hat{\tau}_1}\right) = P(\hat{\tau}_2 < \hat{\tau}_1) = P\left[\hat{\tau}_1, \hat{\tau}_2 \in (H)\right] = \iint_{(H)} \varphi_{\hat{\tau}_1, \hat{\tau}_2}(\tau_1, \tau_2) \mathrm{d}\tau_1 \, \mathrm{d}\tau_2, \quad (32)$$

где $(H) \cong \{\langle \tau_1, \tau_2 \rangle : \tau_2 < \tau_1\}$;

■ ***Рис. 5.*** Индикатор случайного события
■ ***Fig. 5.*** Indicator of a random event

$$P\left(\hat{J}_{\hat{\tau}_1}\right)=P\left(\hat{\tau}_2<\hat{\tau}_1\right)=P\left(\hat{u}<0\right)=\int\limits_{-\infty}^{0}\varphi_{\hat{u}}(u)\mathrm{d}u, \quad (33)$$

где $\hat{u}=\hat{\tau}_2-\hat{\tau}_1$;

$$P\left(\hat{J}_{\hat{\tau}_1}\right)=P\left(\hat{\tau}_2<\hat{\tau}_1\right)=P\left(\hat{\vartheta}\geq 0\right)=\int\limits_{0}^{\infty}\varphi_{\hat{\vartheta}}(\vartheta)\mathrm{d}\vartheta, \quad (34)$$

где $\hat{\vartheta}=\hat{\tau}_1-\hat{\tau}_2$.

При вышеприведенном определении вероятности случайного события $\hat{J}_{\hat{\tau}_1}$ (32), (33) теряется некоторый объем информации о реализации C АСИБ сценария J_m^r ДВ r-го типа на интервале $\left\langle t_1', t_2''\right\rangle$ (см. рис. 4). Для исследования реализации C АСИБ сценария J_m^r ДВ r-го типа на интервале $\left\langle t_1', t_2''\right\rangle$ с учетом сохранения информации преобразуем выражение (32):

$$\iint\limits_{(H)}\varphi_{\langle\hat{\tau}_1,\hat{\tau}_2\rangle}(\tau_1,\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2=$$

$$=\iint\limits_{\tau_2<\tau_1}\varphi_{\langle\hat{\tau}_1,\hat{\tau}_2\rangle}(\tau_1,\tau_2)\mathrm{d}\tau_1\,\mathrm{d}\tau_2=$$

$$=\int\limits_{-\infty}^{\infty}\varphi_{\hat{\tau}_2}(\tau_2)\left[\int\limits_{\tau_1}^{\infty}\varphi_{\hat{\tau}_2/\hat{\tau}_1}(\tau_1;\tau_2)\mathrm{d}\tau_1\right]\mathrm{d}\tau_2=$$

$$=\int\limits_{-\infty}^{\infty}F_{\hat{\tau}_2/\hat{\tau}_1}(\tau_1;\tau_1)\mathrm{d}F_{\hat{\tau}_1}(\tau_1); \quad (35)$$

$$\iint\limits_{(H)}\varphi_{\langle\hat{\tau}_1,\hat{\tau}_2\rangle}(\tau_1,\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2=$$

$$=\iint\limits_{\tau_2<\tau_1}\varphi_{\langle\hat{\tau}_1,\hat{\tau}_2\rangle}(\tau_1,\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2=$$

$$=\int\limits_{-\infty}^{\infty}\varphi_{\hat{\tau}_2}(\tau_2)\left[\int\limits_{\tau_2}^{\infty}\varphi_{\hat{\tau}_2/\hat{\tau}_1}(\tau_1;\tau_2)\mathrm{d}\tau_1\right]\mathrm{d}\tau_2=$$

$$=\int\limits_{-\infty}^{\infty}R_{\hat{\tau}_2/\hat{\tau}_1}(\tau_2;\tau_2)\mathrm{d}F_{\hat{\tau}_2}(\tau_2). \quad (36)$$

В соотношениях (35) и (36) для определения искомой вероятности $p=P\left(\hat{\tau}_2<\hat{\tau}_1\right)$ описаны два пути, которые приводят к одному и тому же результату, но обеспечивают различную его надежность [7].

Наиболее распространенным с практической точки зрения является ситуация, когда случайные величины $\hat{\tau}_1$ и $\hat{\tau}_2$ взаимно независимы, тогда соотношения (35) и (36) примут вид

$$P\left(\hat{\tau}_2<\hat{\tau}_1\right)=\int\limits_{-\infty}^{\infty}F_{\hat{\tau}_2}(\tau_1)\mathrm{d}F_{\hat{\tau}_1}(\tau_1); \quad (37)$$

$$P\left(\hat{\tau}_1>\hat{\tau}_2\right)=\int\limits_{-\infty}^{\infty}R_{\hat{\tau}_1}(\tau_2)\mathrm{d}F_{\hat{\tau}_2}(\tau_2). \quad (38)$$

Введем следующие обозначения:

$$\hat{\omega}_1=\omega_1\left(\hat{\tau}_1\right)=F_{\hat{\tau}_2}(\tau_1); \quad (39)$$

$$\hat{\omega}_2=\omega_2\left(\hat{\tau}_2\right)=R_{\hat{\tau}_2}(\tau_2). \quad (40)$$

Случайные величины $\hat{\omega}_1$ и $\hat{\omega}_2$ называются стохастическими супериндикаторами, и с учетом введенных обозначений запишем

$$\left.\begin{aligned}P\left(\hat{\tau}_2<\hat{\tau}_1\right)=M\left[\hat{\omega}_1\right]=\bar{\omega}_1\\ P\left(\hat{\tau}_1>\hat{\tau}_2\right)=M\left[\hat{\omega}_2\right]=\bar{\omega}_2\end{aligned}\right\}\bar{\omega}_1=\bar{\omega}_2. \quad (41)$$

Соответственно:

$$P\left(\hat{\tau}_2<\hat{\tau}_1\right)=\bar{\omega}_1=\int\limits_{0}^{1}\omega\mathrm{d}F_{\hat{\omega}_1}(\omega)=\bar{\omega}_2=\int\limits_{0}^{1}\omega\mathrm{d}F_{\hat{\omega}_2}(\omega), \quad (42)$$

где $F_{\hat{\omega}_1}(\omega)$, $F_{\hat{\omega}_2}(\omega)$ — функции распределения супериндикаторов $\hat{\omega}_1$ и $\hat{\omega}_2$.

Основные свойства стохастических индикаторов

Константный $\left(\hat{\omega}_1\right)$ и функциональный $\left[\hat{\omega}_J(\tau_1)\right]$ индикаторы, определяющие соответственно апостериорные вероятности $\hat{J}$ и $\hat{J}_{\tau_1}$ обнаружения сценария J_m^r ДВ r-го типа, принимают лишь одно из двух значений (0 или 1), тогда как супериндикаторы $\hat{\omega}_1$ и $\hat{\omega}_2$ могут принимать бесконечное множество значений из интервала (0, 1], т. е. являются случайными величинами более общего типа.

Таким образом, СОИБ будет обнаруживать любой сценарий J_m^r ДВ r-го типа с априорной вероятностью, равной математическому ожиданию условной вероятности $M\left[\hat{\omega}_1\right]$. Если достоверность событий $\hat{J}$ и $\hat{J}_{\tau_1}$ принимает лишь одно из двух значений: 0 или 1 с вероятностями q, p соответственно, — то достоверность события $\hat{J}_{\tau_1}$ распределена на интервале (0, 1] с плотностью $\varphi_{\hat{\omega}_1}(\omega)$ или $\varphi_{\hat{\omega}_2}(\omega)$.

Из равенств (34) и (40) следует, что $\hat{\omega}_J=P\left(\hat{\omega}_J=1\right)$; $\hat{\omega}_J(\tau_1)=P\left[\hat{\omega}_J(\tau_1)=1\right]$, но $\hat{\omega}_1\neq P\left(\hat{\omega}_1=1\right)$; $\hat{\omega}_2\neq P\left(\hat{\omega}_2=1\right)$. Различие является следствием того, что в последнем случае не только процесс обнаружения СОИБ в ЕИП сценария J_m^r ДВ r-го типа неопределенный, но и степень достоверности обнаружения $C=\left\{c_1^r, c_2^r\right\}$ АСИБ в составе сценария J_m^r ДВ r-го типа на интервале времени $\tau_{\text{р.с}}$ реализации сценария J_m^r ДВ r-го типа случайна и может принимать значения, отличные от 0 и 1.

Таким образом, супериндикаторы $\hat{\omega}_1$ и $\hat{\omega}_2$ совмещают в себе свойства и функции $\omega\left(\hat{\tau}_1\right)$ случайного аргумента и случайной функции $\hat{\omega}(\tau_1)$ [см. (39), (40)].

Физический смысл заключается в следующем. Если интервал $\hat{\tau}_2$ времени обнаружения реализующегося c_1^r АСИБ случаен, то в предикате $\hat{\tau}_2 < c_1^r$ константа c_1^r определяет границу детерминированного множества $J = \left(-\infty, c_1^r\right)$, а в предикате $\hat{\tau}_2 < \tau_1$ переменная τ_1 определяет границу переменного множества $J = (-\infty, \tau_1)$, при попадании в которое случайной величины $\hat{\tau}_2$ индикаторы $\hat{\omega}_J$ и $\hat{\omega}_J(\tau_1)$ принимают значение 1. В предикате $\hat{\tau}_2 < \hat{\tau}_1$ переменная $\hat{\tau}_1$ определяет границу «неопределенного» множества $J_{\hat{\tau}_1} = \left(-\infty, \hat{\tau}_1\right)$, при попадании в которое случайной величины $\hat{\tau}_2$ индикаторы $\hat{\omega}_1$, $\hat{\omega}_2$ могут принять уже любые значения из интервала (0, 1].

Применение специализированной модели целенаправленного процесса для построения стохастических временных индикаторов

Математической моделью обнаружения СОИБ любого сценария ДВ и предъявляемых к оперативности их функционирования требований является оперативный Т-процесс [7].

С учетом вышеизложенного приведены стохастические временные индикаторы, характеризующие сценарии ДВ:

— ω_l^{τ} — стохастический супериндикатор продолжительности обнаружения $\hat{\tau}_c$ l-го АСИБ;

— $\omega_{l,l+1}^{t}$ — стохастический супериндикатор очередности обнаружения c_l^r АСИБ в составе реализующегося сценария J_m^r ДВ r-го типа на основе оценивания случайной очередности момента $\hat{t}_l^{\text{зв}}$ завершения l-го АСИБ и момента $\hat{t}_{l+1}^{\text{нч}}$ начала $(l + 1)$-й реализации следующего;

— $\overline{\omega}_{l,l+1}^{\tau}$ — стохастический супериндикатор очередности завершения обнаружения двух АСИБ на основе оценивания продолжительности $\hat{\tau}_l$, $\hat{\tau}_{l+1}$ l-го и $(l + 1)$-го воздействия соответственно.

Для удобства понимания дальнейшего изложения текста обозначим данные индикаторы символом ω_1, который носит название стохастического супериндикатора первого порядка, соответственно примем

$$\overline{\omega}_1 = P\left(\hat{\tau}' \le \hat{\tau}''\right). \tag{43}$$

С учетом сказанного показатель эффективности (ПЭ) $P_{\text{д.ц}}^1$ обнаружения сценария ДВ запишется в виде

$$P_{\text{д.ц}}^1 = P\left(\hat{\tau}' \le \hat{\tau}''\right). \tag{44}$$

В рамках методологии следует отметить, что величина $\hat{\tau}'$ представляет собой минимально необходимый целевой эффект, выраженный во временных единицах с учетом эффекта поглощения по операционным ресурсам [7].

Тогда, например, значение ПЭ $P_{\text{д.ц}}^2$ процесса функционирования СОИБ (вероятность того, что злоумышленник не реализует сценарий ДВ) определяется выражением

$$P_{\text{д.ц}}^2 = 1 - P_{\text{д.ц}}^1. \tag{45}$$

Для определения значения вероятности $P_{\text{д.ц}}^1$ необходимо и достаточно знать законы распределения случайных величин $\hat{\tau}'$ и $\hat{\tau}''$, тогда

$$P_{\text{д.ц}}^1 = P\left(\hat{\tau}' \le \hat{\tau}''\right) = P\left[\left(\hat{\tau}', \hat{\tau}''\right) \in H\right] = \iint\limits_{(H)} \varphi_{\langle \hat{\tau}', \hat{\tau}'' \rangle}(\tau', \tau'')\,d\tau' d\tau'', \tag{46}$$

где $(H) \cong \{\langle \tau', \tau'' \rangle : \tau' < \tau''\}$.

Пусть случайные величины $\hat{\tau}'$ и $\hat{\tau}''$ подчинены смещенным показательным законам распределения с параметрами соответственно λ_1, τ_1 и λ_2, τ_2, т. е.

$$F_{\hat{\tau}'}(\tau') = \left[1 - e^{-\lambda_1\left(\hat{\tau}' - \tau_1\right)}\right]\Delta\left(\hat{\tau}' - \tau_1\right). \tag{47}$$

Тогда функция $F_{\hat{\omega}_1}(\omega)$ распределения стохастического супериндикатора $\hat{\omega}_1$ принимает вид

$$F_{\hat{\omega}_1}(\omega) = F_{\hat{\tau}''}\left[F_{\hat{\tau}'}^{-1}(\omega)\right] = \left[1 - e^{-\lambda_2(\tau_1 - \tau_2)}(1 - \omega)^{\frac{\lambda_2}{\lambda_1}}\right] \times \times \prod\left(\omega;\ \sup\left\{0, F_{\hat{\tau}'}(\lambda_2)\right\}, 1\right) + \Delta(\omega - 1). \tag{48}$$

Плотность распределения

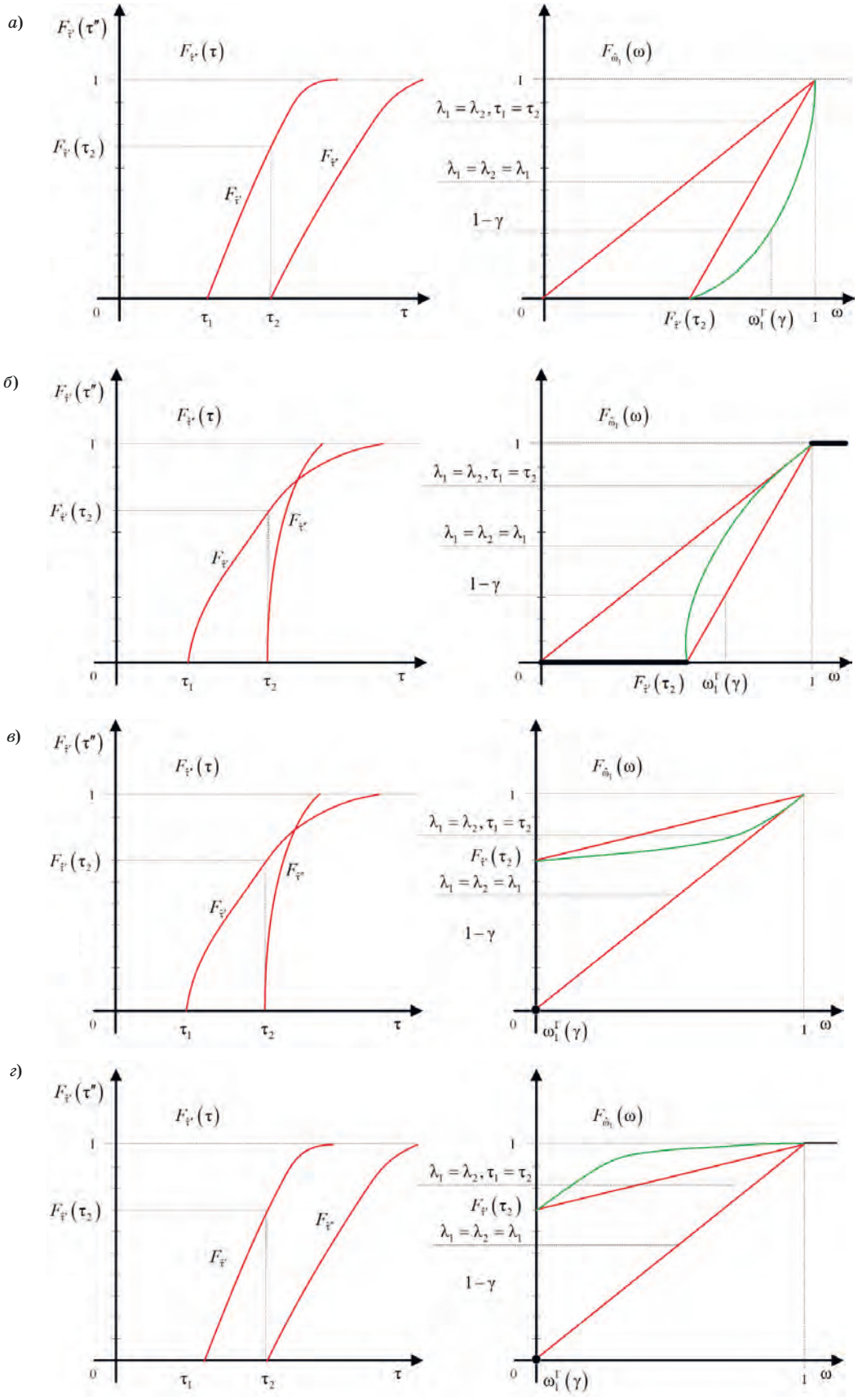
$$\varphi_{\hat{\omega}_1}(\omega) = \frac{\lambda_2}{\lambda_1} e^{\lambda_2(\tau_2 - \tau_1)} \times \times (1 - \omega)^{\frac{\lambda_2}{\lambda_1} - 1} \prod\left(\omega;\ \sup\left\{0, F_{\hat{\tau}'}(\lambda_2)\right\}, 1\right). \tag{49}$$

В результате

$$P_{\text{д.ц}}^1 = \overline{\omega}_1 = \left[1 - \frac{\lambda_2}{\lambda_2 + \lambda_1} e^{-\lambda_1(\tau_2 - \tau_1)}\right] \times \Delta(\tau_2 - \tau_1) + + \frac{\lambda_1}{\lambda_2 + \lambda_1} e^{-\lambda_2(\tau_1 - \tau_2)} \Delta(\tau_1 - \tau_2). \tag{50}$$

Тогда гарантируемая вероятность

$$\omega_1^{\Gamma} = \left[1 - \gamma^{\frac{\lambda_1}{\lambda_2}} e^{\lambda_1(\tau_1 - \tau_2)}\right] \times \times \left[\Delta(\tau_2 - \tau_1) + \Delta\left(R_{\hat{\tau}'}(\tau_1) - \gamma\right)\Delta(\tau_1 - \tau_2)\right], \tag{51}$$



■ ***Рис. 6.*** Функции $F_{\hat{\omega}_1}(\omega)$ для первого (*а*), второго (*б*), третьего (*в*) и четвертого (*г*) случая

■ ***Fig. 6.*** The functions $F_{\hat{\omega}_1}(\omega)$ for the first (*а*), second (*б*), third (*в*), and fourth (*г*) cases

где γ — уровень гарантии (гарантийная вероятность).

В зависимости от соотношений параметров τ_1, τ_2 и λ_1, λ_2 функция распределения $F_{\hat{\omega}_1}(\omega)$ стохастического супериндикатора $\hat{\omega}_1$ принимает конкретный вид, соответствующий одному из четырех вариантов:

1) $\tau_1 < \tau_2$, $\lambda_1 > \lambda_2$;
2) $\tau_1 < \tau_2$, $\lambda_1 < \lambda_2$;
3) $\tau_1 > \tau_2$, $\lambda_1 < \lambda_2$;
4) $\tau_1 > \tau_2$, $\lambda_1 > \lambda_2$.

В частности, если $\tau_1 < \tau_2 \wedge (\lambda_1 > \lambda_2 \vee \lambda_1 < \lambda_2)$, то для случаев 1) и 2) функция распределения $F_{\hat{\omega}_1}(\omega)$ стохастического супериндикатора $\hat{\omega}_1$ примет вид

$$F_{\hat{\omega}_1}(\omega) = \left[1 - e^{\lambda_2(\tau_2-\tau_1)}(1-\omega)^{\frac{\lambda_2}{\lambda_1}}\right] \times \times \prod\left(\omega;\, F_{\hat{\tau}'}(\tau_2),\, 1\right) + \Delta(\omega - 1); \quad (52)$$

$$\varphi_{\hat{\omega}_1}(\omega) = \frac{\lambda_2}{\lambda_1} e^{\upsilon(\tau_2-\tau_1)} \times \times (1-\omega)^{\frac{\lambda_2}{\lambda_1}-1} \prod\left(\omega;\, F_{\hat{\tau}'}(\tau_2),\, 1\right); \quad (53)$$

$$\bar{\omega}_1 = 1 - \frac{\lambda_2}{\lambda_2 + \lambda_1} e^{\lambda_1(\tau_1-\tau_2)}; \quad (54)$$

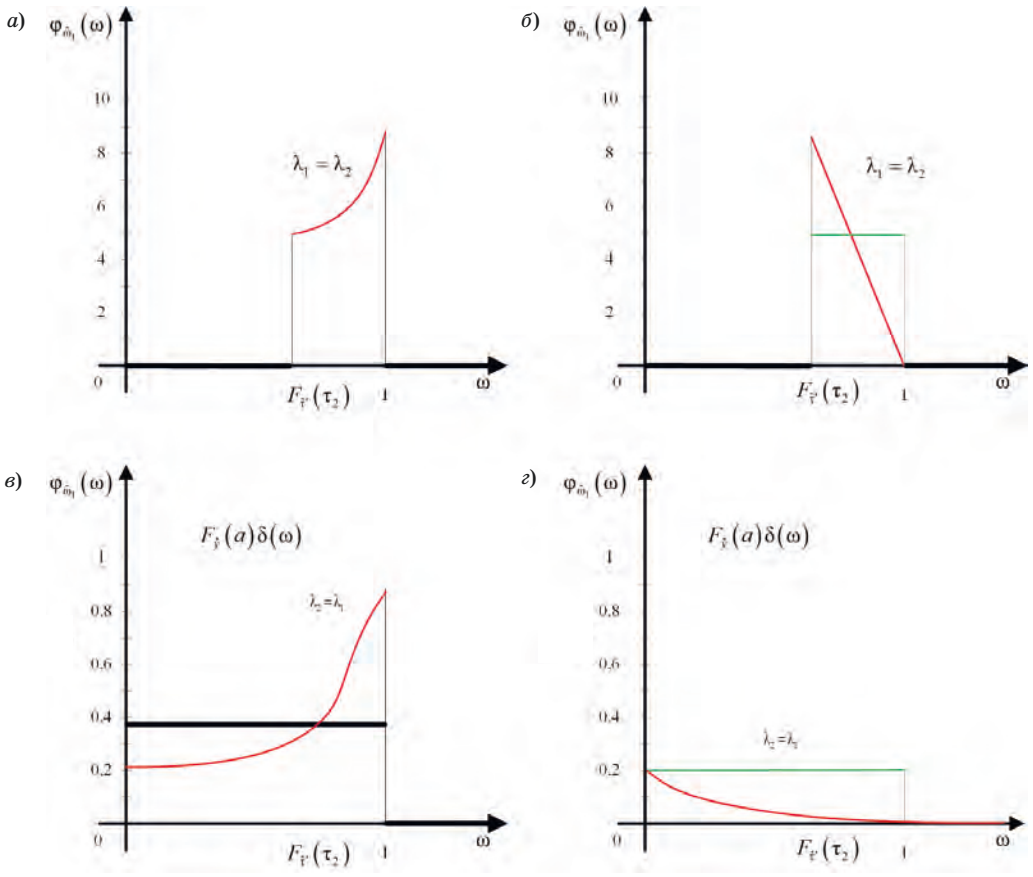
$$\omega_1^{\Gamma}(\gamma) = 1 - \gamma^{\frac{\lambda_1}{\lambda_2}} e^{\lambda_1(\tau_1-\tau_2)}. \quad (55)$$

Для случаев 3) и 4) при условии $\tau_1 > \tau_2 \wedge \wedge (\lambda_1 < \lambda_2 \vee \lambda_1 > \lambda_2)$ функция распределения $F_{\hat{\omega}_1}(\omega)$ стохастического супериндикатора $\hat{\omega}_1$ примет вид

$$F_{\hat{\omega}_1}(\omega) = = \left[1 - e^{\lambda_2(\tau_2-\tau_2)}(1-\omega)^{\frac{\lambda_2}{\lambda_1}}\right] \prod(\omega;\, 0,1) + + \Delta(\omega - 1); \quad (56)$$

$$\varphi_{\hat{\omega}_1}(\omega) = = \frac{\lambda_2}{\lambda_1} e^{\lambda_2(\tau_2-\tau_1)}(1-\omega)^{\frac{\lambda_2}{\lambda_1}-1} \prod(\omega;\, 0,1) + + F_{\hat{\tau}''}(\tau_1)\delta(\omega); \quad (57)$$

$$\bar{\omega}_1 = 1 - \frac{\lambda_1}{\lambda_1 + \lambda_2} e^{\lambda_2(\tau_2-\tau_1)}; \quad (58)$$



■ ***Рис. 7.*** Кривые распределения индикатора $\hat{\omega}_1$ для первого (*а*), второго (*б*), третьего (*в*) и четвертого (*г*) случая
■ ***Fig. 7.*** Distribution curves of the indicator $\hat{\omega}_1$ for the first (*а*), second (*б*), third (*в*), and fourth (*г*) cases

$$\omega_1^{\Gamma}(\gamma)=\left[1-\gamma^{\frac{\lambda_1}{\lambda_2}e^{\lambda_1(\tau_1-\tau_2)}}\right]\Delta\left(F_{\hat{\tau}'}(\tau_1)-\gamma\right). \quad (59)$$

Графики функции $F_{\hat{\omega}_1}(\omega)$ для случаев 1) и 2), когда значения параметров τ_1, τ_2 и λ_1, λ_2 исходных распределений $F_{\hat{\tau}'}(\tau')$ и $F_{\hat{\tau}''}(\tau'')$ совпадают и $\lambda_1=\lambda_2=\lambda_1$, приведены на рис. 6, *а* и *б*.

При совпадении исходных распределений $F_{\hat{\tau}''}(\tau'')$ значений параметров τ_1, τ_2 и λ_1, λ_2 функции $F_{\hat{\omega}_1}(\omega)$ и при $\lambda_2=\lambda_1=\lambda_2$, для случаев 3) и 4) существует обратная зависимость относительно случая 1) и 2), которая отображена на рис. 6, *в* и *г*.

Кривые распределения индикатора $\hat{\omega}_1$ для всех указанных случаев приведены на рис. 7, *а–г*.

Действительно, в зависимости от соотношений параметров τ_1, τ_2 и λ_1, λ_2 функция распределения $F_{\hat{\omega}_1}(\omega)$ стохастического супериндикатора $\hat{\omega}_1$ принимает конкретный вид, соответствующий одному из четырех приведенных выше вариантов.

Заключение

Исследование эффективности процесса функционирования СОИБ комплексно и корректно может быть осуществлено только на методологической основе современной теории эффективности целенаправленных процессов. В этом случае исследователю удается учесть весь комплекс результатов процесса функционирования системы — как положительных (объем и качество целевого эффекта), так и отрицательных (расходы ресурсов и времени).

Разработан подход к оцениванию эффективности процессов функционирования системы обеспечения информационной безопасности на основе теории стохастической индикации, призванной служить инструментом вероятностного анализа случайных явлений.

Формализована стохастическая ситуация для количественного анализа исследования процессов функционирования систем обеспечения информационной безопасности. Рассмотрены основные свойства стохастических индикаторов. Построены стохастические временные индикаторы на основе специализированной модели целенаправленного процесса.

При решении задач исследования эффективности следует использовать специфичные агрегированные модели системы, которые отражают с требуемой адекватностью результаты процесса их функционирования, динамику их получения в ходе операции и их связи с параметрами и эксплуатационно-техническими характеристиками СОИБ и ее процессом функционирования, без детального описания всех элементов системы.

Литература

1. **Daraio C., Simar L.** *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications.* Springer, 2007. 263 p.
2. **Зуев М. Б., Зуев Б. П., Булгакова И. Н.** Усовершенствованный метод освоенного объема для интегральной оценки эффективности и прогнозов результата деятельности в сфере управления. *Управление проектами: идеи, ценности, решения: материалы I Междунар. науч.-практ. конф.*, Санкт-Петербург, 15–17 мая 2019 г. СПб., СПбГАСУ, 2019, с. 80–87.
3. **Юсупов Р. М., Мусаев А. А.** Особенности оценивания эффективности информационных систем и технологий. *Тр. СПИИРАН*, 2017, вып. 1(51), с. 5–34. doi:10.15622/sp.51.1
4. **Арсеньев В. Н., Хомоненко А. Д., Ядренкин А. А.** Взвешенный учет априорной и опытной информации в задаче оценивания эффективности функционирования системы управления при распределении числа испытаний по закону Паскаля. *Информационно-управляющие системы*, 2020, № 3, с. 39–47. doi:10.31799/1684-8853-2020-3-39-47
5. **Беляков М. И.** Модель процесса функционирования системы обеспечения информационной безопасности объекта критической информационной инфраструктуры в задаче оценивания его эффективности. *Вопросы оборонной техники. Серия 16: Технические средства противодействия терроризму*, 2020, № 11-12 (149-150), с. 71–75.
6. **Wetering V., Mikalef P., Adamantia P.** A strategic alignment model for IT flexibility and dynamic capabilities: Toward an assessment tool. *Twenty-Fifth European Conf. on Information Systems (ECIS)*, 2017, pp. 1–17.
7. **Петухов Г. Б., Якунин В. И.** *Методологические основы внешнего проектирования целенаправленных процессов и целеустремленных систем*. М., АСТ, 2006. 504 с.
8. **Фролов О. П., Кузьмин В. Н., Зиннуров С. Х.** Методологический подход к решению слабоформализуемых задач оценивания эффективности и выбора рациональных способов применения космических систем. *Изв. Российской академии ракетных и артиллерийских наук*, 2020, № 3 (113), с. 59–65.
9. **Галанкин А. В., Гончаров А. М., Чащин С. В.** Оценивание эффективности функционирования цифровой сети связи космических войск. *Тр. Военно-космической академии им. А. Ф. Можайского*, 2016, № 650, с. 7–10.

10. **McMahon P.** *15 Fundamentals for Higher Performance in Software Development: Includes discussions on CMMI, Lean Six Sigma, Agile and SEMAT's Essence Framework*. Pem Systems Publ., 2014. 336 p.
11. **Гейда А. С., Лысенко И. В., Юсупов Р. М.** Основные концепты и принципы исследования операционных свойств использования информационных технологий. *Тр. СПИИРАН*, 2015, вып. 5(42), с. 5–36. doi:10.15622/sp.42.1
12. **Гейда А. С., Исмаилова З. Ф., Клитный И. В., Лысенко И. В.** Задачи исследования операционных и обменных свойств систем. *Тр. СПИИРАН*, 2014, вып. 4(35), с. 136–160. doi:10.15622/sp.35.10
13. **Schilke O., Hu S., Helfat C.** Quo vadis, dynamic capabilities? A content-analytic review of the current state of knowledge and recommendations for future research. *Academy of Management Annals*, 2018, vol. 12, no. 1, pp. 390–439.
14. **Garza-Reyes J.** From measuring overall equipment effectiveness (OEE) to overall resource effectiveness (ORE). *Journal of Quality in Maintenance Engineering*, 2015, vol. 21(4), pp. 506–527.
15. **Сухов А. М., Крупенин А. В., Якунин В. И.** Методы анализа и синтеза исследования эффективности процессов функционирования системы обнаружения предупреждения и ликвидации последствий компьютерных атак. *Автоматизация процессов управления*, 2021, № 4 (66), с. 4–14.
16. **Сухов А. М., Герасимов С. Ю., Еремеев М. А., Якунин В. И.** Математическая модель процесса функционирования подсистемы реагирования системы обнаружения, предупреждения и ликвидации последствий компьютерных атак. *Проблемы информационной безопасности. Компьютерные системы*, 2019, № 2, с. 56–64.
17. **Браницкий А. А., Котенко И. В.** Анализ и классификация методов обнаружения сетевых атак. *Тр. СПИИРАН*, 2016, вып. 2(45), с. 207–244. doi:10.15622/SP.45.13
18. **Юрчик П. Ф., Андрющенко В. И., Шастин С. Д.** Формирование архитектуры единого информационного пространства. *Школа Науки*, 2021, № 1 (38), с. 33–36.
19. **Горбачев И. Е., Сухов А. М., Еремеев М. А., Смирнов С. И.** Методика реализации системного подхода при создании облика системы информационной безопасности критической информационной инфраструктуры с учетом экономической целесообразности. *Проблемы информационной безопасности. Компьютерные системы*, 2018, № 2, с. 93–110.
20. **Сухов А. М., Ступин Д. Д., Ломако А. Г.** Модель проактивного обнаружения компьютерных атак. *Проблемы управления и моделирования в сложных системах: тр. XX Междунар. конф.*, Самара, 3–6 сентября 2018 г.; под ред. Е. А. Федосова, Н. А. Кузнецова, С. Ю. Боровика. Самара, 2018, с. 509–512.
21. **Маликов А. В., Авраменко В. С., Саенко И. Б.** Модель и метод диагностирования компьютерных инцидентов в информационно-коммуникационных системах, основанные на глубоком машинном обучении. *Информационно-управляющие системы*, 2019, № 6, с. 32–42. doi:10.31799/1684-8853-2019-6-32-42

UDC 519.718
doi:10.31799/1684-8853-2022-3-31-44

Evaluating the effectiveness of the information security system process based on the theory of stochastic indicators

A. M. Sukhov[a], PhD, Tech., Doctoral Candidate, orcid.org/0000-0003-2233-811X, 19am87@mail.ru
[a]Krasnodar Higher Military School named after General of the Army S. M. Shtemenko, 4, Krasin St., 350065, Krasnodar, Russian Federation

Introduction: Under the conditions of imperfect methods and means of detection and response to computer attacks there is a constant growth of destructive impacts aimed at critical information systems. This generates a need to develop research methods for early warning systems to provide information security in case of malware attacks. One of the effective ways to solve this problem is to use the methods of the theory of stochastic indicators. **Purpose:** The development of a tool for evaluating the effectiveness of the information security system functioning. **Results:** We describe deterministic, random and indefinite components of the information security system functioning. Constant and functional indicators are constructed, their distinctive features are revealed. To solve the problem of evaluating the effectiveness of the process under consideration stochastic superindicators are constructed. We have also described the features of the construction of stochastic indicators of different ranks on the basis of the theory of the effectiveness of targeted processes and purposeful systems. **Practical relevance:** Through the developed stochastic time indicators, the probabilistic and temporal characteristics of the destructive impact are estimated, with the intervals and time points of its occurrence taken into account. This allows the system to be timely warned of a possible destructive impact scenario for the elements of critical information infrastructure.

Keywords — information security system, atomic information security event, destructive impact, efficiency, quality, stochastic indication theory.

For citation: Sukhov A. M. Evaluating the effectiveness of the information security system process based on the theory of stochastic indicators. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 31–44 (In Russian). doi:10.31799/1684-8853-2022-3-31-44

References

1. Daraio C., Simar L. *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*. Springer, 2007. 263 p.
2. Zuev M. B., Zuev B. P., Bulgakova I. N. Upgraded method of the mastered volume for integrated assessment of efficiency and forecasts of result of activity in the management. *Materialy 1-j Mezhdunarodnoj nauchno-prakticheskoj konferencii "Upravlenie proektami: idei, cennosti, resheniya"* [Proc. 1st Intern. Scient. and Pract. Conf. "Project management: ideas, values, solutions"]. Saint-Petersburg, 2019, pp. 80–87 (In Russian).
3. Yusupov R. M., Musaev A. A. Efficiency of information systems and technologies: Features of estimation. *SPIIRAS Proc.*, 2017, vol. 2, no. 51, pp. 5–34 (In Russian). doi:10.15622/sp.51.1
4. Arseniev V. N., Khomonenko A. D., Yadrenkin A. A. Weighed ranking of aprioristic and experimental data in control system functioning efficiency estimation problem with Pascal-distributed number of tests. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 3, pp. 39–47 (In Russian). doi:10.31799/1684-8853-2020-3-39-47
5. Belyakov M. I. Model of the process of functioning of the information security system of a critical information infrastructure object in the assessment of its effectiveness. *Military Enginery. Iss. 16: Counter-terrorism technical devices*, 2020, no. 11-12 (149-150), pp. 37–40 (In Russian).
6. Mikalef P., Adamantia P. A strategic alignment model for IT flexibility and dynamic capabilities: Toward an assessment tool. *Twenty-Fifth European Conference on Information Systems (ECIS)*, 2017, pp. 1–17.
7. Petukhov G. B., Yakunin V. I. *Metodologicheskie osnovy vneshnego proektirovaniya celenapravlennyh processov i celeustremlennyh sistem* [Methodological basis of external designing of targeted processes and purposeful systems]. Moscow. AST Publ., 2006. 504 p. (In Russian).
8. Frolov O. P., Kuzmin V. N., Zinnurov S. H. Methodological approach to solving poorly formalized problems of evaluating the effectiveness and choosing rational ways of using space systems. *Izvestiya Rossijskoj akademii raketnyh i artillerijskih nauk*, 2020, no. 3 (113), pp. 59–65 (In Russian).
9. Galankin A. V., Goncharov A. M., Chashchin S. V. Evaluation of the effectiveness of the functioning of the digital communication network of the space forces. *Proc. of the Mozhaisky Military Space Academy*, 2016, no. 650, pp. 7–10 (In Russian).
10. McMahon P. *15 Fundamentals for Higher Performance in Software Development: Includes discussions on CMMI, Lean Six Sigma, Agile and SEMAT's Essence Framework*. Pem Systems Publ., 2014. 336 p.
11. Geida A. S., Lysenko I. V., Yusupov R. M. Main concepts and principles for information technologies operational properties research. *SPIIRAS Proc.*, 2015, vol. 5, no. 42, pp. 5–36 (In Russian). doi:10.15622/sp.42.1
12. Geida A. S., Ismailova Z. F., Klitnuy I. V., Lysenko I. V. Operational and exchange properties of systems research problems. *SPIIRAS Proc.*, 2014, vol. 4, no. 35, pp. 136–160 (In Russian). doi:10.15622/sp.35.10
13. Schilke O., Hu S., Helfat C. Quo vadis, dynamic capabilities? A content-analytic review of the current state of knowledge and recommendations for future research. *Academy of Management Annals*, 2018, vol. 12, no. 1, pp. 390–439.
14. Garza-Reyes J. From measuring overall equipment effectiveness (OEE) to overall resource effectiveness (ORE). *Journal of Quality in Maintenance Engineering*, 2015, vol. 21(4), pp. 506–527.
15. Sukhov A. M., Krupenin A. V., Yakunin V. I. The analysis and synthesis methods of research of the operation processes efficiency of the computer attacks detection, prevention and consequences elimination system. *Automated Control Systems*, 2021, no. 4 (66), pp. 4–14 (In Russian).
16. Sukhov A. M., Gerasimov S. Yu., Eremeev M. A., Yakunin V. I. Mathematical model of the process of functioning of detection system prevention and mitigation of computer attacks. *Information Security Problems. Computer Systems*, 2019, no. 2, pp. 56–64 (In Russian).
17. Branitskiy A. A., Kotenko I. V. Analysis and classification of methods for network attack detection. *SPIIRAS Proc.*, 2016, vol. 2, no. 45, pp. 207–244 (In Russian). doi:10.15622/SP.45.13
18. Yurchik P. F., Andryushchenko V. I., Shastin S. D. Formation of the architecture of a single information space. *School of Science*, 2021, no. 1 (38), pp. 33–36 (In Russian).
19. Gorbachev I. E., Sukhov A. M., Eremeev M. A., Smirnov S. I. The implementation of a systematic approach in creation of system of information security of critical information infrastructure taking into account economic feasibility. *Information Security Problems. Computer Systems,* 2018, no. 2, pp. 93–110 (In Russian).
20. Sukhov A. M., Stupin D. D., Lomako A. G. A model of proactive detection of computer attacks. *Trudy XX Mezhdunarodnoj konferencii "Problemy upravleniya i modelirovaniya v slozhnyh sistemah". Pod redakciej E. A. Fedosova, N. A. Kuznecova, S. Yu. Borovika* [Proc. of the XX Intern. Conf. "Complex systems: Control and modeling problems"]. Samara, 2018, pp. 509–512 (In Russian).
21. Malikov A. V., Avramenko V. S., Saenko I. B. Model and method for diagnosing computer incidents in information and communication systems based on deep machine learning. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 6, pp. 32–42 (In Russian). doi:10.31799/1684-8853-2019-6-32-42

UDC 003.26
doi:10.31799/1684-8853-2022-3-45-54

Articles

The variant of post-quantum cryptosystem based on burst-correcting codes and on the complete decoding problem

A. A. Ovchinnikov[a], *PhD, Tech., Associate Professor, orcid.org/0000-0002-8523-9429, mldoc@guap.ru*
[a]*Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation*

Introduction: *Today the investigations of post-quantum cryptosystems secure against quantum computations is the area of great interest. An important direction here is code-based cryptography utilizing the mathematical problems from error-correcting coding theory. The improvement of existing code-based systems may be achieved both in practical part (reducing the key sizes) and theoretically by using more complicated mathematical code-based tasks.* ***Purpose:*** *The development of public-key code-based cryptosystem using low-density parity-check codes with burst correction; the estimation of the parameters of the obtained system.* ***Results:*** *The variant of code-based cryptosystem using random block permutation low-density parity-check codes is proposed. The cryptocomplexity of the system is supposed to be based on the complete decoding problem, which is believed to be a harder mathematical problem than those used in existing systems. With high probability, the analysis of the system by using decoding methods is not possible at all, which both increases the long-term cryptocomplexity of the system and allows to reduce the key size. The evaluation of the underlying code selection is performed, the approaches to the selection of the parameters of the proposed system on the basis of the required level of cryptocomplexity are considered.* ***Practical relevance:*** *The proposed system allows to reduce the public-key size as compared to the classical McEliece system, cryptocomplexity also comparable, with the underlying mathematical problem to be more stable against perspective attacks.*

Keywords *— post-quantum cryptography, code-based public-key systems, low-density parity-check codes, burst error correction.*

For citation: Ovchinnikov A. A. The variant of post-quantum cryptosystem based on burst-correcting codes and on the complete decoding problem. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 45–54. doi:10.31799/1684-8853-2022-3-45-54

Introduction

The concept of public-key cryptography is usually connected with groundbreaking paper by W. Diffie, M. Hellman “New directions in cryptography” published in 1976 [1]. According to this concept each part has the pair of long-term keys: public key and correspondent private (secret) key. In case of secrecy providing the recipient’s public key is used during encryption, while the correspondent private key is used for decryption. Today the most widely spread public-key system is RSA whose strength is based on hardness of integer factorization. However, this problem is not belongs to NP-hard problems [2], besides, the quantum polynomial-time Shor’s algorithm is known for this task, so in middle-term perspective the strength of RSA becomes under question both in terms of classical computation architectures and by using powerful enough quantum computers. It worth to mention that there are intensive arithmetic with big integers (order of thousands of bits) being used in RSA system, so practical implementations of this system are rather slow.

As the result in 2016 NIST initiated the competition on adoption the new post-quantum cryptography standard [3]. One of the main directions within post-quantum cryptography is code-based cryptography, utilizing the problems from error-correcting codes theory.

The first code-based cryptosystem was proposed by R. McEliece in 1978. Being extremely computationally efficient, McEliece system, nevertheless, did not found wide practical usage, which is traditionally explained by relatively large key sizes, primarily for public key. Possible directions of McEliece system improvement are usage of error-correcting codes classes allowing decreasing the public key size, as well as selection of more complicated mathematical problems for system’s strengthening.

In this paper, the public-key system based on specific class of error-correcting low-density parity-check codes for bursts error correction is considered. The system uses the hard problem of complete decoding, which is NP-hard and potentially harder than the bounded-distance decoding problem used in McEliece cryptosystem.

Code-based hard problems

For investigating and understanding the details of different code-based cryptosystems the basics of underlying hard problems should be considered.

Public-key cryptography is based on the concept of one-way trap-door functions. Briefly the construction of such functions may be described as follows:

— (P, S) — key pair, where P — public key, S — private (secret) key;

— $E_P(m) = c$ — polynomial-time function, mapping the message m into ciphertext c using P;

— $D_S(c) = m$ — polynomial-time function, which is inverse to E, and uses S.

From the point of view of cryptographic strength the following should be provided:

— with knowledge of P, calculation of S should be computationally hard;

— with knowledge of c, and without knowledge of S, calculation of $E^{-1}(m) = c$ should be computationally hard.

By computational hardness it is supposed the exponential-time complexity of the correspondent problem, however, the specificity of one-way trap-door functions is that their inversion should be hard in general, but feasible (polynomial-time) with knowledge of secret S.

In the complexity theory there are approaches for problems classification by so-called "feasible" and "hard", one of the most widely used approach considers the following classes:

— class P (polynomial) — problems which may be solved by polynomial time on deterministic Turing machine;

— class NP (non-deterministic polynomial) — problems which may be solved by polynomial time on non-deterministic Turing machine (note P $\subseteq$ NP);

— class NPC (NP-complete) — problems which are in NP, and any other problem from NP can be reduced to them by polynomial time;

— class NP-hard — problems which may be not from NP, but any other NP-complete problem can be reduced to them by polynomial time.

More formal and accurate mathematical definitions are out of the scope of this paper, in cryptography the NP-hard problems are usually considered, but we will not make distinction between NP-complete and NP-hard problems.

Within this classification the problems from P are considered as feasible, while NPC or NP-hard contain hard problems (for which only exponential-time solutions are known in general case), however, polynomial-time specific cases are possible. We should also mention the existence of problems (denote them as "< NPC", which means "hard problems but simpler than NPC"), for which the polynomial-time solution in general case is unknown, but these problems are simpler than NPC in the sense that if their polynomial solution would be found this will not help to solve the problems from NPC. For example, such problems are integer factorization or discrete logarithm problem that are used in most practically spread number-theoretic cryptosystems.

Consequently, the following classes may be used to construct the one-way trap-door functions:

— < NPC — widely used in cryptography for today, but it is believed that there are the possibility of finding the polynomial-time solutions for these problems, besides, number-theoretic problems from this class have quantum polynomial complexity (may be solved by polynomial time using quantum computer);

— NPC — it is believed that polynomial solution for this class do not exist at all (though this is not proved mathematically), there are no polynomial time quantum algorithms known for this class.

From the classification given above it follows that NP-complete (or NP-hard) problems are preferable for usage in cryptography, but the distinction should be made between the cryptosystem (i.e. trap-door function) and underlying hard mathematical problem — it may be turned out that trap-door function does not belong to the same class as correspondent hard problem, for example the Merkle — Hellman system was broken by A. Shamir by polynomial time [1], though correspondent subset-sum problem is NP-hard.

Next we describe several hard problems from the coding theory, for this goal some definitions and terms should be given.

Linear (n, k)-code is k-dimension subspace of n-dimension linear vector space over the field F (in this paper only binary codes over $GF(2)$ are considered) [4, 5]. We assume $k < n$, then k is the number of information symbols, n is codelength, the value $r = n - k$ defines the number of redundant symbols, $R = k/n$ is code rate. Since linear code is linear vector space, it may be defined by its basis $\mathbf{G}$, which is $(k \times n)$-matrix called the generator matrix of the code. Basis of the orthogonal space is $(r \times n)$-matrix $\mathbf{H}$, which is called the parity-check matrix of the code, and $\mathbf{G}\mathbf{H}^{\mathrm{T}} = \mathbf{0}$. If $\mathbf{m}$ is k-bit information vector, then $\mathbf{a} = \mathbf{m}\mathbf{G}$ is codeword of length n, the vector $\mathbf{S} = \mathbf{b}\mathbf{H}^{\mathrm{T}}$ is called the syndrome for arbitrary vector $\mathbf{b}$ of length n, and $\mathbf{S} = \mathbf{0}$ iff $\mathbf{b}$ is codeword.

Let C is the set of codewords, $\mathbf{a} \in C$ — any codeword of length n, $\mathbf{b}$ is arbitrary vector of length n. The difference between $\mathbf{b}$ and $\mathbf{a}$ may be described by the so-called error vector $\mathbf{e} = \mathbf{b} - \mathbf{a}$ (we assume binary arithmetic which uses XOR), or $\mathbf{b} = \mathbf{a} + \mathbf{e}$.

The problem of minimal distance decoding is an optimization problem

$$\hat{\mathbf{a}} = \arg\min_{\mathbf{a} \in C} d(\mathbf{a}, \mathbf{b}), \tag{1}$$

where $d(\mathbf{a}, \mathbf{b})$ is Hamming distance between $\mathbf{a}$ and $\mathbf{b}$.

The problem of bounded-distance decoding, or decoding in sphere with radius t is an optimization problem (1) with additional constraints:

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a} \in C, d(\mathbf{a}, \mathbf{b}) \le t} d(\mathbf{a}, \mathbf{b}). \tag{2}$$

Note that the solution of (2) is not always exists, and $d(\mathbf{a},\mathbf{b}) = W(\mathbf{b} - \mathbf{a}) = W(\mathbf{e})$, where $W(\mathbf{e})$ is Hamming weight of $\mathbf{e}$.

The minimal distance d_0 of the code is the minimal pairwise Hamming distance between codewords. Then the code can correct any combination of t errors or less, where $d_0 = 2t + 1$, this means that if no more than t symbols are incorrect in codeword $\mathbf{a}'$, i.e. $\mathbf{b} = \mathbf{a}' + \mathbf{e}$, $W(\mathbf{e}) \le t$, then the problem (2) of bounded-distance decoding in sphere with radius t always has exactly one solution, which is $\hat{\mathbf{a}} = \mathbf{a}'$.

Linear (n, k)-code split the overall n-dimensional vector space into 2^r disjoint sets, one of them is the set of codewords and others are cosets. All vectors from the coset has the same syndrome (which is zero vector for the set of codewords). From any coset one representative may be chosen which is called the coset leader (zero codeword for the set of codewords). Since there are 2^r leaders and also 2^r different syndromes, the one-to-one mapping may be set between them, allowing to define the syndrome decoding procedure as calculation the syndrome $\mathbf{S} = \mathbf{bH}^{\mathrm{T}}$ for the vector $\mathbf{b}$, then the leader of the coset with correspondent $\mathbf{S}$ is considered as error vector $\mathbf{e}$, and the decoded codeword is $\hat{\mathbf{a}} = \mathbf{b} - \mathbf{e}$. If the coset leader is chosen as the vector with minimal weight from the coset, then syndrome decoding coincides with minimal distance decoding [4–6].

Note that in fact the list of coset leaders coincides with the set of errors correctable by the code. We will call the decoding, allowing correction of any coset leader, as complete decoding. Clearly, bounded-distance decoding is incomplete: only the subset of leaders with weight of no more than t may be corrected.

For the random linear code it is proved that the following problems are NP-hard:

— minimal distance decoding;

— complete decoding;

— calculation of the code's minimal distance;

— calculation of the non-zero codeword of minimal weight.

Note that the bounded-distance decoding problem is not in the list, though there are different points of view concerning NP-hardness of this problem, however, to the author's knowledge, formal proof of any correspondent hypothesis is unknown. It should be mentioned that the listed problems are hard for random linear codes, while for some specific code constructions simple solutions are known, this allows usage of coding problems in construction of public-key cryptosystems.

Classical code-based public-key cryptosystems

The idea of the McEliece system [1, 7, 8] is to select the error-correcting code, for which effective (polynomial-time) decoding algorithm is known, and to hide the structure of this code in linear code of random structure. This idea is realized as follows.

1. Key generation.

Each entity U performs the following.

— Select generator $(k \times n)$-matrix $\mathbf{G}$ of linear code, which can correct t errors (has minimal distance $d_0 \ge 2t + 1$), and for which the polynomial-time bounded-distance decoding procedure ψ is known (in the sphere of radius t).

— Compute $\mathbf{G}' = \mathbf{MGP}$, where $\mathbf{M}$ — non-singular $(k \times k)$-matrix, $\mathbf{P}$ — $(n \times n)$-permutation matrix.

— Public key is $P_{\mathrm{U}} = (\mathbf{G}', t)$, private key is $S_{\mathrm{U}} = (\mathbf{M}, \mathbf{G}, \mathbf{P})$.

2. Encryption.

Entity A encrypts k-bit message $\mathbf{m}$, using authentic public key P_{B} of entity B.

— A computes $\mathbf{c} = \mathbf{mG}' + \mathbf{e}$, where $\mathbf{e}$ is random binary vector of length n and weight t.

3. Decryption.

Entity B decrypts $\mathbf{c}$, using his private key S_{B}.

— Compute $\mathbf{x} = \mathbf{cP}^{-1}$.

— Compute $\psi(\mathbf{x}) = \mathbf{m}$.

McEliece proposed to use Goppa codes as private code. This codes are cyclic and can be decoded in polynomial time by decoders constructed using algebra for polynomials [4, 5]. Public key here is the code equivalent to private code (i.e. obtained by the coordinates permutation). It is supposed that the code equivalent to Goppa code can not be distinguished from the random code, though it is known that this is not true in some cases [7]. Additional requirement to private code is that code construction should allow exponentially large key space for given parameters of the code.

Analysis of McEliece cryptosystem may be performed in two directions. First, this is the recovering of the private code's structure from the public code. In fact this is the analysis of masking transformation, which is permutation in case of McEliece cryptosystem. In worst case this requires considering all permutations of length n, which is clearly infeasible.

Second, and this is counted as the main attack on McEliece system, is an attempt to correct t errors in ciphertext $\mathbf{c}$ and find the codeword in code $\mathbf{G}'$, i.e. solving the decoding problem in the sphere of radius t for the code which considered as random. Best known approach to solve this task for today is information set decoding [8–11]. Note that equivalent code has the same minimal distance as initial code, so bounded-distance decoding will find the correct codeword with probability 1, so the attack is limited only by computational complexity.

In the first variant of the system McEliece proposed to use (1024, 524)-code correcting 50 errors. Comparatively up-to-date review of decoding methods given in [8] mentioned that this parameters are

correspondent to cryptocomplexity equal to 2^{53}, to achieve level of 2^{94} the matrix size should be 1036×2048 (correcting 92 errors), and matrix size 2056×4096 (correcting 170 errors) provide system strength of 2^{171}. In general the key sizes of this system have the order of hundreds of thousands bits. In many situations this is not excessive requirement, but traditional point of view is that this is the main drawback of the McEliece system.

The following directions of McEliece system improvement may be formulated:

1) reducing the key sizes by usage of special classes of Goppa codes, or alternative error-correction codes;

2) increasing of system's strength, first of all by strengthening the masking transformation between public and private keys.

In 1986 H. Niederreiter proposed the code-based system, for which later its equivalence to McEliece system was proven [8], but having some practical advantages. In this paper we do not consider this approach.

In the last decade the significant direction of McEliece system evolution is usage of block-circulant matrices for public and private codes, such matrices define the so called quasi-cyclic (QC) codes and allow significant reduction of key sizes during storage and transportation by means of circulant structure. To provide the polynomial-time decoding procedure, the private key is selected as sparse matrix, in this keys the decoding algorithms for low-density parity-check (LDPC) codes may be used [8, 12, 13]. In some cases of such systems the masking transformation is no longer the permutation matrix and selected in a special way (however, this transformation matrices should also be sparse to avoid large increasing of the number of errors corrected during decryption), however, in all such systems the underlying problem is bounded-distance decoding.

Public-key cryprosystems based on complete decoding problem

As it was mentioned in the previous section, there are modifications of McEliece system considering other classes of codes, and in some cases the special transformation matrices are considered instead of permutation matrix to hide the secret key in the public key. However, the fundamentally qualitative modification would be consideration totally random matrix for masking operation instead of permutation matrix or its analogues. In this case not only the public key is no longer defines the equivalent codes, but the number of qualitative new properties of the system are appeared.

Initially this approach was proposed by E. Krouk in 1993 [14] and later considered in the number of publications [15]. Let us describe the general structure of the system.

1. Key generation.

Each entity U performs the following.

— Select generator $(k \times n)$-matrix $\mathbf{G}$ of linear codes, for which the polynomial-time decoding procedure ψ is known, which corrects errors from the set E.

— Compute $\mathbf{G}' = \mathbf{GM}$, где $\mathbf{M}$ — $(n \times n)$ non-singular matrix.

— Define the set $E' = \{\mathbf{e}' : \mathbf{e}' = \mathbf{eM}, \mathbf{e} \in E\}$.

— Public key is $P_U = (\mathbf{G}', E')$, private key is $S_U = (\mathbf{G}, \mathbf{M})$.

2. Encryption.

Entity A encrypts k-bit message $\mathbf{m}$, using authentic public key P_B of entity B.

— A computes $\mathbf{c} = \mathbf{mG}' + \mathbf{e}'$, where $\mathbf{e}' \in E'$.

3. Decryption.

Entity B decrypts $\mathbf{c}$, using his private key S_B.

— Compute $\mathbf{x} = \mathbf{cM}^{-1}$.

— Compute $\psi(\mathbf{x}) = \mathbf{m}$.

The problem with implementation of described system is that there are two generalized sets of errors: the set E of errors, which should be corrected during decryption and the set E' of errors used during encryption (in McEliece system both sets consist of error vectors of weight t). Both sets should be exponentially large to avoid brute force, and at the same time they should have compact representation.

In described variant the vectors from these sets are connected with help of multiplication by $\mathbf{M}$, but this matrix is the part of private key, while vectors from E' should be generated by the party possessing only the public key.

From the other hand, suppose that this problem is somehow solved. Then, if we consider for example that E is the set of vectors of weight t, as in McEliece system, vector $\mathbf{e}' = \mathbf{eM}$, where $\mathbf{M}$ is random, has random weight, which is more probable close to $n/2$. Besides, the matrix $\mathbf{G}' = \mathbf{GM}$ defines the code with minimal distance which is more probable less than in private code $\mathbf{G}$. Thus, if the system is analyzed through decoding (to reconstruct $\mathbf{m}$ from $\mathbf{c}$), one should correct approximately $n/2$ errors in the code with probably small minimal distance, instead of solving the problem (2). This may lead to the situation when error vector is not within coset leaders at all, thus even solving the minimum distance decoding problem (1), which is complete decoding problem, will not give the correct codeword. In this case the problem of breaking the system is at least not simpler than complete decoding (though one should take in mind the possibility of breaking the system through analysis of the structure of public codes and matrix $\mathbf{M}$), thus we will call such system as based on the problem of complete decoding.

Next we describe more practical variant of the system based on the problem of complete decoding, giving an example of defining E and E' [15].

1. Key generation.

Each entity U performs the following.

— Select $(k \times n)$-generator matrix $\mathbf{G}$ of linear (n, k)-code, for which polynomial-time algorithm for correcting errors from some set E is known.

— Select random non-singular $(n \times n)$-matrix $\mathbf{M}_2$.

— Define the set $\tilde{E}$ and matrix $\mathbf{M}_1$ such that for any $\tilde{\mathbf{e}} \in \tilde{E}$ the vector $\tilde{\mathbf{e}}\mathbf{M}_1$ belongs to E (note that $\mathbf{M}_1$ may be singular).

— Compute $\mathbf{M} = \mathbf{M}_1\mathbf{M}_2$ (note that $\mathbf{M}$ singular if $\mathbf{M}_1$ is singular).

— Compute $(k \times n)$-matrix $\mathbf{G}' = \mathbf{G}\mathbf{M}_2$.

— Public key is $P_U = (\mathbf{G}', \mathbf{M}, \tilde{E})$, private key is $S_U = (\mathbf{G}, \mathbf{M}_1, \mathbf{M}_2)$.

2. Encryption.

Entity A encrypts k-bit message $\mathbf{m}$, using authentic public key P_B of entity B.

— A computes $\mathbf{c} = \mathbf{m}\mathbf{G}' + \mathbf{e}'$, where $\mathbf{e}' = \tilde{\mathbf{e}}\mathbf{M}$ for random $\tilde{\mathbf{e}} \in \tilde{E}$.

3. Decryption.

Entity B decrypts $\mathbf{c}$, using his private key S_B.

— Compute $\mathbf{x} = \mathbf{c}\mathbf{M}_2^{-1}$.

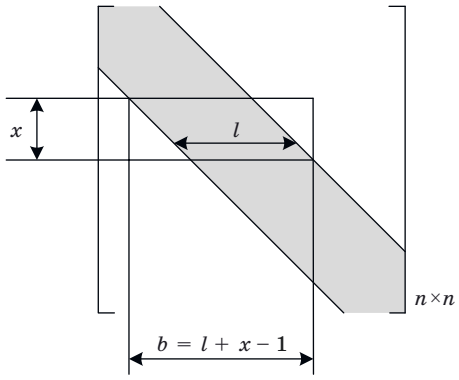
— Compute $\psi(\mathbf{x}) = \mathbf{m}$.

In this variant of the system the set E' is defined by vectors $\tilde{\mathbf{e}}\mathbf{M}$, which in turn requires effective description of $\tilde{E}$. Besides, the matrix $\mathbf{M}_1$ should be determined, mapping vectors from $\tilde{E}$ into E.

In particular, the set E itself may be selected as $\tilde{E}$, for example, consisting of all vectors of fixed weight, as in classical McEliece system. In this paper we consider the variant of the system which based on error-correcting codes which correct error bursts.

The effect of grouping errors in bursts (or packets) is typical for the most real communication channels, however, the codes that can correct such erroneous combinations are less investigated, and in practice the data transmitted via the channel is decorrelated using the interleaving procedure, and then the codes for independent errors correction are applied. In the case of cryptosystem development the errors in bursts may be formed artificially, in this case the positions and lengths of the bursts may be controlled.

The term of error burst itself may be defined in different ways. In this paper we define the burst of length b as binary error vector $\mathbf{e} = (e_0, \ldots, e_{n-1})$, in which the last non-zero element is placed no more than in b positions from the first. That is, if i is the minimal index for which $e_i = 1$, and j is maximal such index, then $\mathbf{e}$ forms the (single) error burst of length $b = j - i + 1$ at position i (thus two adjacent non-zero element form the packet of length 2). We will assume that positions of $\mathbf{e}$ from index i to j are filled by 1 and 0 with probability 1/2. Note that under the term "burst" one may consider not only the overall sequence $\mathbf{e}$, but its erroneous subsequence $(e_i, \ldots, e_j)$ without leading and ending zeros, the concrete sense of this term will be clear from the context.



■ Matrix $\mathbf{M}_1$

Similar to the fact that the minimal distance d_0 defines the maximal number t of independent errors, which can be corrected in any combination by minimal distance decoding, for each linear code the maximal correctable burst length b may be determined — this means that all possible error bursts of length no more than b are in different cosets and may be chosen as leaders. However, as it was mentioned earlier, finding the minimal distance of the random code is NP-hard, while maximal correctable burst length may be found in polynomial-time, using procedure from [16] (though the degree of the polynomial is rather large).

Let the set E consists of vectors which form error bursts of length no more than b. As the set $\tilde{E}$ we will also consider the set of bursts, but their length may differ from b and is defined by $\mathbf{M}_1$.

Consider $\mathbf{M}_1$ as $(n \times n)$-matrix in Figure. Here positions filled by random binary digits are marked in grey, other positions are zero. Clearly, such matrix through multiplication by it defines the mapping from bursts of length x into bursts of length b.

Then the above system may be additionally determined as follows:

— the set E: set of error bursts of length no more than b;

— matrix $\mathbf{G}$ defines the code, for which the polynomial-time procedure of correcting the error bursts of length b is known;

— the set $\tilde{E}$: set of error bursts of length no more than x. Clearly, public key is $P_U = (\mathbf{G}', \mathbf{M}, x)$.

In the next sections we will consider the selection of the code for the proposed system and estimation of its parameters.

Selection of the code for the system

Estimation of the quantitative parameters of the system considered in the previous section: burst

lengths x and b, cardinalities of sets E' and $\tilde{E}$, and finally selection of k and n, defining the key sizes, depends on the selection of class of burst-error-correction code. This class should contain exponential number of codes for given b, k, n, and admit polynomial-time procedure of correcting the bursts of length b. One of the variant of such a class is the class of low-density parity-check codes.

Low-density parity-check codes (LDPC codes) were proposed by R. Gallager in early 60-s [8, 12]. LDPC-code is defined by its parity-check matrix $\mathbf{H}$, containing low number of nonzero elements. The term "low number" is not formally defined, moreover, in the number of works on modified McEliece systems based on such codes the term "middle density" (MDPC) is used [7, 17, 18], but in both cases we may admit that we consider the codes with relatively sparse parity-check matrix, for which the decoders utilizing its sparseness show rather high correcting capability (low error probability). In general, LDPC codes are usually defined and analyzed as probabilistic ensembles of random codes with specific parameters, which is additional advantage for their usage as secret keys in code-based systems.

One of the most often used construction of LDPC codes is block-permutation construction, where the parity-check matrix has the form

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{1,1} & \mathbf{H}_{1,2} & \dots & \mathbf{H}_{1,\rho} \\ \mathbf{H}_{2,1} & \mathbf{H}_{2,2} & \dots & \mathbf{H}_{2,\rho} \\ \dots & \dots & \dots & \dots \\ \mathbf{H}_{\gamma,1} & \mathbf{H}_{\gamma,2} & \dots & \mathbf{H}_{\gamma,\rho} \end{bmatrix},$$

where $\mathbf{H}_{i,j}$ are sub-blocks of some structure. Usually some degree of $(m \times m)$-matrix of cyclic permutation is used as sub-blocks:

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

then the parity-check matrix of LDPC codes has the form

$$\mathbf{H} = \begin{bmatrix} \mathbf{C}^{i_{11}} & \mathbf{C}^{i_{12}} & \dots & \mathbf{C}^{i_{1\rho}} \\ \mathbf{C}^{i_{21}} & \mathbf{C}^{i_{22}} & \dots & \mathbf{C}^{i_{2\rho}} \\ \dots & \dots & \dots & \dots \\ \mathbf{C}^{i_{\gamma 1}} & \mathbf{C}^{i_{\gamma 2}} & \dots & \mathbf{C}^{i_{\gamma\rho}} \end{bmatrix}. \qquad (3)$$

It is known that some specific combinations of non-zero elements in the parity-check matrix may degrade the LDPC code decoder's performance, the simplest restriction to avoid some "bad" combinations is absence of two rows or two columns in the parity-check matrix having more than one common non-zero positions. If this is holds and we consider the matrix as incidence matrix of bipartite graph (the so called Tanner graph) there are no cycles of length 4 in the graph (for simplicity we will say that there are no cycles of length 4 in the matrix).

Traditionally LDPC codes are used to correct independent errors, however, in [16, 19, 20] the capability of these codes (in particular, block-permutation constructions) to correct bursts of errors was analyzed. In [16] both the procedure of determining the maximal length of correctable burst and decoding procedure for block-permutation LDPC code are described.

For the system described in the previous section we will use the ensemble of codes defined by (3). The ensemble is defined by the values of γ, ρ and m. As the additional requirement we demand the absence of 4-cycles in matrix (3).

As follows, for the fixed values of γ, ρ and m, which are selected mainly from the cryptocomplexity point of view, the probability of random selection of matrix with no cycles of length 4 should be estimated, as well as expected correctable burst lengths.

The software for determining the correctable burst length was implemented in Microsoft Visual Studio Enterprise 2017 using C++, Release x64. Experiments were hold using the computer with Windows 10 Pro, 16 Gb RAM, CPU Intel Core i7-4770K@3,50 GHz.

The matrices are considered with 3×4, 4×6, 4×8 blocks, with block sizes $m = 20, 31, 50, 61, 110, 127$. The distribution of correctable burst lengths, as well as the average time of determining the burst length are given in Table 1. For each set of parameters 100 random matrices were generated without cycles of length 4. One may note that for the block sizes which are prime numbers the lengths of correctable burst in all experiments are $b = m - 1$ [this is the maximal possible length of correctable burst for the block-permutation codes with parity-check matrix (3) with block size m]. In other cases it was found that for considered values of γ and ρ the burst lengths correctable by random codes is not significantly less than block size m, i.e. $b \approx m$.

The estimation of probability P of selecting the matrix containing cycles of length 4 in the Tanner graph was also performed, the results are given in Table 2. As can be seen from the table, when the block sizes are small, the probability of selecting the matrix with cycles of length 4 is rather high. But this probability decreases with block sizes growth, besides, determination of cycle existence in all considered cases takes less than microsecond, thus even if the probability of matrix without cycles

■ ***Table 1***. Examples of estimations of correctable burst lengths distributions and average processing time for different number of blocks in the parity-check matrix

$\gamma \times \rho$	Parameter	Block size m					
		20	31	50	61	110	127
3 × 4	Length b	19 — 10% 18 — 64% 16 — 19% 15 — 7%	30 — 100%	49 — 9% 48 — 75% 45 — 12% 40 — 4%	60 — 100%	109 — 8% 108 — 70% 105 — 14% 100 — 6% 99 — 2%	127 — 100%
	Time, s	0.01	0.04	0.38	0.63	11.64	12.43
4 × 6	Length b	19 — 21% 18 — 71% 16 — 4% 15 — 4%	30 — 100%	49 — 13% 48 — 80% 45 — 7%	60 — 100%	109 — 11% 108 — 81% 105 — 7% 100 — 1%	126 — 100%
	Time, s	0.04	0.18	2.68	2.86	50.92	80
4 × 8	Length b	19 — 2% 18 — 83% 16 — 12% 15 — 3%	30 — 100%	49 — 1% 48 — 88% 45 — 11%	60 — 100%	109 — 1% 108 — 84% 105 — 12% 100 — 3%	126 — 100%
	Time, s	0.10	0.44	6.26	6.75	80	171

■ ***Table 2***. Probability of existence of cycle of length 4

$\gamma \times \rho$	Parameter	Block size m					
		20	31	50	61	110	127
3 × 4	Probability P	0.62	0.46	0.31	0.26	0.15	0.13
	Time, μs	0.25	0.28	0.23	0.26	0.26	0.28
4 × 6	Probability P	0.99	0.96	0.85	0.78	0.57	0.51
	Time, μs	0.55	0.58	0.61	0.52	0.57	0.59
4 × 8	Probability P	0.999	0.997	0.97	0.95	0.8	0.75
	Time, μs	0.58	0.85	0.93	0.71	0.83	0.8

of length 4 is rather small, this matrix may be generated expectably fast by the sequence of random guesses (total number of such matrices is exponentially high). For example, for $\gamma = 4$, $\rho = 8$, $m = 20$ the probability of absence of 4-cycles is less than 0.001, but total number of matrices is $20^{32} \approx 2^{138}$, and appropriate random matrix is easy to find.

Summarizing, the results show that for small values of γ and ρ it takes not a lot of time to find the random matrix (3) without 4-cycles and with $b = m - 1$.

Estimation of system's parameters

In this section we estimate the parameters of the cryptosystem, basing on required security level. We will consider the following attacks, which complexity should be exponential of order not less than 2^{128}:

— search on private matrices;
— search within the set E';
— search within the set $\tilde{E}$.

Let us consider the search on private matrices. There are $m^{\gamma\rho}$ matrices (3) of $\gamma \times \rho$ blocks, where each block is defined by integer from the set $\{0, ..., m - 1\}$. For example we will take matrices with 3×6 and 4×8 blocks. From $m^{18} = 2^{128}$ and $m^{32} = 2^{128}$ we have the correspondent block sizes $m \approx 2^7 = 128$ and $m = 2^4 = 16$. To increase the probability of selecting the code with maximal correctable burst length we set m as prime number, i.e. $m = 127$ and $m = 17$. Particular selection of block size should be additionally agreed with the length b of correctable burst, which should be provided, so given estimations may be considered as lowest possible values for m.

Let us now consider brute-force search within the set E'. It consists of vectors $\mathbf{e}' = \tilde{\mathbf{e}}\mathbf{M}$, where $\mathbf{M} = \mathbf{M}_1\mathbf{M}_2$. Despite the special construction of ma-

trix $\mathbf{M}_1$ (see Figure), $\mathbf{M}_2$ is random matrix, thus even though $\tilde{\mathbf{e}}$ is error burst of length x, $\mathbf{e}'$ is random vector with expected weight $n/2$, which makes impossible both enumerating these vectors and breaking the system by decoding — any code including $\mathbf{G}'$ is not able to correct error vectors of such weight.

Finally, consider the set $\tilde{E}$. It consists of vectors which are bursts of length x. If starting position of the burst is fixed, there are about 2^x such vectors (in fact we should fill the positions within burst by random bits with probability 1/2 and obtain the bursts of weight near, so the number of such vectors is slightly less than 2^x). The number of burst locations (starting positions) within the error vector $\tilde{\mathbf{e}}$ is $n - x + 1$. This number is not very large, but we take it into account. Thus the complexity of brute force search of $\tilde{\mathbf{e}}$ is given by $(n - x + 1)2^x = 2^{128}$ (this equation is approximate, we do not take into account the weight of bursts, starting and ending 1's and so on).

Consider example with $\gamma = 3$, $\rho = 6$. Select $m = 127$, then the correctable burst length is $b = 126 = l + x - 1$, where l is the width of diagonal in matrix $\mathbf{M}_1$ in Figure. For such parameters we have $n = m\rho = 762$, and $(763 - x)2^x = 2^{128}$ and $x \approx 119$, hence $l = 8$.

For such parameters $\mathbf{M}_1$ is matrix 762×762, containing diagonal of width $l = 8$. Analysis of how the structure of $\mathbf{M}_1$ affects the system's strength is important question, but it does not considered in this paper since it requires more thorough and sophisticated analysis. Nevertheless, it seems that usage of rather large matrix with relatively thin diagonal of random elements while other elements are zero may be not secure. Thus consider the possibility of increasing the width of diagonal in $\mathbf{M}_1$.

Let us take, for example, $l = 30$. Then, having $x = 119$, the length of the burst correctable by the secret code should be $b = 149$ and hence the block size $m \geq 150$. We obtain the following parameters: $\gamma = 3$, $\rho = 6$, $l = 30$, $b = 149$, $m = 150$, $n = 900$, $x = 119$, then the number of bursts of length x is estimated as $2^{128.6}$, which corresponds to the required security level. To define the matrix $\mathbf{H}$ (3) it is enough to store only degrees of correspondent matrices $\mathbf{C}$, for our parameters we get $3 \cdot 6 \cdot \lceil \log_2 150 \rceil = 144$ bits. The size of $\mathbf{G}'$ (public key, which is more important from the point of view of key size), is 450×900 bits (which is less than initial parameters of McEliece system with key size 450×900 and cryptocomplexity 2^{53}). Note that we should also count the size of matrix $\mathbf{M}$ since it's the part of public key.

Similarly consider the case $\gamma = 4$, $\rho = 8$, with $m = 17$, which was defined earlier, and $b = 16$, $n = 136$. Then $x \approx 125$, and taking into account $b = l + x - 1$ it is impossible to have $b = 16$. Hence, the block size m should be significantly increased, as well as the burst length b.

Set $x = 125$ and $l = 30$, then we have $b = 155$ and $m \geq 156$, this gives parameters: $\gamma = 4$, $\rho = 8$, $l = 30$, $b = 155$, $m = 156$, $n = 1248$, $x = 125$, then the number of bursts of length x is 2^{135}. The size of private key is 256, public matrix $\mathbf{G}'$ — 624×1248 bit.

As can be seen from the estimations, the cryptocomplexity of proposed system depends on the number of bursts of length x, which defines the complexity of brute force search within the set $\tilde{E}$, and also on possibility of attacking the structure of $\mathbf{M}_1$, which depends on the value of l. These parameters define the values of b and m, such that enumerating the matrices $\mathbf{H}$ should be infeasible.

In Table 3 the key sizes, number of errors that should be decoded by adversary to break the system, cryptocomplexity of McEliece system, system based on quasi-cyclic LDPC codes (QC-LDPC) described in [8], and proposed system are collected.

Note that the public key size is defined by the size of $(k \times n)$-matrix $\mathbf{G}'$, however in proposed system the $(n \times n)$-matrix $\mathbf{M}$ is also the part of the public key, at the same time in QC-LDPC system significantly larger codes are used, but using the block-circulant structure of the matrix (including public matrix) the required storage for the keys may be significantly reduced. From the other hand, one should make the distinction between storage needs and memory which is used during encryption and especially decryption processes, when the decoding procedure should be

■ ***Table 3.*** Comparison of code-based cryptosystems

System	k	n	t	Complexity	Attack	Public key size, Kbyte	Private key size, Kbyte
McEliece	524	1024	50	2^{53}	Bounded-distance decoding (t errors)	67	67
	1036	2048	92	2^{94}		265	265
	2056	4096	170	2^{171}		1052	1052
QC-LDPC	9857	19 714	134	2^{128}		1.2	1.2
Bursts (proposed)	450	900	450	2^{128}	Complete decoding ($n/2$ errors) or brute force search on $\tilde{E}$	152	0.144
	624	1248	624	2^{135}		291	0.25

used. With increasing of matric sizes the required memory is also increases, but estimations depend on particular implementations and their optimizations.

At the same time, further development of the proposed system in the direction of using quasi-cyclic codes is of interest.

Conclusion

In the paper the code-based cryptosystem is proposed which uses burst-correction codes. The underlying hard mathematical problem is complete decoding problem. It is supposed that systems based on this problem can achieve better cryptocomplexity than code-based McEliece cryptosystem based on bounded-distance decoding.

The future investigations and development can be made in using quasi-cyclic codes and considering the version of the system in Niederreiter mode.

Financial support

The paper was prepared with the financial support of the Russian Science Foundation, project No. 22-19-00305 "Spatial-temporal stochastic models of wireless networks with a large number of users".

References

1. Stinson D. R., Paterson M. B. *Cryptography. Theory and Practice*. CRC Press, 2018. 598 p.
2. Cormen T. H., Leiserson C. E., Rivest R. L., Stein C. *Introduction to Algorithms*. MIT Press, 2022. 1312 p.
3. *NIST Post-quantum Cryptography Project*. Available at: https://csrc.nist.gov/projects/post-quantum-cryptography (accessed 21 April 2022).
4. Lin S. *Fundamentals of Classical and Modern Error-Correcting Codes*. Cambridge University Press, 2022. 800 p.
5. Moon T. K. *Error Correction Coding: Mathematical Methods and Algorithms*. Wiley, 2020. 992 p.
6. Chailloux A., Debris-Alazard T., Etinski S. Classical and quantum algorithms for generic syndrome decoding problems and applications to the lee metric. *Proc. 12th Intern. Conf. "Post-Quantum Cryptography", PQCrypto 2021*, J. H. Cheon, J.-P. Tillich (eds), LNCS, 2021, vol. 12841, pp. 44–62. doi: 10.1007/978-3-030-81293-5_3
7. Sendrier N. Code-based cryptography: State of the art and perspectives. *IEEE Security & Privacy*, 2017, vol. 15, no. 4, pp. 44–50. doi: 10.1109/MSP.2017.3151345
8. Baldi M. *QC-LDPC Code-Based Cryptography*. Springer, 2014. 120 p.
9. Canto Torres R., Sendrier N. Analysis of information set decoding for a sub-linear error weight. *Proc. 7th Intern. Conf. "Post-Quantum Cryptography", PQCrypto 2016*, T. Takagi (ed), LNCS, 2016, vol. 9606, pp. 144–161. doi:10.1007/978-3-319-29360-8_10
10. Both L., May A. Decoding linear codes with high error rate and its impact for LPN security. *Proc. 9th Intern. Conf. "Post-Quantum Cryptography", PQCrypto 2018*, T. Lange, R. Steinwandt (eds), LNCS, 2018, vol. 10786, pp. 25–46. doi:0.1007/978-3-319-79063-3_2
11. Kirshanova E. Improved quantum information set decoding. *Proc. 9th Intern. Conf. "Post-Quantum Cryptography", PQCrypto 2018*, T. Lange, R. Steinwandt (eds), LNCS, 2018, vol. 10786, pp. 507–527. doi:10.1007/978-3-319-79063-3_24
12. Lin S., Ryan W. *Channel Codes: Classical and Modern*. Cambridge University Press, 2009. 710 p.
13. Baldi M., Barenghi A., Chiaraluce F., Pelosi G., Santini P. LEDAkem: A post-quantum key encapsulation mechanism based on QC-LDPC codes. *Proc. 9th Intern. Conf. "Post-Quantum Cryptography", PQCrypto 2018*, T. Lange, R. Steinwandt (eds), LNCS, 2018, vol. 10786, pp. 3–24. doi:10.1007/978-3-319-79063-3_1
14. Krouk E. A new public-key cryptosystem. *Sixth Joint Swedish-Russian Intern. Workshop on Information Theory*, Moelle, Sweden, 1993, pp. 285–286.
15. Krouk E., Ovchinnikov A. Code-based public-key cryptosystem based on bursts-correcting codes. *The Thirteen Advanced Intern. Conf. on Telecommunications*, 2017, pp. 93–95.
16. Veresova A. M., Ovchinnikov A. A. About one algorithm for correcting bursts using block-permutation LDPC-codes. *2019 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, Saint-Petersburg, Russia, 2019, pp. 1–4.
17. Sendrier N., Vasseur V. About low DFR for QC-MDPC decoding. *Proc. 11th Intern. Conf. "Post-Quantum Cryptography", PQCrypto 2020*, J. Ding, J.-P. Tillich (eds), LNCS, 2020, vol. 12100, pp. 20–34. doi:10.1007/978-3-030-44223-1_2
18. Eaton E., Lequesne M., Parent A., Sendrier N. QC-MDPC: A timing attack and a CCA2 KEM. *Proc. 9th Intern. Conf. "Post-Quantum Cryptography", PQCrypto 2018*, T. Lange, R. Steinwandt (eds), LNCS, 2018, vol. 10786, pp. 47–76. doi:10.1007/978-3-319-79063-3_3
19. Krouk E. A., Ovchinnikov A. A. Exact burst-correction capability of gilbert codes. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 1, pp. 80–87 (In Russian). doi:10.15217/issn1684-8853.2016.1.80
20. Ovchinnikov A., Fominykh A. About burst decoding for block-permutation LDPC codes. *Internet of Things, Smart Spaces, and Next Generation Networks and Systems: 20-th Intern. Conf., NEW2AN 2020, and 13-th Conf., ruSMART 2020*, Saint-Petersburg, Russia, 2020, pp. 393–401. doi:10.1007/978-3-030-65726-0_35

УДК 003.26
doi:10.31799/1684-8853-2022-3-45-54

Вариант постквантовой системы на основе кодов, исправляющих пакеты ошибок, и задачи полного декодирования

А. А. Овчинников[a], канд. техн. наук, доцент, orcid.org/0000-0002-8523-9429, mldoc@guap.ru
[a]Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: важным направлением в исследовании постквантовых систем, устойчивых к квантовым вычислениям, является кодовая криптография на основе задач теории помехоустойчивого кодирования. Улучшение существующих кодовых систем может вестись как в практической части (уменьшение размеров ключей), так и с точки зрения использования более трудных математических кодовых задач. **Цель:** построение кодовой системы с открытым ключом на основе низкоплотностных кодов, исправляющих пакеты ошибок; оценка параметров полученной системы. **Результаты:** предложен вариант кодовой системы на основе случайных блочно-перестановочных низкоплотностных кодов. Стойкость системы предполагается основанной на задаче полного декодирования, что является более сложной математической задачей по сравнению с существующими системами. При этом с высокой вероятностью анализ системы на основе методов декодирования вообще не представляется возможным, что как повышает перспективную стойкость системы, так и позволяет уменьшить размеры ключей. Проведена оценка выбора кодов с требуемыми характеристиками, рассматриваются подходы к выбору параметров предложенной системы на основе требуемого уровня стойкости. **Практическая значимость:** предложенная система позволяет уменьшить размеры открытых ключей по сравнению с классической системой МакЭлиса при сравнимой стойкости, при этом используемая трудная математическая задача представляется более устойчивой к перспективным атакам.

Ключевые слова — постквантовая криптография, кодовые системы, коды с малой плотностью проверок на четность, исправление пакетов ошибок.

Для цитирования: Ovchinnikov A. A. The variant of post-quantum cryptosystem based on burst-correcting codes and on the complete decoding problem. *Информационно-управляющие системы*, 2022, № 3, с. 45–54. doi:10.31799/1684-8853-2022-3-45-54
For citation: Ovchinnikov A. A. The variant of post-quantum cryptosystem based on burst-correcting codes and on the complete decoding problem. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 45–54. doi:10.31799/1684-8853-2022-3-45-54

УДК 621.396.965
doi:10.31799/1684-8853-2022-3-55-66

Научные статьи
Articles

Достоверность селекции целей в сети разнесенных радиолокационных станций при совместной обработке радиолокационной информации в условиях ретранслированных помех

Д. В. Левин[а], *канд. техн. наук, orcid.org/ 0000-0002-3480-087X, dm.181@yandex.ru*
А. В. Паршуткин[а], *доктор техн. наук, профессор, orcid.org/ 0000-0001-7535-4880*
А. В. Тимошенко[б], *доктор техн. наук, профессор, orcid.org/ 0000-0002-9791-142X*
[а]*Военно-космическая академия им. А. Ф. Можайского, Ждановская наб., 13, Санкт-Петербург, 197198, РФ*
[б]*Национальный исследовательский университет «Московский институт электронной техники», Шокина пл., 1, Зеленоград, Москва, 124498, РФ*

Введение: *для выявления целей на коротких интервалах наблюдения используются методы обработки радиолокационной информации, основанные на отождествлении отметок разных радиолокационных станций и параметрической идентификации целей. Однако взаимное сравнение результативности применения таких методов в условиях ретранслированных помех не проводилось.* ***Цель:*** *сравнение достоверности селекции целей на фоне ретранслированных помех в сети территориально разнесенных радиолокационных станций обзора пространства при реализации метода корреляционного эллипсоида, метода строба селекции целей и метода пространственного разноса измеренных положений цели.* ***Результаты:*** *приведены решающие правила разбиения пространства разностей координат целей на подобласти принятия решения об истинности целей. Проведено имитационное моделирование селекции целей в условиях ретранслированных помех и получены зависимости изменения вероятности ошибочной селекции ложных отметок от нормированной дальности, отсчитываемой от середины базы разноса пары территориально разнесенных радиолокационных станций. Приведены количественные оценки достоверности селекции целей для различных условий радиолокационного наблюдения парой территориально разнесенных радиолокационных станций. Выигрыш в увеличении нормированной дальности при применении метода корреляционного эллипсоида по сравнению с методами строба и пространственного разноса измеренных положений цели составил от 36 до 46 %. Показано, что в большинстве практических ситуаций можно использовать наиболее простой в реализации метод пространственного разноса измеренных положений цели, а при решении наиболее важных задач — метод корреляционного эллипсоида.* ***Практическая значимость:*** *результаты исследований могут применяться при разработке алгоритмов селекции целей в условиях ретранслированных помех.*

Ключевые слова *— селекция целей, радиолокационная информация, ретранслированная помеха, ложная отметка, корреляционный эллипсоид.*

Для цитирования: Левин Д. В., Паршуткин А. В., Тимошенко А. В. Достоверность селекции целей в сети разнесенных радиолокационных станций при совместной обработке радиолокационной информации в условиях ретранслированных помех. *Информационно-управляющие системы*, 2022, № 3, с. 55–66. doi:10.31799/1684-8853-2022-3-55-66

For citation: Levin D. V., Parshutkin A. V., Timoshenko A. V. Reliability of target selection in the network of geographically separated radar stations in joint processing of radar information in the conditions of relayed interference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 55–66 (In Russian). doi:10.31799/1684-8853-2022-3-55-66

Введение

Современные наземные радиолокационные станции (РЛС) применяются для решения задач контроля пространства на обширных территориях. Для надежного обеспечения указанных задач создаются наземные многопозиционные комплексы, организованные на базе независимо функционирующих РЛС, связанных единой информационной сетью с каналами взаимного обмена данными [1, 2].

Под сетью разнесенных РЛС будем понимать множество независимо функционирующих территориально разнесенных разнодиапазонных РЛС, имеющих совокупность перекрывающихся зон обзора и общую систему обработки результатов радиолокационных измерений [3, 4].

В каждой отдельной РЛС прием сигналов, отраженных от целей, ведется на фоне различных излучений естественного и искусственного происхождения и собственных тепловых шумов приемного устройства и антенно-фидерного тракта. По результатам радиолокационного наблюдения в каждой РЛС формируются истинные и ложные отметки целей. При росте числа РЛС, объединяемых в сеть, повышается число ложных отметок, перегружающих систему обработки информации в сети РЛС. Поэтому при разработке современных наземных РЛС особое внимание уделяется помехоустойчивости к широкому классу преднамеренных и непреднамеренных помех. Однако в каждой отдельной РЛС защита от ретранслированных помех затрудняется сходством помехи с зондирующим сигналом [5–7]. В работах [4, 8]

было показано, что селекция целей на фоне ложных отметок осуществляется в основном по траекторным признакам. Известен метод первичного захвата физическими или математическими стробами с последующим сопровождением целей [9–12]. Кроме того, в РЛС обзора пространства реализуется метод идентификации измерений с существующими траекториями при сопровождении типовых целей [13–15]. Однако указанные методы селекции целей требуют длительного сопровождения как целей, так и ложных отметок.

На коротких интервалах радиолокационного наблюдения именно в сетях РЛС за счет многоракурсного наблюдения за одной и той же областью пространства появляется возможность моментальной селекции целей на фоне ложных отметок. Для этого могут использоваться методы селекции целей, основанные на отождествлении отметок разных РЛС и параметрической идентификации целей [2, 4, 8, 16–18]. При этом взаимное сравнение результативности применения указанных методов селекции целей в условиях ретранслированных помех не проводилось.

Целью статьи является сравнение достоверности селекции целей на фоне ретранслированных помех в сети территориально разнесенных РЛС обзора пространства при реализации следующих трех методов совместной обработки радиолокационной информации, различающихся порядком отнесения отметок к реально существующему радиолокационному объекту: корреляционного эллипсоида, строба селекции целей, пространственного разноса измеренных положений цели.

Постановка задачи селекции целей на фоне ложных отметок в сети территориально разнесенных РЛС

Рассмотрим ситуацию, когда несколько РЛС из состава сети территориально разнесенных РЛС осуществляют контроль заданной области пространства. При этом в общую систему обработки результатов радиолокационных измерений поступают отметки от радиолокационных объектов и ложные отметки, обусловленные помехами. При этом будем полагать, что ложные отметки равномерно распределены по наблюдаемой области пространства.

Независимое функционирование РЛС в составе сети приводит к тому, что единичные измерения параметров обнаруженных радиолокационных сигналов будут выполнены в отличающиеся моменты времени и в различных условиях радиоэлектронной обстановки. Соответственно, радиолокационная информация о заданной области пространства формируется разными РЛС в разные моменты времени и с разной в общем случае периодичностью [2, 8, 19–22]. Кроме того, измерения сопровождаются случайными и систематическими ошибками, а также погрешностями пересчета к единой системе координат и единому времени.

Порядок селекции целей на фоне ложных отметок в сети территориально разнесенных РЛС на основе сопоставления результатов измерений разных РЛС можно представить в следующем виде.

Пусть A — все множество отметок, сформированных в сети территориально разнесенных РЛС при наблюдении радиолокационных объектов, а также обусловленных помехами. Тогда в систему обработки радиолокационной информации поступает множество отметок a_i, $a_i \in A$ с некоторым набором признаков, измеренных разными РЛС.

Пусть совокупность всех РЛС в сети образует множество K. Без потери общности можно полагать, что все РЛС сети, наблюдающие за заданной областью пространства, можно разбить на пары $\{k_l, k_j\}$, где $k_l \in K$, $k_j \in K$. Это позволяет рассматривать всю сеть как некоторое конечное множество пар РЛС [8].

Множество признаков, которыми описываются отметки в конкретной РЛС k_j, $k_j \in K$, может включать параметры координатные (координаты и проекции вектора скорости движения цели) и некоординатные (эффективная площадь рассеяния, параметры поляризационной матрицы рассеяния отметки и др.). Рассмотрим только координатные признаки. Пусть в РЛС k_j, $k_j \in K$, объект наблюдения a_i, $a_i \in A$, описывается вектором из M признаков $\mathbf{X}_{ij} = \{x_{1i}, x_{2i}, ..., x_{Mi}\}$ с компонентами x_{mi}, которые представляют измеренное значение некоторой непрерывной величины, например координаты дальности до цели $m = 1, 2, ..., M$ [9–11].

Установим для сети разнесенных РЛС начало координат X_0 и базис $\alpha_1, \alpha_2, ..., \alpha_M$. В выбранном базисе M-мерного пространства отметка радиолокационного объекта a_i представляется точкой с координатами $x_{1i}, x_{2i}, ..., x_{Mi}$. Множеством отметок A отображается все множество наблюдаемых радиолокационных объектов и ложных отметок.

Для каждой пары РЛС можно ввести такую максимальную размерность координатных параметров, которая обеспечивается каждой станцией в паре. Поэтому в дальнейшем без потери общности изложения будем полагать, что разными РЛС в паре формируются сообщения о целях одинаковой размерности M.

Принятие решения об истинности радиолокационного объекта может основываться на том факте, что истинные значения координат отметок, принадлежащих одному и тому же объекту, при отсутствии ошибок измерений должны совпадать. В этом случае по всем $m \in M$ имеет ме-

сто $\Delta x_{mir} = x_{mi} - x_{mr} = 0$, где Δx_{mir} — разница между истинными координатами цели, представленной отметками a_i и a_r, $a_i \in A$, $a_r \in A$, наблюдаемой парой разнесенных РЛС [8, 12, 14].

Основным признаком, отличающим ложные отметки, образованные ретранслированными помехами, от радиолокационных объектов, может служить несовпадение их измеренных координат, полученное при многоракурсном наблюдении. Поэтому при несовпадении ракурсов наблюдения в паре РЛС $\{k_l, k_j\}$ координаты ложных отметок в разных РЛС совпадать не будут, $x_{mi} - x_{mr} \neq 0$ даже при отсутствии погрешностей измерений. Наличие таких погрешностей приводит к тому, что даже отметки от одного радиолокационного объекта, сформированные разными РЛС, после синхронизации и приведения к единой системе координат в общем случае совпадать не будут [8, 12, 14].

Задача селекции целей на фоне ложных отметок в сети разнесенных РЛС может быть сформулирована следующим образом: необходимо путем сравнения координатных параметров отметок, полученных от разных РЛС, принять решение, что это либо отметки от одного объекта и их рассогласование обусловлено погрешностями измерений, либо это ложные отметки, обусловленные помехами.

Далее будут рассмотрены указанные во введении методы селекции целей, достоверность которых будем сравнивать по вероятности ошибочной селекции ложной отметки.

Статистически оптимальный метод селекции целей на фоне ложных отметок

Для отыскания статистически оптимального метода селекции целей необходимо рассматривать результаты измерений координатных параметров радиолокационных объектов и ложных отметок как случайные величины. Для их описания используются многомерные законы распределения координатных параметров.

Пусть отметки a_i и a_r, $a_i \in A$, $a_r \in A$, наблюдаемые парой РЛС $\{k_j, k_l\}$, $k_j \in K$, $k_l \in K$, описываются M-мерными плотностями вероятности $\omega_{ij}(\mathbf{X})$ и $\omega_{rl}(\mathbf{X})$, где $\mathbf{X} = \{x_1, x_2, ..., x_M\}$. Введем $2M$-мерную плотность вероятности $\omega_{ir}(\mathbf{X}, \mathbf{X})$, описывающую совместное распределение пары отметок $\{a_i, a_r\}$. В случае статистической независимости результатов радиолокационных измерений разными РЛС в рассматриваемой паре $\omega_{ir}(\mathbf{X}, \mathbf{X}) = \omega_{ij}(\mathbf{X}) \cdot \omega_{rl}(\mathbf{X})$ [8, 12, 14].

Представим координатные признаки $\mathbf{X}_{ij}$ и $\mathbf{X}_{rl}$, измеряемые парой РЛС $\{k_j, k_l\}$, $k_j \in K$, $k_l \in K$, суммой детерминированной составляющей, определяемой координатами цели $\mathbf{X}_{\text{ц}}$, и некоторой погрешности измерений в РЛС δ_j и δ_l соответственно [2, 8, 23, 24]:

$$\mathbf{X}_{ij} = \mathbf{X}_{\text{ц}} + \delta_j; \; \mathbf{X}_{rl} = \mathbf{X}_{\text{ц}} + \delta_l. \qquad (1)$$

В дальнейшем будем полагать, что систематические ошибки измерений в каждой РЛС значительно меньше случайных ошибок.

Часто предполагается, что в пределах рассматриваемой области пространства плотности распределения координатных параметров отметок $\omega_{ij}(\delta_j)$ и $\omega_{rl}(\delta_l)$ в первом приближении не зависят от $\mathbf{X}_{\text{ц}}$. В этом случае для снижения размерности исследуемых плотностей вероятности целесообразно проводить анализ разности их измеренных координат $\Delta\mathbf{X}_{ir} = \mathbf{X}_{ij} - \mathbf{X}_{rl} = \{\Delta x_{mir}, m = 1, 2, ..., M\}$. Причем в силу несовпадения ракурсов наблюдения РЛС разброс случайной величины $\Delta\mathbf{X}_{ir}$ для ложных отметок будет превосходить разброс отметок истинного радиолокационного объекта.

При статистической постановке задача селекции целей сводится к различению двух гипотез [25, 26]:

H_1 — совместное распределение пары отметок $\{a_i, a_r\}$ соответствует наблюдению одного радиолокационного объекта;

H_2 — совместное распределение пары отметок $\{a_i, a_r\}$ соответствует наблюдению ложных отметок.

Тогда полный класс оптимальных решающих правил селекции целей можно представить сравнением отношения правдоподобия с порогом в следующем виде:

— гипотеза H_1 принимается при условии $\omega_{ir}(\Delta\mathbf{X}_{ir}/H_1)/\omega_{ir}(\Delta\mathbf{X}_{ir}/H_2) \geq \gamma$;

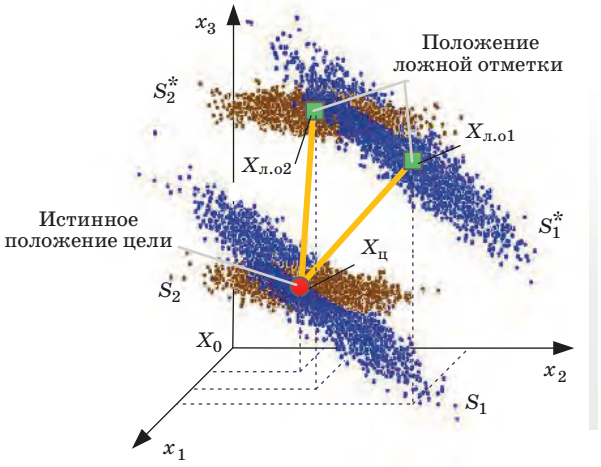
— гипотеза H_2 принимается при условии $\omega_{ir}(\Delta\mathbf{X}_{ir}/H_1)/\omega_{ir}(\Delta\mathbf{X}_{ir}/H_2) < \gamma$,

где $\omega_{ir}(\Delta\mathbf{X}_{ir}/H_1)$, $\omega_{ir}(\Delta\mathbf{X}_{ir}/H_2)$ — условные плотности распределения разности координат пар отметок при реализации гипотез H_1 и H_2 соответственно; γ — порог принятия решения, определяемый выбранным критерием оптимальности.

Метод корреляционного эллипсоида

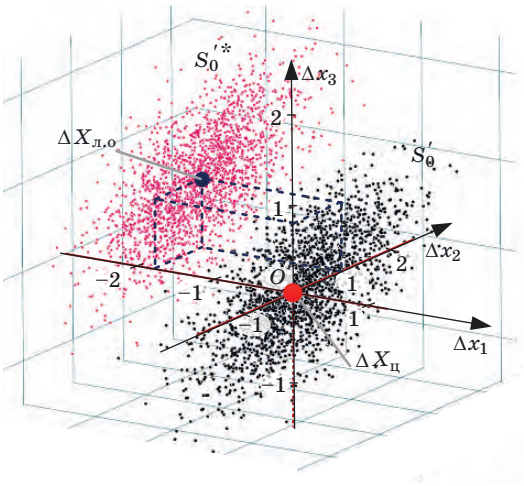
Как правило, погрешности радиолокационных измерений в топоцентрической системе координат РЛС можно полагать независимыми и нормально распределенными, но не равноточными. Примеры входных реализаций, приведенных в общий базис для пары РЛС, показаны на рис. 1, где S_j, S_j^* — множества измеренных координат цели (отметок цели) и имитируемой цели (ложных отметок), обусловленных ретранслированными помехами наблюдаемых РЛС k_j, $j = 1, 2$.

При проведении многократных статистических испытаний по измерению координат цели



■ ***Рис. 1.*** Множества измеренных координат цели и ложной отметки, обусловленной ретранслированными помехами, в трехмерной системе координат

■ ***Fig. 1.*** A set of measured coordinates target and false mark caused by relayed interference in three-dimension coordinate system



■ ***Рис. 2.*** Множество разностей координат для отметок цели и пары ложных отметок

■ ***Fig. 2.*** A set of coordinate differences for target marks and pair of false marks

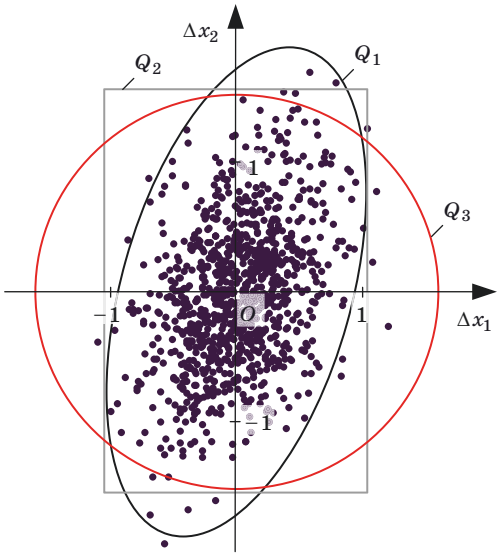
$X_ц$ множества отметок цели S_j и ложных отметок S_j^* будут сформированы преимущественно в пределах эллипсоидов рассеяния. Причем форма и пространственная ориентация эллипсоидов рассеяния определяются точностными характеристиками и ракурсом наблюдения цели РЛС k_j [8, 24, 27–29].

Очевидно, что погрешности измерений, независимые в топоцентрической системе координат РЛС, при развороте базиса получаются коррелированными. Отсутствие провала в центре плотности распределения $\omega_{ir}(\Delta\mathbf{X}_{ir}/H_2)$ обусловлено тем, что ложные отметки, образованные ретранслированными помехами, могут приближаться друг к другу сколь угодно близко, в отличие от реальных радиолокационных объектов [4, 5, 10, 15, 22].

Плотности распределения разности координат $\Delta\mathbf{X}_{ir}$ образуются сверткой нормальных законов распределения и также имеют M-мерный нормальный закон распределения. Пример множества разностей координат для отметок цели и пары ложных отметок представлен на рис. 2.

На рисунке введены следующие обозначения: $\Delta X_ц$ — истинное положение цели; $\Delta X_{л.о}$ — положение ложной отметки, обусловленной ретранслированной помехой; S_0 — множество отметок, характеризующих разность измеренных координат цели территориально разнесенных РЛС; S_0^* — множество отметок, характеризующих разность измеренных координат пары ложных отметок, обусловленных ретранслированной помехой. Данный пример показывает, что дисперсия разности координат $\Delta\mathbf{X}_{ir}$ в случае наблюдения фиксированной пары ложных отметок совпадает с дисперсией разности координат цели. Распределения разностей координат целей и ложных отметок отличаются только математическими ожиданиями. Для цели математическое ожидание разности координат территориально разнесенных РЛС стремится к нулю, а для любой пары ложных отметок будет представлять случайную величину. Это позволяет полагать, что в первом приближении в области допустимых погрешностей разности радиолокационных измерений условная плотность распределения $\omega_{ir}(\Delta\mathbf{X}_{ir}/H_2)$ может рассматриваться как постоянная величина. В этом случае оптимальное решающее правило заключается в сравнении условной плотности распределения $\omega_{ir}(\Delta\mathbf{X}_{ir}/H_1)$ с величиной порога γ^*. Следовательно, при селекции целей в условиях ретранслированных помех оптимальное решающее правило заключается в нахождении такой подобласти M-мерного пространства, для которой $\omega_{ir}(\Delta\mathbf{X}_{ir}/H_1) \geq \gamma^*$ и форма которой определяется только плотностью распределения разности координат для истинной цели. С учетом сделанных выше предположений оптимальное решающее правило можно представить как проверку попадания разности координат $\Delta\mathbf{X}_{ir}$ в корреляционный эллипсоид [1, 2, 8, 12].

Пример множества разностей координат цели в плоскости $O\Delta x_1\Delta x_2$ приведен на рис. 3. Для двумерного пространства наблюдения РЛС эллипсоид рассеяния вырождается в эллипс. На рисунке показаны примеры разбиения пространства на подобласти, соответствующие разным методам селекции целей, где Q_1 — корреляционный эллипс, Q_2 — прямоугольный строб в выбранном базисе двумерного пространства, Q_3 — область



■ ***Рис. 3.*** Пример разбиения пространства на подобласти принятия решения об истинности наблюдаемой цели

■ ***Fig. 3.*** An example of dividing space into subdomains of deciding on truth of observed target

допустимых значений пространственного разноса измеренных положений цели.

Главные оси эллипса рассеяния в плоскости $O\Delta x_1\Delta x_2$ прямоугольной системы координат в общем случае не совпадают с осями прямоугольной системы координат. Их размеры и ориентация определяются среднеквадратическими отклонениями (СКО) измерения прямоугольных координат $\sigma_{\Delta x1}$ и $\sigma_{\Delta x2}$ и величиной коэффициента взаимной корреляции $r_{\Delta x1\Delta x2}$ [1, 2, 8, 12].

Если центр эллипса на плоскости $O\Delta x_1\Delta x_2$ находится в точке (0; 0), то его оси симметрии будут повернуты на угол α, определяемый из выражения

$$\operatorname{tg}(2\alpha) = \frac{2R_{\Delta x1\Delta x2}\sigma_{\Delta x1}\sigma_{\Delta x2}}{\sigma^2_{\Delta x1} - \sigma^2_{\Delta x2}}, \qquad (2)$$

где $R_{\Delta x1\Delta x2}$ — корреляционный момент, связанный коэффициентом корреляции Пирсона $r_{\Delta x1\Delta x2}$ соотношением $r_{\Delta x1\Delta x2} = R_{\Delta x1\Delta x2}/(\sigma_{\Delta x1} \cdot \sigma_{\Delta x2})$; $\sigma^2_{\Delta x1}$, $\sigma^2_{\Delta x2}$ — дисперсии ошибок измерения координат Δx_1 и Δx_2.

Коэффициент корреляции Пирсона r_{xy} между множествами нормально распределенных измерений координат по осям $O\Delta x_1$ и $O\Delta x_2$ определяется по формуле

$$r_{\Delta x1\Delta x2} = \frac{\sum(\Delta x_1 - M[\Delta x_1])(\Delta x_2 - M[\Delta x_2])}{\sqrt{\sum(\Delta x_1 - M[\Delta x_1])^2(\Delta x_2 - M[\Delta x_2])^2}}, \qquad (3)$$

где $M[\Delta x_1]$, $M[\Delta x_2]$ — математические ожидания нормально распределенных измерений координат по осям $O\Delta x_1$ и $O\Delta x_2$ соответственно; Δx_1, Δx_2 — значения измеренных координат.

Тогда, учитывая каноническое уравнение эллипса в системе координат $O\Delta x_1\Delta x_2$, условие попадания точки в эллипс рассеяния задается выражением

$$\frac{(\Delta x_1\cos(\alpha) + \Delta x_2\sin(\alpha))^2}{a^2} + \frac{(-\Delta x_1\sin(\alpha) + \Delta x_2\cos(\alpha))^2}{b^2} \le 1, \qquad (4)$$

где a, b — большая и малая полуоси эллипса рассеяния на $O\Delta x_1\Delta x_2$, которые определяются по следующим формулам:

$$a = \sqrt{\sigma^2_{\Delta x1} + \sigma^2_{\Delta x2} + \sqrt{\left(\sigma^2_{\Delta x1} - \sigma^2_{\Delta x2}\right) + 4K^2_{\Delta x1\Delta x2}\sigma^2_{\Delta x1}\sigma^2_{\Delta x2}}};$$
$$b = \sqrt{\sigma^2_{\Delta x1} + \sigma^2_{\Delta x2} - \sqrt{\left(\sigma^2_{\Delta x1} - \sigma^2_{\Delta x2}\right) + 4K^2_{\Delta x1\Delta x2}\sigma^2_{\Delta x1}\sigma^2_{\Delta x2}}}. \qquad (5)$$

Коэффициент c представляет собой отношение полуосей эллипса к соответствующим СКО, т. е. $c = a/\sigma_{\Delta x1} = b/\sigma_{\Delta x2}$, причем вероятность попадания в корреляционный эллипс p заданного размера c случайного измерения определяется из выражения

$$p = 1 - \exp(-c^2/2), \qquad (6)$$

откуда размер эллипса c, требуемый для обеспечения заданной вероятности попадания, задается по формуле

$$c = \sqrt{-2\ln(1-p)}. \qquad (7)$$

Сравнение разбиения плоскости на подобласти принятия решения об истинности наблюдаемой цели для трех рассматриваемых методов селекции целей на фоне ложных отметок представлено на рис. 4. При этом размеры эллипса, прямоугольника и круга выбраны таким образом, чтобы вероятности пропуска истинной цели были для всех трех фигур одинаковыми. В этом случае вероятности ошибочного принятия ложной отметки за цель $P_{\text{ош.л.о}}$ определяются соотношением площадей соответствующих фигур, которые для рассмотренного примера соотносятся как 1:1,35:1,38 соответственно. Тогда методы Q_2 и Q_3 по показателю $P_{\text{ош.л.о}}$ несколько уступают методу Q_1, и можно констатировать, что метод пространственного разноса измеренных положений цели Q_3 при равноточных некоррелированных измере-

ниях может давать значения, близкие к потенциально достижимым, поскольку корреляционный эллипсоид вырождается в сферу. В большинстве же случаев неравенства погрешностей измерений по дальности и угловым координатам (особенно при наблюдении целей на предельной дальности) метод стробов может обеспечивать меньшее число ошибок селекции целей при удачном выборе базиса, если границы строба селекции цели параллельны полуосям эллипсоида рассеяния величины $\Delta\mathbf{X}_{ir}$.

Методы строба селекции целей и пространственного разноса измеренных положений цели

При селекции целей методом строба учитывается тот факт, что отметки одной и той же цели должны иметь одинаковые значения всех компонентов. Решение о селекции целей принимается путем сравнения всех компонентов отметок независимо. Если при этом совпадают все соответствующие компоненты, то принимается решение об истинности цели, наблюдаемой разнесенными РЛС [1–3, 8–10].

На основе ошибок измерений, пересчитанных в базис с координатами $\mathbf{X}_{ir}$ в M-мерном пространстве, задаются размеры допустимых ошибок по каждой координате $\Delta x_{m\,\text{доп}}$.

В работе [8] показано, что для селекции целей на фоне ложных отметок целесообразно использовать величину пространственного разноса отметок от разных РЛС $\Delta r_{\text{р.и.п.ц}}$. Под величиной $\Delta r_{\text{р.и.п.ц}}$ понимается декартово расстояние между измеренными в единой системе координат отметками цели от разнесенных РЛС:

$$\Delta r_{\text{р.и.п.ц}} = \sqrt{\sum_{i=1}^{3}(x_{i1} - x_{i2})^2}. \qquad (8)$$

Поскольку ошибки измерения сферических координат целей в каждой РЛС распределены по многомерному нормальному закону, то плотность вероятности $\Delta r_{\text{р.и.п.ц}}$ описывается законом распределения Рэлея [4, 8]. Дисперсия распределения случайной величины $\Delta r_{\text{р.и.п.ц}}$ может быть оценена выражением [8]

$$\sigma^2_{\Delta r} = \sigma^2_{R1} + \sigma^2_{R2} + R_1[\text{tg}^2(\sigma_{\alpha 1}) + \text{tg}^2(\sigma_{\beta 1})] + \\ + R_2[\text{tg}^2(\sigma_{\alpha 2}) + \text{tg}^2(\sigma_{\beta 2})], \qquad (9)$$

где индексы 1 и 2 характеризуют СКО измерения дальности и угловых координат первой и второй РЛС соответственно; R_1 и R_2 — измеренные расстояния от $\text{РЛС}_{1,2}$ до цели.

Правилом принятия решения об истинности цели, одновременно наблюдаемой территориально разнесенными РЛС в условиях ретранслированных помех, является непревышение текущей величины $\Delta r_{\text{р.и.п.ц}\,ir}$ для пары отметок разнесенных РЛС с пороговым значением $\mu_{\text{доп}\,ir}$. Причем значение $\mu_{\text{доп}\,ir}$ определяется через вероятность правильной селекции целей $P_{\text{сел}}$ для закона распределения Рэлея с соответствующей $\sigma_{\Delta r\,ir}$ [8].

Результаты имитационного моделирования селекции целей на фоне ложных отметок, обусловленных ретранслированными помехами

Для оценивания достоверности различных методов селекции целей проведено имитационное моделирование процессов селекции целей на фоне ложных отметок, обусловленных ретранслированными помехами, при реализации в сети разнесенных РЛС метода корреляционного эллипсоида, метода строба селекции целей и метода пространственного разноса измеренных положений цели. При моделировании принят ряд допущений:

— одна физически существующая цель находится в области контроля двух разнесенных РЛС, каждая из которых измеряет дальность и угловые координаты целей, пересчитываемые в базис x, y, z;

— случайные ошибки единичных измерений сферических координат целей (δ_R, δ_α, δ_β) в каждой РЛС имеют нормальный закон распределения с нулевым математическим ожиданием и СКО измерений сферических координат (σ_R, σ_α, σ_β), функционально задаваемыми через положение цели относительно антенной системы РЛС и величины отношения сигнал/(шум+помеха) $q_{\text{с/ш+п}}$ в точке приема отраженного сигнала;

— расстояние до цели характеризуется нормированной дальностью $R_\text{ц}$, которая отсчитывается от середины базы разноса двух разнесенных РЛС и задается в относительных величинах по формуле

$$R_\text{ц} = R_\text{тек}/R^*_\text{max}, \qquad (10)$$

где $R_\text{тек}$ — текущее расстояние до физически существующей цели; R^*_max — расстояние от середины базы разноса двух территориально разнесенных РЛС до крайней точки зоны ответственности двух РЛС;

— относительное среднее удаление пары ложных отметок $\Delta r_\text{л.о}$, обусловленных ретранслированными помехами, задается по формуле

$$\Delta r_\text{л.о} = \Delta r_{\text{л.о ср}}/\Delta r_{\text{РЛС min}}, \qquad (11)$$

где $\Delta r_{\text{л.о ср}}$ — среднее на интервале радиолокационного наблюдения расстояние между ложными отметками, формируемыми в РЛС_1 и РЛС_2 из-за ретранслированных помех, определяется плотностью ложных отметок в контролируемой области пространства; $\Delta r_{\text{РЛС min}}$ — наименьшее значение разрешающей способности по дальности двух территориально разнесенных РЛС, объединенных в единую информационную сеть; в процессе каждого цикла статистических испытаний положение цели в контролируемой области пространства фиксировано, а погрешности измерений и положение ложной отметки задаются случайно. При этом погрешности измерений формировались по нормальному закону распределения, а положение ложных отметок — по равномерному с фиксированным средним расстоянием между ними.

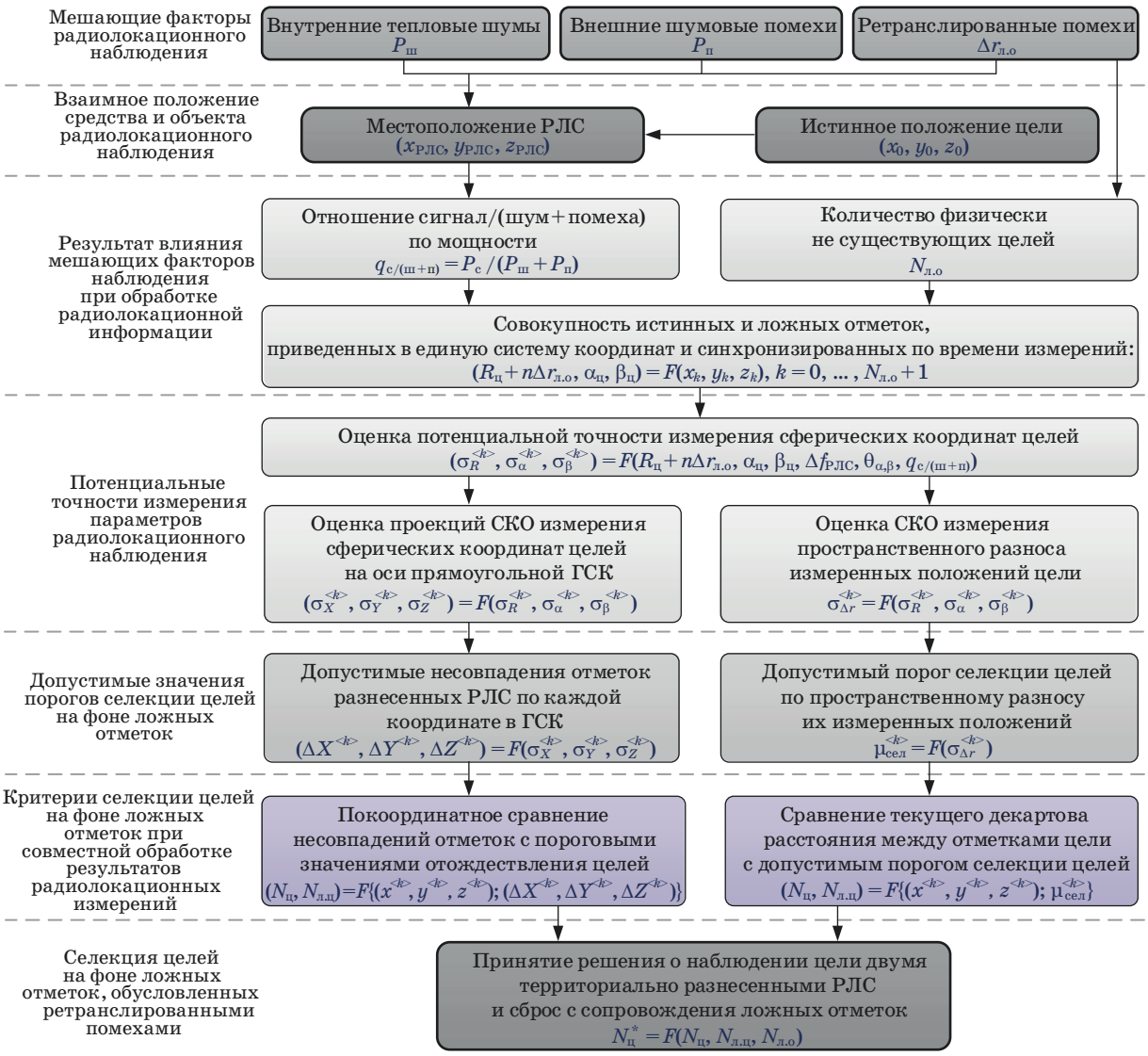
При известных положениях каждой независимо функционирующей РЛС ($x_{\text{РЛС}}$, $y_{\text{РЛС}}$, $z_{\text{РЛС}}$) в геоцентрической системе координат (ГСК) и положении физически существующей цели (x_0, y_0, z_0), наблюдаемой в условиях ретранслированных помех, измеренные сферические координаты ($R_{\text{ц}}$, $\alpha_{\text{ц}}$, $\beta_{\text{ц}}$) функционально задаются через топоцентрические прямоугольные координаты (x_k, y_k, z_k), где k — номер отметки цели и ложных отметок в РЛС, причем $k = 0, \ldots, N_{\text{л.о}} + 1$, где $N_{\text{л.о}}$ — количество ложных отметок, обусловленных ретранслированными помехами (т. е. количество физически не существующих целей, которые независимо функционирующие РЛС наблюдают в условиях ретранслированных помех). Причем топоцентрические прямоугольные координаты (x_k, y_k, z_k) в очередном цикле обзора формируются как результат измерения местоположения цели с учетом внутренних тепловых шумов и внешних шумовых помех на выходе системы обработки радиолокационной информации РЛС $q_{\text{с/ш+п}}$ [19–22]. Величина $q_{\text{с/ш+п}}$ задается отношением мощности принятого радиолокационного сигнала $P_{\text{с}}$, отраженного от цели, к суммарной мощности внутренних тепловых шумов приемного устройства РЛС и внешних шумовых помех ($P_{\text{ш}} + P_{\text{п}}$), отражающих помеховую обстановку в районе расположения наземной РЛС [8–11].

При известном отношении сигнал/шум $q_{\text{с/ш+п}}$ на выходе системы обработки радиолокационной информации в сети разнесенных РЛС могут быть выполнены оценки характеристик эллипсоидов рассеяния для ограниченных областей пространства. Связано это с тем, что отметка цели будет сформирована с учетом разрешающей способности РЛС по дальности $\delta r_{\text{РЛС}}$ и по угловым координатам ($\delta\alpha_{\text{РЛС}}$, $\delta\beta_{\text{РЛС}}$), а также с потенциальной точностью выполняемых радиолокационных измерений (σ_R, σ_α, σ_β), где σ_R — СКО измерений в канале дальности; σ_α, σ_β — СКО измерений в угловых каналах по азимуту и углу места соответственно. Разрешающая способность РЛС по дальности $\delta r_{\text{РЛС}}$ прямо пропорциональна ширине спектра зондирующего сигнала $\Delta f_{\text{РЛС}}$ [10, 13, 20]. Разрешающая способность РЛС по угловым координатам ($\delta\alpha_{\text{РЛС}}$, $\delta\beta_{\text{РЛС}}$) обратно пропорциональна ширине приемной диаграммы направленности по уровню половинной мощности в азимутальной θ_α и угломестной θ_β плоскостях [10, 13, 20]. Кроме того, потенциальная точность измерения сферических координат (σ_R, σ_α, σ_β) определяется величиной $q_{\text{с/ш+п}}$.

В обобщенном виде схема имитационного моделирования селекции целей на фоне ложных отметок, обусловленных ретранслированными помехами, в сети двух разнесенных РЛС обзора пространства представлена на рис. 4.

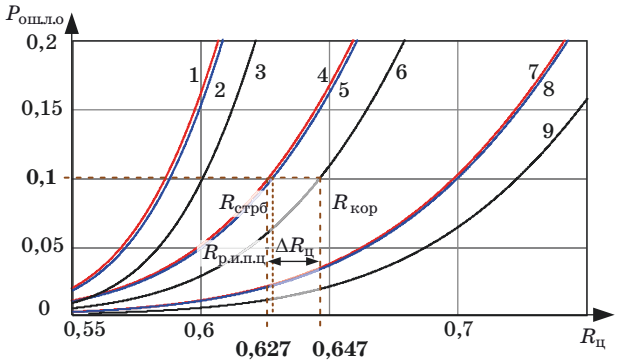
В результате проведения многократных статистических испытаний получены зависимости изменения вероятности сброса с сопровождения ложных траекторий цели, обусловленных ретранслированными помехами, в сети разнесенных РЛС для различных условий радиолокационного наблюдения. На рис. 5 приведены графики зависимостей изменения вероятности сброса с сопровождения ложных траекторий целей от расстояния до цели $R_{\text{ц}}$ при различных величинах относительного среднего удаления пары ложных отметок $\Delta r_{\text{л.о}}$ и пролете цели по маршруту, проекция которого на поверхности земли проходит перпендикулярно базе разноса между двумя территориально разнесенными РЛС. При моделировании величина $\Delta r_{\text{л.о1}}$ задавалась равной трем (кривые *1*, *2*, *3*), $\Delta r_{\text{л.о2}} = 3{,}5$ (кривые *4*, *5*, *6*), $\Delta r_{\text{л.о3}} = 5$ (кривые *7*, *8*, *9*). Кривые получены при реализации метода строба селекции целей $R_{\text{стрб}}$ (кривые *1*, *4*, *7*), метода пространственного разноса измеренных положений цели $R_{\text{р.и.п.ц}}$ (кривые *2*, *5*, *8*) и метода корреляционного эллипсоида $R_{\text{кор}}$ (кривые *3*, *6*, *9*).

Представленные на рис. 5 графики зависимостей позволяют сделать вывод, что сброс до 90 % ложных траекторий целей, формируемых из-за ретранслированных помех при разных значениях относительного удаления ложной отметки $\Delta r_{\text{л.о}}$, осуществляется для нормированной дальности $R_{\text{стрб}} = 0{,}587 \div 0{,}698$, $R_{\text{р.и.п.ц}} = 0{,}589 \div 0{,}700$ и $R_{\text{кор}} = 0{,}603 \div 0{,}722$. Результаты имитационного моделирования наблюдения целей территориально разнесенными РЛС при изменении относительного угла пролета от –15 до 15 градусов в диапазоне заданных значений $\Delta r_{\text{л.о}} = 3 \div 5$ показали, что изменение нормированной дальности для метода строба селекции целей и метода пространственного разноса измеренных положений цели осуществляется в интервалах $R_{\text{стрб (р.и.п.ц)}} = 0{,}560 \ldots 0{,}705$, а для метода корреляционного эллипсоида $R_{\text{кор}} = 0{,}590 \ldots 0{,}730$.



■ ***Рис. 4.*** Схема моделирования селекции целей в условиях ретранслированных помех

■ ***Fig. 4.*** Simulation scheme of target selection in conditions relayed interference



■ ***Рис. 5.*** Зависимости изменения вероятности ошибочной селекции ложной отметки

■ ***Fig. 5.*** Dependence of change in probability of erroneous selection of false mark

Анализ результатов моделирования показывает, что на вероятность селекции целей существенное влияние оказывает плотность ложных отметок. Так, увеличение относительного удаления ложной отметки в $\Delta r_{л.о}$ в 1,2 и 1,7 раза при реализации указанных методов селекции целей увеличивает нормированную дальность сброса до 90 % ложных траекторий целей, обусловленных ретранслированными помехами, на 7 и 19 % для указанных методов селекции целей.

Выигрыш в увеличении нормированной дальности $\Delta R_ц$ при применении метода корреляционного эллипсоида по сравнению с методами строба и пространственного разноса измеренных положений цели при фиксированных условиях наблюдения составляет от 36 до 46 %.

Заключение

Совместная обработка радиолокационной информации в сети разнесенных РЛС может быть использована для повышения помехоустойчивости в условиях ретранслированных помех. Сравнение методов селекции целей на фоне ложных отметок демонстрирует, что наибольшая достоверность обеспечивается при реализации метода корреляционных эллипсоидов. Методы стробов селекции целей и пространственного разноса измеренных положений цели дают близкие результаты с точки зрения обеспечения вероятности ошибочной селекции ложных отметок.

Учет пространственного согласования результатов радиолокационных измерений, выполняемых разнесенными РЛС, позволяет выявлять ложные траектории целей, обусловленные наличием ретранслированных помех. На основе имитационного моделирования показано, что в двухпозиционном радиолокационном комплексе реализация метода корреляционного эллипсоида по сравнению с методом строба селекции целей по-разному сказывается на качестве его функционирования в условиях ретранслированных помех. Основной причиной отличия в результативности сброса ложных траекторий целей является порядок формирования для двух разнесенных РЛС единого строба сопровождения целей. Выигрыш в увеличении нормированной дальности при применении метода корреляционного эллипсоида по сравнению с методами строба и пространственного разноса измеренных положений цели составил от 36 до 46 %.

Небольшие различия в реальных дальностях селекции целей, реализованные тремя рассмотренными методами, показывают, что в большинстве практических ситуаций можно использовать наиболее просто реализуемый из них метод пространственного разноса измеренных положений цели. При решении особо важных задач радиолокационного наблюдения целесообразно применять метод корреляционного эллипсоида, который обеспечивает максимальную достоверность селекции целей на фоне ложных отметок.

Литература

1. **Татарский Б. Г., Ильчук П. А., Ильчук А. Р.** Идентификация параметров сигналов, отраженных от множества целей, в распределенной радиолокационной системе методом сравнения. *Информационно-измерительные и управляющие системы*, 2018, т. 16, № 1, с. 33–41.
2. **Хомяков А. В., Филипченков В. И., Мамон Ю. И.** Алгоритмы совместной траекторной обработки в многопозиционном радиолокационном комплексе. *Изв. Тульского государственного университета. Технические науки*, 2016, № 2, с. 305–314.
3. **Shi Y., Park S., Song T.** Multitarget tracking in cluttered environment for a multistatic passive radar system under the DAB/DVB network. *EURASIP Journal on Advances in Signal Processing*, 2017, no. 11. doi:10.1186/s13634-017-0445-4
4. **Паршуткин А. В., Левин Д. В., Галандзовский А. В.** Имитационная модель обработки радиолокационной информации в сети радиолокационных станций в условиях сигналоподобных помех. *Информационно-управляющие системы*, 2019, № 6, с. 22–31. doi:10.31799/1684-8853-2019-6-22-31
5. **Агиевич С. Н., Луценко С. А.** Применение ретранслированных помех в целях воздействия на спутниковые системы радиосвязи с фазоманипулированными широкополосными сигналами. *Изв. Тульского государственного университета. Технические науки*, 2018, № 12, с. 411–416.
6. **Chen X., Chen B.** Anti-jamming approach based on radar transmitted waveform matching. *EURASIP Journal on Advances in Signal Processing*, 2021, no. 59. doi:10.1186/s13634-021-00776-w
7. **Куликов Г. В., Лелюх А. А., Граченко Е. Н.** Помехоустойчивость когерентного приемника сигналов с квадратурной амплитудной модуляцией при наличии ретранслированной помехи. *Радиотехника и электроника*, 2020, т. 65, № 8, с. 804–808. doi:10.31857/S0033849420070074
8. **Левин Д. В.** Моделирование селекции целей территориально-разнесенными радиолокационными станциями при совместной обработке их радиолокационных измерений в условиях ретранслированных помех. *Вопросы радиоэлектроники*, 2020, № 11, с. 6–13. doi:10.21778/2218-5453-2020-11-6-13
9. **Shi C., Zhou J., Wang F.** Adaptive resource management algorithm for target tracking in radar network based on low probability of intercept. *Multidimensional Systems and Signal Processing*, 2018, no. 29, pp. 1203–1226. doi:10.1007/s11045-017-0494-8
10. **Zhu Y., Zhao L., Zhang Y.** Receiver selection for multi-target tracking in multi-static Doppler radar systems. *EURASIP Journal on Advances in Signal Processing*, 2021, no. 118. doi: 10.1186/s13634-021-00826-3
11. **Смирнов Е. Е., Поздняков А. А.** Методика сопровождения близко расположенных объектов с различными характеристиками движения радиолокационной станцией дальнего обнаружения. *Электромагнитные волны и электронные системы*, 2021, т. 26, № 4, с. 42–53. doi:10.18127/j15604128-202104-05
12. **Васильев К. К., Маттис А. В., Саверкин О. В.** Стробирование радиолокационных отметок при траек-

торной фильтрации в связанных координатах. *Изв. высших учебных заведений России. Радиоэлектроника*, 2019, т. 22, № 5, с. 71–79. doi:10.32603/1993-8985-2019-22-5-71-79

13. **Алёшкин А. П., Владимиров В. В., Невзоров В. И., Савочкин П. В.** Метод повышения разрешающей способности и точности радиолокационных угловых измерений на основе последовательной пространственно-временной обработки принимаемых сигналов. *Информационно-управляющие системы*, 2020, № 2, с. 37–45. doi:10.31799/1684-8853-2020-2-37-45
14. **Смирнов Е. Е., Поздняков А. А., Паршин М. С.** Модель классификации объектов наблюдения в условиях пересечения их траекторий движения на основе совместного анализа траекторной и поляризационной информации. *Информационно-измерительные и управляющие системы*, 2021, т. 19, № 4, с. 14–26. doi:10.18127/j20700814-202104-02
15. **Zarai K., Cherif A.** Adaptive filter based on Monte Carlo method to improve the non-linear target tracking in the radar system. *Aerospace Systems*, 2021, no. 4 (7), pp. 67–74. doi:10.1007/s42401-020-00080-9
16. **Пальгуев Д. А., Шентябин А. Н.** К вопросу оценки вероятности объединения радиолокационной информации при третичной обработке в сетевых структурах. *Радиопромышленность*, 2020, т. 30, № 2. с. 32–41. doi:10.21778/2413-9599-2020-30-2-32-41
17. **Кирюшкин В. В., Коровин А. В., Журавлев А. В.** Межпозиционное отождествление результатов измерений и определение координат воздушных целей в многопозиционной радиолокационной системе в условиях многоцелевой обстановки. *Журнал Сибирского федерального университета. Серия: Техника и технологии*, 2019, т. 12, № 6, с. 708–718. doi:10.17516/1999-494X-0170
18. **Журавлев А. В., Кирюшкин В. В., Коровин А. В.** Алгоритм межпозиционного отождествления результатов измерений в суммарно-дальномерной многопозиционной радиолокационной системе в условиях многоцелевой обстановки. *Радиотехника*, 2019, т. 83, № 6 (8), с. 180–189. doi:10.18127/j00338486-201906(8)-16
19. **Бучинский Д. И., Вознюк В. В., Фомин А. В.** Исследование помехоустойчивости приемника сигналов с многопозиционной фазовой манипуляцией к воздействию помех с различной структурой. *Тр. Военно-космической академии им. А. Ф. Можайского*, 2019, № 671, с. 120–127.
20. **Алёшкин А. П., Балашов В. М., Владимиров В. В.** Синтезирование искусственной апертуры в цифровой антенной решетке путем экстраполяции функции раскрыва на основе последовательной обработки отраженных радиолокационных сигналов. *Вопросы радиоэлектроники*, 2021, № 4, с. 16–22. doi:10.21778/2218-5453-2021-4-16-22
21. **Бучинский Д. И., Вознюк В. В.** Помехоустойчивость когерентного демодулятора двоичных фазоманипулированных сигналов с расширенным спектром при воздействии гауссовской помехи с ограниченным по полосе равномерным спектром. *Тр. Военно-космической академии им. А. Ф. Можайского*, 2020, № 675, с. 69–76.
22. **Петешов А. В., Полубехин А. И., Румянцев В. Л.** Структура систем корреляционной пространственно-временной обработки сигналов многоканальных радиолокационных систем. *Изв. Тульского государственного университета. Технические науки*, 2020, № 9, с. 215–225.
23. **Сиразиев Л. Р., Черемисин О. П.** Адаптивный алгоритм обработки сигналов с неизвестными законами временной модуляции в многопозиционных радиолокационных системах. *Электромагнитные волны и электронные системы*, 2018, т. 23, № 6, с. 4–13. doi:10.18127/j15604128-201806-01
24. **Shepeta A. P., Nenashev V. A.** Accuracy characteristics of object location in a two-position system of small onboard radars. *Информационно-управляющие системы*, 2020, № 2, с. 31–36. doi:10.31799/1684-8853-2020-2-31-36
25. **Исаев И. Д., Савельев А. Н., Семенов А. Н.** Анализ калибровочных характеристик для кадров радиолокационной информации наземной многопозиционной радиолокационной системы. *Информационно-измерительные и управляющие системы*, 2020, т. 18, № 6, с. 51–64. doi:10.18127/j20700814-202006-06
26. **Нахмансон Г. С., Акиньшин Д. С.** Обнаружение траекторий движущихся прямолинейно воздушных целей при вторичной обработке радиолокационной информации. *Изв. высших учебных заведений России. Радиоэлектроника*, 2019, т. 22, № 5, с. 61–70. doi:10.32603/1993-8985-2019-22-5-61-70
27. **Дроздов Д. О., Татарский Б. Г.** Алгоритм отождествления первичных измерений с траекториями сопровождаемых целей. *Информационно-измерительные и управляющие системы*, 2018, т. 16, № 7, с. 12–19. doi:10.18127/j20700814-201807-02
28. **Зябиров Э. В., Аравин А. В., Михайлов С. В., Филюшкин И. П.** Выбор вида и параметров стробов при отождествлении координатной информации от средств обнаружения воздушных целей в комплексе средств автоматизации батарейного командного пункта. *Изв. высших учебных заведений. Поволжский регион. Технические науки*, 2018, № 4 (48), с. 88–95. doi:10.21685/2072-3059-2018-4-8
29. **Борисов Е. Г.** Совместная обработка измерений в дальномерно-доплеровской многопозиционной радиолокационной системе. *Научный вестник Московского государственного технического университета гражданской авиации*, 2020, т. 23, № 2, с. 8–19. doi:10.26467/2079-0619-2020-23-2-8-19

UDC 621.396.965
doi:10.31799/1684-8853-2022-3-55-66

Reliability of target selection in the network of geographically separated radar stations in joint processing of radar information in the conditions of relayed interference

D. V. Levin[a], PhD, Tech., Lecturer, orcid.org/ 0000-0002-3480-087X, dm.181@yandex.ru
A. V. Parshutkin[a], Dr. Sc., Tech., Professor, orcid.org/ 0000-0001-7535-4880
A. V. Timoshenko[b], Dr. Sc., Tech., Professor, orcid.org/ 0000-0002-9791-142X
[a]A. F. Mozhaiskiy Military Space Academy, 13, Zhdanovskaia Emb., 197198, Saint-Petersburg, Russian Federation
[b]National Research University of Electronic Technology, Shokin Sq., 1, 124498, Moscow, Zelenograd, Russian Federation

Introduction: Radar information processing methods are used to identify targets at short observation intervals, based on the identification of marks of different radar stations and the parametric identification of a target. However, a mutual comparison of the effectiveness of using such methods of target selection in the conditions of relayed interference has not been carried out. **Purpose:** The comparison of the reliability of target selection in the conditions of relayed interference in a network of geographically separated radar stations for space surveillance with the implementation of correlation ellipsoid method, strobe method of target selection and spatial separation of measured target positions method. **Results:** We give the decision rules for dividing the space of coordinate difference of a target into subdomains to make a decision about the truth of the target. We have carried out simulation modeling of the target selection in the conditions of relayed interference and have obtained the dependencies for the change in the probability of erroneous selection of false marks on the normalized range which is measured from the middle of the spacing base of the two geographically separated radar stations. There are also quantitative estimates of the reliability of target selection for various conditions of radar surveillance by two geographically separated radar stations. The gain in increasing the normalized range with the use of the correlation ellipsoid method as compared to strobe methods of target selection and spatial separation of measured target positions method ranged from 36% to 46%. It is shown that in most practical situations one can use the simplest method, that is the spatial separation of measured target positions method, and when solving most important problems the correlation ellipsoid method can be used. **Practical relevance:** Research results can be used in the development of target selection algorithms in the conditions of relayed interference.

Keywords — target selection, radar information, relayed interference, false marks, correlation ellipsoid.

For citation: Levin D. V., Parshutkin A. V., Timoshenko A. V. Reliability of target selection in the network of geographically separated radar stations in joint processing of radar information in the conditions of relayed interference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 55–66 (In Russian). doi:10.31799/1684-8853-2022-3-55-66

Reference

1. Tatarsky B. G., Ilchuk P. A., Ilchuk A. R. The identification of signals reflected from multiple targets in a distributed radar system, the method of comparison. *Information-measuring and Control Systems*, 2018, vol. 16, no. 1, pp. 33–41 (In Russian).
2. Khomyakov A. V., Filipchenkov V. I., Mamon Yu. I. Algorithms joint trajectory processing in multiposition radar complex. *Proc. of the TSU*, 2016, no. 2, pp. 305–314 (In Russian).
3. Shi Y., Park S., Song T. Multitarget tracking in cluttered environment for a multistatic passive radar system under the DAB/DVB network. *EURASIP Journal on Advances in Signal Processing*, 2017, no. 11. doi:10.1186/s13634-017-0445-4
4. Parshutkin A. V., Levin D. V., Galandzovskiy A. V. Simulation model of radar data processing in a station network under signal-like interference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 6, pp. 22–31 (In Russian). doi:10.31799/1684-8853-2019-6-22-31
5. Agievich S. N., Lutsenco S. A. Application of repeater jamming for effects on satellite radiocommunication systems with direct sequence spread spectrum. *Proc. of the TSU*, 2018, no. 12, pp. 411–416 (In Russian).
6. Chen X., Chen B. Anti-jamming approach based on radar transmitted waveform matching. *EURASIP Journal on Advances in Signal Processing*, 2021, no. 59. doi:10.1186/s13634-021-00776-w
7. Kulikov G. V., Lelyukh A. A., Grachenko E. N. Noise immunity of a coherent signal receiver with quadrature amplitude modulation in the presence of relayed interference. *Journal of Communications Technology and Electronics*, 2020, vol. 65, no. 8, pp. 804–808 (In Russian). doi:10.31857/S0033849420070074
8. Levin D. V. Simulation of target selection by geographically separated radar stations when processing of their radar measurements together in conditions relayed interference. *Questions of Radio Electronics*, 2020, no. 11, pp. 6–13 (In Russian). doi:10.21778/2218-5453-2020-11-6-13
9. Shi C., Zhou J., Wang F. Adaptive resource management algorithm for target tracking in radar network based on low probability of intercept. *Multidimensional Systems and Signal Processing*, 2018, no. 29, pp. 1203–1226. doi:10.1007/s11045-017-0494-8
10. Zhu Y., Zhao L., Zhang Y. Receiver selection for multi-target tracking in multi-static Doppler radar systems. *EURASIP Journal on Advances in Signal Processing*, 2021, no. 118. doi:10.1186/s13634-021-00826-3
11. Smirnov E. E., Pozdniakov A. A. Method of tracking of close objects with different moving characteristics by a long-range radar station. *Electromagnetic Waves and Electronic Systems*, 2021, vol. 26, no. 4, pp. 42–53 (In Russian). doi:10.18127/j15604128-202104-05
12. Vasiliev K. K., Mattis A. V., Saverkin O. V. Strobing of radar marks for trajectory filtration in a body-fixed frame. *Journal of the Russian Universities. Radioelectronics*, 2019, vol. 22, no. 5, pp. 71–79 (In Russian). doi:10.32603/1993-8985-2019-22-5-71-79
13. Aleshkin A. P., Vladimirov V. V., Nevzorov V. I., Savochkin P. V. Method for increasing the resolution and accuracy of radar angular measurements based on sequential spatio-temporal processing of received signals. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 2, pp. 37–45 (In Russian). doi:10.31799/1684-8853-2020-2-37-45
14. Smirnov E. E., Pozdniakov A. A., Parshin M. S. Model of classification of observation objects under conditions of intersection of their motion paths based on joint analysis of trajectory and polarization information. *Information-measuring and Control Systems*, 2021, vol. 19, no. 4, pp. 14–26 (In Russian). doi:10.18127/j20700814-202104-02
15. Zarai K., Cherif A. Adaptive filter based on Monte Carlo method to improve the non-linear target tracking in the radar system. *Aerospace Systems*, 2021, no. 4 (7), pp. 67–74. doi:10.1007/s42401-020-00080-9
16. Palguyev D. A., Shentyabin A. N. More on assessing the probability of radar information association during tertiary information processing in network structures. *Radio Industry (Russia)*, 2020, vol. 30, no. 2, pp. 32–41 (In Russian). doi:10.21778/2413-9599-2020-30-2-32-41
17. Kiryushkin V. V., Korovin A. V., Zhuravlev A. V. Sensor-to-sensor data association and preliminary estimation of coordinates of air target in multisensory radar in multitarget conditions. *Journal of Siberian Federal University.*

Engineering and Technologies, 2019, vol. 12, no. 6, pp. 709–718 (In Russian). doi:10.17516/1999-494X-0170
18. Zhuravlev A. V., Kiryushkin V. V., Korovin A. V. Sensor-to-sensor data association in total-range-measurements multisensory radar in multitarget conditions. *Radioengineering*, 2019, vol. 83, no. 6 (8), pp. 180–189 (In Russian). doi:10.18127/j00338486-201906(8)-16
19. Buchinskiy D. I., Voznyuk V. V., Fomin A. V. Research of noise stability of the receiver with M-PSK modulation under the influence of interference with different structure. *Proc. of the Mozhaisky Military Space Academy*, 2019, no. 671, pp. 120–127 (In Russian).
20. Aleshkin A. P., Balashov V. M., Vladimirov V. V. Signals synthesis of artificial aperture in a digital antenna array by extrapolation of the opening function based on sequential processing of reflected radar signals. *Questions of Radio Electronics*, 2021, no. 4, pp. 16–22 (In Russian). doi:10.21778/2218-5453-2021-4-16-22
21. Buchinskiy D. I., Voznyuk V. V. Immunity of a coherent demodulator of binary phasomanipulated signals with a spread spectrum under exposure to gaussian interference with a band-bounded uniform spectrum. *Proc. of the Mozhaisky Military Space Academy*, 2020, no. 675, pp. 69–76 (In Russian).
22. Peteshov A. V., Polubahin A. I., Rumyantsev V. L. The structure of correlation systems spatio-temporal processing of multichannel radar signals. *Proc. of the TSU*, 2020, no. 9, pp. 215–225 (In Russian).
23. Siraziev L. R., Cheremisin O. P. Adaptive algorithm of processing of signals with unknown modulation law in multistatic radar systems. *Electromagnetic Waves and Electronic Systems*, 2018, vol. 23, no. 6, pp. 4–13 (In Russian). doi:10.18127/j15604128-201806-01
24. Shepeta A. P., Nenashev V. A. Accuracy characteristics of object location in a two-position system of small onboard radars. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 2, pp. 31–36. doi:10.31799/1684-8853-2020-2-31-36
25. Isaev I. D., Savelyev A. N., Semenov A. N. Analysis of calibration methods to raw radar frames received from multi-radar surface movement radars. *Information-measuring and Control Systems*, 2020, vol. 18, no. 6, pp. 51–64 (In Russian). doi:10.18127/j20700814-202006-06
26. Nakhmanson G. S., Akinshin D. S. Detection of the trajectories of moving rectilinearly air targets in the secondary processing of radar information. *Journal of the Russian Universities. Radioelectronics*, 2019, vol. 22, no. 5, pp. 61–70 (In Russian). doi:10.32603/1993-8985-2019-22-5-61-70
27. Drozdov D. O., Tatarsky B. G. Algorithm for identification primary measurements with tracked trajectories. *Information-measuring and Control Systems*, 2018, vol. 16, no. 7, pp. 12–19 (In Russian). doi:10.18127/j20700814-201807-02
28. Zyabirov E. V., Aravin A. V., Mikhaylov S. V., Filyushkin I. P. The choice of the form and parameters of strobes at the identification of coordinate information from air target sensors in a complex of automation equipment of a battery command post. *University Proceedings. Volga Region. Technical Science*, 2018, no. 4 (48), pp. 88–95 (In Russian). doi:10.21685/2072-3059-2018-4-8
29. Borisov E. G. Joint processing of measurements in a rangefinder-Doppler multistatic radar system. *Civil Aviation High Technologies*, 2020, vol. 23, no. 2, pp. 8–19 (In Russian). doi:10.26467/2079-0619-2020-23-2-8-19

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (ius.spb@gmail.com).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

Инвестиционная платформа «Инвестиционный Компас» на службе внедрения разработок российских ученых

Современный исторический этап проходит в ситуации недофинансирования экономики в развитие предпринимательства в России, острой необходимости противостояния западным ограничениям, требований развития реального импортозамещения посредством разработки и реализации соответствующих инновационных проектов. В этих обстоятельствах важно создавать благоприятные условия российской науке для популяризации и внедрения своих инновационных научных разработок, привлечения к ним внимания инвесторов, используя для этого все имеющиеся возможности и допущения, следующие из действующего законодательства России.

Акционерное общество «*Специализированный Регистратор «КОМПАС»* (https://zao-srk.ru/) (АО «СРК»), являясь профессиональным участником финансового рынка, подконтрольным Банку России, исторически развивалось как организация, старающаяся в высшей степени реализовать установки президента Российской Федерации и правительства Российской Федерации. И в настоящее сложное для российской экономики время оно прилагает максимум усилий для создания инфраструктуры, способствующей внедрению научных и инновационных разработок российских ученых для развития в России импортозамещения и инновационного предпринимательства. Так, АО «СРК» одним из первых в России создало инвестиционную платформу «*Инвестиционный Компас*» (https://in-ko.ru/) — новый современный цифровой сервис, обеспечивающий содействие в привлечении и осуществлении инвестирования. Эта первая в России полноценная инвестиционная платформа, созданная АО «СРК» во исполнение приоритетов, заявленных президентом России в своих указах и поручениях, действующая по принципам коллективного инвестирования, призвана обеспечивать финансирование реализации различных проектов в интересах России.

Порядок функционирования в России инвестиционных платформ регламентирован Федеральным законом от 02 августа 2019 года № 259-ФЗ «О привлечении инвестиций с использованием инвестиционных платформ и о внесении изменений в отдельные законодательные акты Российской Федерации». Инвестиционная платформа «Инвестиционный Компас» обеспечивает мобилизацию финансовых ресурсов лиц, зарегистрировавшихся на ней в качестве потенциальных инвесторов, *для финансирования проектов внедрения научных и инновационных разработок, для реализации проектов импортозамещения, инфраструктурных, инновационных, индустриальных и иных проектов субъектов малого и среднего предпринимательства, а также стартапов*. Способами привлечения инвестиций, предусмотренными инвестиционной платформой «Инвестиционный Компас», являются реализация (размещение) авторами проектов эмиссионных ценных бумаг (акций), реализация (размещение) утилитарных цифровых прав, а также предоставление лицам, привлекающим инвестиции, займов.

Для популяризации научных и инновационных разработок на сайте инвестиционной платформы «Инвестиционный Компас» создан специальный раздел, где авторы перспективных идей, разработок, новаций могут размещать информацию о них для доведения до сведения широкого круга потенциальных инвесторов; для разработки совместно с представителями предпринимательского сообщества проектов, необходимых для внедрения результатов своих научных изысканий; для привлечения инвестиций на внедрение разработок. С использованием возможностей инвестиционной платформы «Инвестиционный Компас» авторы смогут донести до широкого круга предпринимателей и инвесторов информацию о результатах своих научных изысканий, а также указать в публикации удобные способы контактирования предпринимателей и инвесторов с авторами.

Инвестиционную платформу могут использовать авторы журнала для размещения развернутой аннотации (анонса) статьи с изложением информации о сути научной разработки, со ссылкой на статью в журнале. Мы предлагаем авторам размещение информации на сайте инвестиционной платформы «Инвестиционный Компас» на безвозмездной основе. Российские ученые также имеют возможность самостоятельно или с помощью специалистов оператора платформы разработать инвестиционный проект внедрения результатов своих научных изысканий и разместить его на сайте инвестиционной платформы «Инвестиционный Компас» для привлечения инвестиций. Автором проекта на инвестиционной платформе (лицом, привлекающим инвестиции) может быть российское юридическое лицо любой организационно-правовой формы или индивидуальный предприниматель.

Алехин Андрей Юрьевич, генеральный директор
АО «Специализированный Регистратор «КОМПАС»
alehin@zao-srk.ru

Масштабный образовательный интенсив — первая летняя «Школа академического совершенства»

15–30 августа 2022 г.

Первая летняя «Школа академического совершенства» — это образовательный проект, который объединяет и развивает людей, работающих в сфере науки и образования: более 20 000 участников и 50 экспертов примут участие в более чем 40 мероприятиях.

Вас ждут семинары ученых и экспертов-практиков, курсы повышения квалификации, онлайн мастер-классы и мастерские, а также четыре тематических трека: академическое мастерство, современное управление, открытая наука и образовательные интенсивы.

Формат проведения

Онлайн

Кому будет интересно

Научно-педагогическим работникам вузов

Ученым вузов и научных институтов

Молодым ученым и аспирантам

Руководителям научных и научно-образовательных организаций высшего и среднего звена

Всем заинтересованным в саморазвитии и развитии людей в сфере науки и образования

Программа

Трек 1. Академическое мастерство

Кому будет интересно: работникам вузов и научных институтов

Развиваем компетенции: академическое письмо, работа научной команды, цифровые инструменты педагога, педагогический дизайн, эффективные практики преподавания в высшей школе, научная коммуникация и многое другое

Трек 2. Современное управление

Кому будет интересно: руководителям в области науки, образования и инновационной деятельности

Лучшие практики управления в образовании, науке, коммерциализации технологий. Передовые управленческие технологии в проектной деятельности, кадровой политике, цифровых решениях

Трек 3. Открытая наука

Кому будет интересно: молодым ученым и состоявшимся исследователям

Школы от уникальных научных коллективов и ведущих ученых, системные дискуссии о путях развития в науке, семинар представителей различных индустрий о взаимодействии науки и бизнеса

Трек 4. Образовательные интенсивы

Что вас ждет: интерактивные образовательные курсы, семинары, стратегические интенсивы

Возможность системно «прокачать» практические компетенции по различным областям научно-образовательной и исследовательской деятельности, выстроить продуктивные партнерства и сети, построить свои стратегии и планы

Организаторы

Тюменский государственный университет

Академия управления WINbd

Западно-Сибирский межрегиональный научно-образовательный центр

«Школа» создана учеными и экспертами-практиками, чтобы объединить профессионалов и передать лучшие практики: новые форматы управления проектами и коллективами, трансформация методов преподавания, работа со знаниями и технологиями.

В пространстве «Школы» вы сможете развить собственные компетенции, организовать развитие целых коллективов и команд. Присоединяйтесь!

Контакты

Сайт «Школы академического совершенства»: sae.utmn.ru

По вопросам регистрации участников: cdc@scitech.ru

БАЛОНИН
Николай
Алексеевич

Профессор кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1982 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Автоматика и телемеханика».
В 2008 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 100 научных публикаций, в том числе трех монографий.
Область научных интересов — теория динамических систем, теория идентификации, теория операторов, теория матриц, вычислительные методы, интернет-робототехника, интернет-книги с исполняемыми алгоритмами, научные социальные сети.
Эл. адрес: korbendfs@mail.ru

ЛЕБЕДЕВ
Илья
Сергеевич

Профессор, главный научный сотрудник Санкт-Петербургского Федерального исследовательского центра РАН.
В 1998 году окончил Санкт-Петербургское высшее военное училище ПВО по специальности «Инженер-математик».
В 2012 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 80 научных публикаций.
Область научных интересов — информационные технологии, машинное обучение, компьютерная лингвистика.
Эл. адрес: isl_box@mail.ru

ЛЯШКОВ
Михаил
Андреевич

Аспирант кафедры математического моделирования и информационных технологий Тамбовского государственного университета им. Г. Р. Державина.
В 2018 году окончил Тамбовский государственный технический университет по специальности «Информационная безопасность автоматизированных систем».
Является автором четырех научных публикаций.
Область научных интересов — методы и системы искусственного интеллекта в поддержке принятия решений.
Эл. адрес:
iwishcoolwork@gmail.com

КОВАЛЕВА
Ольга
Александровна

Профессор кафедры математического моделирования и информационных технологий Тамбовского государственного университета им. Г. Р. Державина, профессор Тамбовского государственного технического университета.
В 2005 году окончила Тамбовский государственный университет им. Г. Р. Державина по специальности «Математика».
В 2019 году защитила диссертацию на соискание ученой степени доктора технических наук.
Является автором более 150 научных публикаций и 22 объектов интеллектуальной собственности.
Область научных интересов — математическое и компьютерное моделирование динамических систем.
Эл. адрес:
solomina-oa@yandex.ru

ЛЕВИН
Дмитрий
Викторович

Преподаватель кафедры систем и средств радиоэлектронной борьбы Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург.
В 2006 году окончил Военно-космическую академию им. А. Ф. Можайского по специальности «Автоматизированные системы обработки информации и управления».
В 2014 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 35 научных публикаций.
Область научных интересов — радиоэлектронная защита радиоэлектронных систем, техническая защита информации.
Эл. адрес: dm.181@yandex.ru

ОВЧИННИКОВ
Андрей
Анатольевич

Заведующий кафедрой безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2000 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Информационные системы в экономике».
В 2004 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 50 научных публикаций и шести международных патентов на изобретения.
Область научных интересов — теория помехоустойчивого кодирования, кодовая криптография.
Эл. адрес: mldoc@mail.ru

ПАРШУТКИН Андрей Викторович

Профессор кафедры систем и средств радиоэлектронной борьбы Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург.
В 1980 году окончил Военный инженерно-космический институт им. А. Ф. Можайского по специальности «Радиотехнические системы и средства контроля».
В 2010 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 155 научных публикаций и четырех патентов на изобретения.
Область научных интересов — радиоэлектронная защита радиоэлектронных систем, техническая защита информации.
Эл. адрес: andydc2010@mail.ru

СЕРГЕЕВ Александр Михайлович

Доцент кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2004 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети».
В 2020 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 35 научных публикаций.
Область научных интересов — численные методы, теория вычислительных процессов, проектирование специализированных процессоров.
Эл. адрес: asklab@mail.ru

ТИМОШЕНКО Александр Васильевич

Профессор, начальник лаборатории Московского института электронной техники.
В 1981 году окончил Минское высшее инженерное зенитно-ракетное училище по специальности «Радиотехнические средства».
В 2004 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 160 научных публикаций.
Область научных интересов — разработка и испытание радиоинформационных систем.
Эл. адрес: u567ku78@gmail.ru

ПЧЕЛИНЦЕВ Сергей Юрьевич

Аспирант кафедры математического моделирования и информационных технологий Тамбовского государственного университета им. Г. Р. Державина.
В 2018 году окончил Тамбовский государственный технический университет по специальности «Информационная безопасность автоматизированных систем».
Является автором восьми научных публикаций и одного свидетельства о регистрации программы для ЭВМ.
Область научных интересов — методы и системы искусственного интеллекта в поддержке принятия решений.
Эл. адрес: veselyrojer@mail.ru

СУХОВ Александр Максимович

Докторант Краснодарского высшего военного училища им. генерала армии С. М. Штеменко.
В 2009 году окончил Краснодарское высшее военное училище (Военный институт) им. генерала армии С. М. Штеменко по специальности «Комплексная защита объектов информатизации».
В 2018 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 30 научных публикаций.
Область научных интересов — теория эффективности целенаправленных процессов, информационная безопасность.
Эл. адрес: 19am87@mail.ru

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

Статьи должны содержать изложение новых научных результатов. Название статьи должно быть кратким, но информативным. В названии недопустимо использование сокращений, кроме самых общепринятых (РАН, РФ, САПР и т. п.).

Текст рукописи должен быть оригинальным, а цитирование и самоцитирование корректно оформлено.

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word шрифтом Times New Roman размером 13, поля не менее двух сантиметров.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, ORCID и электронный адрес одного из авторов. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы. Предоставляйте подрисуночные подписи и названия таблиц на русском и английском языках.

Статьи авторов, не имеющих ученой степени, рекомендуется публиковать в соавторстве с научным руководителем, наличие подписи научного руководителя на рукописи обязательно; в случае самостоятельной публикации обязательно предоставляйте заверенную по месту работы рекомендацию научного руководителя с указанием его фамилии, имени, отчества, места работы, должности, ученого звания, ученой степени.

Простые **формулы** набирайте в Word, сложные с помощью редактора Mathtype или Equation. Для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта в Mathtype никогда не пользуйтесь вкладкой Other, Smaller, Larger, используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; пробелы в формуле ставьте только после запятых при перечислении с помощью Ctrl+Shift+Space (пробел); не отделяйте пробелами знаки: + = − ×, а также пространство внутри скобок; для выделения греческих символов в Mathtype полужирным начертанием используйте Style → Other → bold.

Для набора формул в Word никогда не используйте вкладки: «Уравнение», «Конструктор», «Формула» (на верхней панели: «Вставка» — «Уравнение»), так как этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Подробнее см. pdf-файл «Правила подготовки рукописей» (стр. 11) на сайте https://guap.ru/ric

Иллюстрации:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, поддающихся редактированию, используя векторные программы: Visio (*.vsd, *.vsdx); Adobe Illustrator (*.ai); Coreldraw (*.cdr, версия не выше 15); Excel (*.xls); Word (*.docx); AutoCad, Matlab (экспорт в PDF, EPS, SVG, WMF, EMF); Компас (экспорт в PDF), веб-портал DRAW.IO (экспорт в PDF);

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

Наличие подрисуночных подписей и названий таблиц на русском и английском языках обязательно (желательно не повторяющих дословно комментарии к рисункам в тексте статьи).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, e-mail), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png, *.jpg с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40×55 мм;

— экспертное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц, doi;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц, doi;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Список литературы оформляйте двумя отдельными блоками по образцам lit.dot на сайте журнала (http://i-us.ru/paperrules): Литература и References.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Руководство для авторов».

Контакты

Куда: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Кому: Редакция журнала «Информационно-управляющие системы»
Тел.: (812) 494-70-02
Эл. почта: ius.spb@gmail.com
Сайт: www.i-us.ru