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

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

1(122)/2023

doi:10.31799/1684-8853-2023-1

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

SCIENTIFIC JOURNAL

1(122)/2023

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 09.01.23. Passed for printing 03.03.23. Signed for printing: 07.03.2023.
Format $60\times84_{1/8}$. Phototype CentSchbkCyrill BT. 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:10.31799/1684-8853-2023-1 ISSN 1684-8853 (print); ISSN 2541-8610 (online)

НАУЧНЫЙ ЖУРНАЛ

1(122)/2023

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

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

Редактор: А. Г. Ларионова
Корректор: Т. В. Звертановская
Дизайн: М. Л. Черненко, Ю. В. Умницына
Компьютерная верстка: Ю. В. Умницына

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 09.01.23. Подписано в печать 03.03.23. Дата выхода в свет: 07.03.2023.
Формат 60×841/8. Гарнитура CentSchbkCyrill BT. Печать цифровая.
Усл. печ. л. 8,7. Уч.-изд. л. 11,9. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 47.
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:10.31799/1684-8853-2023-1-2-16
EDN: KOMNBV

Научные статьи
Articles

Задачи разрешимые и неразрешимые. Алгоритм Прокруста получения матриц семейства Адамара

Н. А. Балонин[а], *доктор техн. наук, профессор, orcid.org/0000-0001-7338-4920, korbendfs@mail.ru*
Дж. Себерри[б], *доктор наук, профессор, orcid.org/0000-0002-9558-4293*
М. Б. Сергеев[а], *доктор техн. наук, профессор, orcid.org/0000-0002-3845-9277*
[а]*Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ*
[б]*Университет Вуллонгонг, Вуллонгонг, Новый Южный Уэльс 2522, Австралия*

Введение: *развитие теории матриц Адамара столкнулось с препятствием, обусловленным не столько природой целочисленной задачи, сколько искусственным ограничением решения квадратичных уравнений переборным путем. Игнорирование прямого пути, отказ от иррациональности привели к появлению мнения, что гипотеза существования матриц Адамара недоказуема.* ***Цель:*** *обосновать разрешимость задачи Адамара ортогональными матрицами за счет выявления их устойчивой связи с матрицами с иррациональными элементами.* ***Результаты:*** *показано, что иррациональность проявляется в квадратичной норме столбцов матрицы Адамара второго порядка. Проанализирован перенос итерационных алгоритмов вычисления корней на матричный случай. Предложен алгоритм Прокруста минимизации максимального по абсолютному значению элемента ортогональной матрицы. Поскольку матрицы Адамара определены инвариантами вложенных в ее структуру матриц меньшего порядка, алгоритм оказывается универсальной основой для их совместного нахождения. Гипотеза о существовании матриц Адамара рассматривалась в оперативной области итерационных алгоритмов, определенных над полем вещественных чисел, дающих преимущества перед инструментами в форме конечных полей и групп.* ***Практическая значимость:*** *ортогональные последовательности, получаемые из строк (столбцов) матриц Адамара, и сами матрицы Адамара высоких порядков имеют большое практическое значение для задач помехоустойчивого кодирования, сжатия, маскирования и обработки изображений.*

Ключевые слова *— матрицы Адамара, конференц-матрицы, критские матрицы, алгоритм Прокруста, конечные поля, симметрии матриц.*

Для цитирования: Балонин Н. А., Себерри Дж., Сергеев М. Б. Задачи разрешимые и неразрешимые. Алгоритм Прокруста получения матриц семейства Адамара. *Информационно-управляющие системы*, 2023, № 1, с. 2–16. doi:10.31799/1684-8853-2023-1-2-16, EDN: KOMNBV

For citation: Balonin N. A., Seberry J., Sergeev M. B. Solvable and unsolvable problems. Using Procrustes analysis algorithm for obtaining a family of Hadamard matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 2–16 (In Russian). doi:10.31799/1684-8853-2023-1-2-16, EDN: KOMNBV

Введение

Три знаменитые задачи античной древности на удвоение куба, трисекцию угла и квадратуру круга известны тем, что их не решить при помощи циркуля и линейки. Эти неразрешимые задачи были решены иначе выдающимися учеными, сумевшими систематизировать средства достижения цели и их эффективность. Изучение таких задач привело к понятию иррационального числа.

В современных обозначениях задача на удвоение куба сводится к решению уравнения $x^3 = 2$. Задача на трисекцию угла α связана с тригонометрическим уравнением $x^3 = 3x + 2\cos(\alpha)$. Если принять за единицу измерения радиус круга и обозначить через x длину стороны искомого квадрата задачи на квадратуру круга, то задача сводится к решению уравнения $x^2 = \pi$. Таким образом, неразрешимость задачи на квадратуру круга следует из неалгебраичности (трансцендентности) числа π, которая была доказана в 1882 г. Линдеманом. Из его теоремы следует, что осуществить решение нельзя с помощью прямых, окружностей или любых других алгебраических кривых и поверхностей, например эллипсов, гипербол или кубических парабол.

Аристотель в IV в. до н. э. писал: «Посредством геометрии нельзя доказать, что два куба составляют один куб». Однако эту неразрешимость следует понимать как неразрешимость при использовании *только* циркуля и линейки. Простейший механический способ решения предложил Леонардо да Винчи.

Противоречивый ход развития математики сначала отвергает очевидное (у диагонали равнобедренного прямоугольного треугольника есть длина, невыразимая числом), а потом находит решение, в котором сам этот «неразрешимый» треугольник становится генератором новых чи-

сел. Мы их «включаем» в систему за счет того, что множество рациональных чисел «расширяется» иррациональными числами, образуя новое понятие — вещественное число. С корнем из –1 это произошло позднее и иначе — комплексным числам приписали положение на плоскости. В этом проявляется пестрота математики.

Как известно, иррациональное число не может быть записано в виде обыкновенной дроби m/n, где m, n — целые числа, но может быть представлено в виде бесконечной непериодической десятичной дроби. Иррациональными являются, среди прочих, отношение длины окружности к диаметру круга (число π), число Эйлера e, золотое сечение φ, квадратный корень из двух. Все квадратные корни натуральных чисел, кроме полных квадратов, иррациональны. Каждое иррациональное число является либо алгебраическим, либо трансцендентным. Множество алгебраических чисел является счетным множеством корней полиномов с целыми коэффициентами. Простейшими являются квадратичные иррациональности. Поскольку множество вещественных чисел несчетно, то множество иррациональных чисел также несчетно.

К. Гаусс, рассматривая построение правильного семнадцатиугольника, пользовался тем, что с помощью циркуля и линейки можно выполнить все четыре арифметических действия и осуществить извлечение квадратного корня. Все остальное надо делать иначе. Подобное надо изучать подобным, иначе возникает коллизия. Однако не стоит думать, что переход к алгебраической форме записи автоматически упрощает рассмотрение проблемы. Так, например, количество и характер решений задачи, записанной в виде пары квадратичных уравнений, становится очевиднее, если видеть за уравнениями пару окружностей, описываемых ими. Две разные окружности со смещенными центрами могут пересекаться либо в одной, либо в двух точках. Для установления этого обстоятельства достаточно не столько углубленного знания геометрии, сколько жизненного опыта.

Иррациональные числа настолько абстрактны, что для их обозначения у нас нет привычных средств записи. Их обозначают, по сути, уравнениями (формулами), коэффициентами которых являются рациональные числа. Решения уравнений, включающих многочлены с рациональными коэффициентами, не всегда столь очевидны, как решения задач геометрии. Но если они имеют вещественные значения, то, в отличие от уравнений Диофанта, этот случай, согласно теореме Тарского, классифицируется как более простой.

Цель данной работы состоит в расширении теории матриц Адамара и демонстрации преодоления принципиальных трудностей их поиска за счет использования выявленной связи с иррациональными матрицами соседних порядков. Эта связь образует доказательную базу существования матриц Адамара, снимая ограничение, связанное с использованием комбинаторных методов.

Теорема Тарского и матричные квадратичные уравнения

А. Тарский обнаружил, что использование языка формул открывает путь для установления истинности утверждений относительно бесконечного числа объектов совершением конечного числа манипуляций.

В самом деле, значения полинома в промежутках между его корнями могут быть какими угодно, но они не меняют знака. Поэтому, несмотря на то, что функция определена над бесконечным множеством точек, существует конечный алгоритм построения таблицы Тарского (орнамента), помогающей вынести суждение о качествах всех их [1]. Иными словами, Тарский свел бесконечномерную задачу к задаче построения конечного орнамента, описываемого орнаментальными инвариантами (присущими узору параметрами).

Задачи на поиск корня уравнения $x^2 = n$, где n — целое число, относятся к числу фундаментальных, сложивших основание современной теории чисел и алгебраической геометрии. Итерационные алгоритмы систематизировал Ньютон, хотя их начали предлагать еще в глубокой древности. Таков, например, алгоритм Герона вычисления приближения к корню квадратному, начиная с некоторого любого положительного числа.

Задачи на поиск экстремума или (в иной трактовке) *неподвижной точки* отображения, задаваемого алгоритмом Герона, тесно взаимосвязаны. Поэтому в основании итерационных алгоритмов могут звучать как оптимизационные мотивы, так и мотивы, навеянные тематикой обобщений метода Герона. Важно понять, какой именно оптимизации и какого именно отображения, а также изучить их свойства.

Матричное расширение задачи

$$\mathbf{H}^{\mathrm{T}}\mathbf{H} = n\mathbf{I}, \qquad (1)$$

где $\mathbf{I}$ — единичная матрица того же порядка n, что и искомая матрица $\mathbf{H}$, уже при $n = 2$ расширяет пространство аргументов, элементов матрицы, среди которых находятся и целочисленные [2, 3].

Этот переход заметил в свое время основатель теории матриц Дж. Сильвестр в связи с получе-

нием на порядках, кратных степеням 2, орнаментов (матриц с элементами 1 и –1) [4], инверсия которых сводится к транспонированию и масштабированию элементов. Узор из знаков является общим инвариантом как прямой, так и обратной матриц.

Заинтересованный в геометрическом толковании алгебраических задач Ж. Адамар дополнил это наблюдение [5]. Во-первых, он особо выделил порядки, кратные четырем, доказав, что только для них возможно целочисленное решение задачи матрицей с элементами 1 и –1. Во-вторых, он отметил важное свойство получаемых таким образом решений: они обладают еще одним инвариантом — максимумом детерминанта на классе матриц с элементами, по модулю не превышающими единицу. С тех пор задача остановилась в своем развитии, поскольку перекочевала в класс заведомо трудных задач с сугубо целочисленными решениями.

Комбинаторные методы, связанные с перебором индексов (адресов) отрицательных (положительных) элементов в матрице, изначально слабые, сильно развились с появлением эффективных алгоритмов теорий конечных полей и групп [6, 7]. Например, Н. Ито установил взаимно однозначное соответствие между орнаментами матриц Адамара и конструкциями, которые можно построить, вооружившись арифметикой дициклических групп [8, 9]. Тем не менее эта попытка и иные приемы использования конечномерной математики не дают основания однозначно судить, разрешима ли задача Адамара на всех выделенных им порядках или нет.

Это все большей частью методы ускорения решения при некоторых благоприятных условиях, но условия не одинаковы для различных областей порядков вида $n = 4t$, где t — натуральное число. Перебор, как бы сложно ни был оформлен использующий его метод через теорию групп или полей, не дает повода определенно заключить, закончится ли поиск нужной для решения комбинацией элементов 1 и –1. В итоге с помощью комбинаторного подхода отказались от попыток доказать гипотезу Адамара, что отмечается в публикациях, начиная с основополагающих работ Р. Пэли и Дж. Вильямсона [10, 11].

Между тем породивший это направление скалярный случай относится в настоящее время к тривиальным, разрешимым даже не одним известным способом.

Мы отмечаем отход от общего пути решения (итерацией), чем и вызвано данное недоразумение. В самом деле, достаточно нормировать столбцы матрицы второго порядка, делая ее ортогональной в смысле $\underline{\mathbf{H}}^{\mathrm{T}}\underline{\mathbf{H}} = \mathbf{I}$, чтобы элементы стали обратно пропорциональными корню квадратному из двух. Выход темы орнаментов, поднятой Сильвестром, за пределы достижимости их итерационными методами не обязателен. Можно сформулировать ту же проблему иначе, а именно как поиск ортогональной матрицы $\underline{\mathbf{H}}$ с минимальным максимальным элементом [12], и тогда задача не относится к классу целочисленных задач Диофанта.

Такая оптимизационная, а не переборная задача имеет с точки зрения фиксации орнамента (узора) то же самое решение, но по значениям элементов оно окажется рациональным или иррациональным. Целочисленное толкование задачи появилось ввиду того, что, если элементы матрицы равны друг другу по абсолютной величине, их значения можно исключить делением на неудобное нам иррациональное число. Если это исключение постулировать и положить в основу определения матрицы Адамара, как сделано в большинстве работ, мы отсекаем все то, с чего это начиналось. Трактовка проблемы как переборной задачи, осуществленная на раннем этапе ее постановки, оказалась удобной для приложения эффектных методов комбинаторной теории. Но она не согласована с целью поиска ответа на вопрос, а всегда ли есть искомое решение?

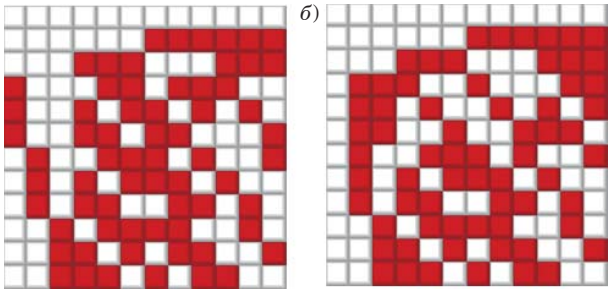
Мы намерены дать более полное представление об этом пути развития теории ортогональных матриц, поскольку это кажется нам назревшим и более важным, чем очередной комбинаторный алгоритм.

Инварианты матриц семейства Адамара

Для матриц Адамара (1) порядка n с элементами 1, –1, полученных каким-либо методом, не обнаруживаются, на первый взгляд, свойства для их идентификации. Однако вычисленные Адамаром матрицы порядков 12 и 20 эквивалентными преобразованиями путем умножения строк или столбцов на –1 и структурирующими перестановками можно привести к нормальному виду с «каймой» из единиц, располагаемых в первых ее строке или столбце.

После добавления к матрице каймы обнаруживаются два инварианта, характерных для всех матриц Адамара. Ими являются количество единиц (или –1), определяемое в любой ее строке или столбце как $k = n/2$ (без учета каймы), и $\lambda = n/4$ — количество единиц (или –1), совпадающих по местоположению в любой паре строк и столбцов [2, 3]. Проверить это обстоятельство можно на примере портретов взятой из статьи Адамара матрицы порядка 12 ($k = 6$ и $\lambda = 3$) (рис. 1, *а*) и ее нормальной формы (рис. 1, *б*) [13].

В статьях встречаются как нормальные формы, так и инварианты блочных конструкций [12, 14]. Допустим, блочно-составная матрица



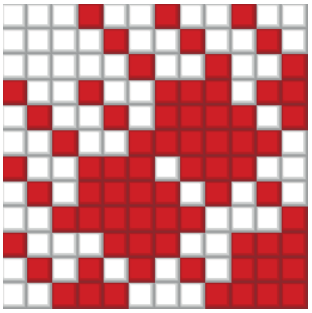
■ ***Рис. 1.*** Портреты матрицы Адамара порядка 12 (*а*) и ее нормальной формы (*б*)

■ ***Fig. 1.*** Portraits of the Hadamard matrix of order 12 (*а*) and its normal form (*б*)

Адамара с частными инвариантами блоков k_1, k_2, ..., подсчитываемыми для элемента –1, не приведена к нормальному виду, как это видно на портрете матрицы конструкции Пропус [12], представленной на рис. 2.

Естественно, суммарный инвариант $k = k_1 + k_2 + \ldots$ будет отличаться от инварианта нормальной формы. Формула для подсчета второго инварианта $\lambda = k - n/4$ универсальна, но в силу отличия k она дает иную (тоже верную) оценку, опираясь на иное упорядочивание 1 и –1 в пределах выделяемой блоками части всей матрицы. Симметричная конструкция Пропус (см. рис. 2) опирается на равенство двух (из четырех) средних блоков и удобна тем, что ее инварианты $k_1 = (v - x)/2$, $k_2 = k_3 = (v - y)/2$, $k_4 = (v - z)/2$, где $v = n/4$ — размер блока, можно классифицировать с помощью точек Гаусса на сфероиде $x^2 + 2y^2 + z^2 = n$, связывая разрешимость этого уравнения с теоремами Гаусса и Лиувилля [15, 16].

Как видно, диагональный блок размера 3×3 блочно-составной матрицы порядка 12 состоит из единиц и $k_1 = 0$. Три оставшихся блока содержат отрицательный элемент на диагонали. Суммарное $k = 3$, а для нормальной формы $k = n/2 = 6$.



■ ***Рис. 2.*** Матрица Адамара порядка 12 конструкции Пропус

■ ***Fig. 2.*** Propus construction of Hadamard matrix of order 12

Соответственно, $\lambda = k - n/4 = 3 - 3 = 0$ (для нормальной формы $\lambda = 3$), и ортогональность всей матрицы гарантируется конструкцией этого массива. Для кососимметричного массива Себерри [2] базовым является уравнение сферы $x^2 + y^2 + z^2 = n - 1$, а не сфероида из работы [17]. Все остальные соображения сохраняются.

Комбинаторика толкует о недоказуемости гипотезы Адамара. Да, но недоказуемость с использованием каких инструментов? Решение задачи о трисекции угла не позволяет говорить о неразрешимости вообще. Неразрешимость — следствие неверного выбора инструментов. Действительно, вопрос о разрешимости квадратичного матричного уравнения перебором не решить, хотя можно найти большое количество строк и столбцов квадратной матрицы с $k = n/2$ положительными и отрицательными элементами. Однако обеспечение второго инварианта $\lambda = n/4$ и тем самым получение матрицы Адамара не гарантировано.

Это те же «циркуль и линейка», но для поставленной Адамаром задачи, имеющей и другую формулировку. Согласно ей матрицы Адамара являются матрицами максимума детерминанта на множестве матриц с ортогональными строками и столбцами. При этом значения элементов не превышают единицу по абсолютной величине. Следовательно, имеется возможность находить матрицу Адамара путем оптимизации детерминанта ортогональной матрицы того же порядка. Про матрицы *максимума детерминанта* известно то, что они существуют всегда и имеют элементы 1 и –1 [3], а на порядках $4t$ совпадают с матрицами Адамара.

Инвариантов у таких матриц много, и они могут находиться переборными процедурами, но будут ли инварианты ортогональными? Ответ на данный вопрос получить с помощью комбинаторной математики невозможно, поскольку алгоритмы на основе переборного подхода могут гарантировать только оптимум, но не ортогональность.

Теперь зададимся вопросом: а является ли матрицей матрица Адамара?

Орнаменты иррациональных матриц семейства Адамара

То, что мы рассматриваем не матрицу, а гиперобъект как узор, подчеркивается тем, что он может быть представлен не одной ортогональной матрицей, а несколькими матрицами смежных порядков. То есть порядок матрицы (размер узора) в этой задаче понятие весьма растяжимое. Проекция какого-либо математического объекта — это тень, обладающая тем качеством, что

она не обязательно совпадает со всем объектом, но несет о нем информацию. Присмотримся к основе, взятой с обратным знаком от матрицы Адамара без нормализующей каймы порядка $m = n - 1$.

Количество положительных элементов в каждой строке у матрицы сохранится, как и второй инвариант, вычисляемый по ним. Эту матрицу можно сделать ортогональной по строкам (столбцам), заменив значения элементов 1 и –1 на $a = 1$ и $-b$. Тогда скалярное произведение двух соседних строк матрицы порядка $m = 4\lambda - 1$ будет иметь λ значений a^2, $2(k - \lambda) = 2\lambda$ произведений $-ab$ ($k - \lambda$ элементов a каждой из строк умножено на $-b$) и $\lambda - 1$ значений b^2. Таким образом, условие ортогональности определяется как $(\lambda - 1)b^2 - 2\lambda ab + \lambda a^2 = 0$.

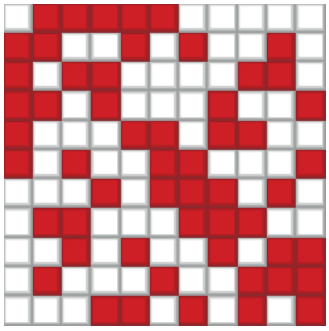
Положительный корень приведенного полинома $b = \frac{\lambda}{\lambda + \sqrt{\lambda}}\alpha$ является модульным уровнем ортогональных матриц, построенных на основе (core) матрицы Адамара. Едва ли две тесно связанные матрицы стоит рассматривать как самостоятельные объекты. Это две проекции гиперобъекта, который мы склонны рассматривать как две ортогональные по столбцам матрицы смежных порядков, но они не независимы, а сосуществуют друг с другом согласно тождеству инвариантов.

Как видно, матрица Адамара, не меняя внутренней сути, не является уже целочисленной матрицей, поскольку значение b иррациональное. Все, что будет далее говориться об итерационном процессе, дающем такую матрицу точно, независимо от значений ошибок вычислений, дает нам возможность забыть на время задачу Диофанта — можно находить иррациональную проекцию, а не целочисленную.

Однако и это еще не все, ведь ровно половину первой строки (или столбца) матрицы занимают элементы со значением 1, кроме первого элемента. Это означает, что эквивалентными операциями перестановок можно произвести разделение строк и столбцов матрицы на начинающиеся с $-b$, а потом с $a = 1$. Для удобства будем называть ее второй нормальной формой матрицы Адамара, или нормальной формой ее основы (рис. 3).

Чтобы далее не путаться, основу порядка $4t - 1$ будем называть матрицей Мерсенна $\mathbf{M}$, а основу порядка $4t - 2$ будем называть матрицей Эйлера $\mathbf{E}$ [12]. Операцию перехода от одной матрицы к другой в силу ее важности назовем «метаморфозой» гиперобъекта.

После *отделения* бинарной каймы и ортогонализации вторичным изменением уровня $b = \frac{\lambda}{\lambda + \sqrt{2\lambda}}\alpha$ останется разделенная на четыре части матрица блочных видов: знакосимметричного $\mathbf{E} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C}^{\mathrm{T}} & -\mathbf{D}^{\mathrm{T}} \end{pmatrix}$ или знакококосимметричного $\mathbf{E} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C}^{\mathrm{T}} & \mathbf{D}^{\mathrm{T}} \end{pmatrix}$. Они сводимы к версии с $\mathbf{A} = \mathbf{D}$, $\mathbf{B} = \mathbf{C}$ [12] размера $m - 1$.



■ ***Рис. 3.*** Портрет нормальной формы основы матрицы Адамара

■ ***Fig. 3.*** Portrait of normal form of Hadamard matrix core

Метаморфоза заключается в эквивалентных преобразованиях и перестановках строк и столбцов для выделения блоков, а также в инверсии их знаков и адаптации модуля уровня b. Указанные преобразования позволяют сохранить ортогональность строк и столбцов усекаемой или расширяемой матрицы. Матрица Мерсенна отличается от модульно одноуровневой матрицы Адамара тем, что имеет два уровня 1 и $-b$. Матрица Эйлера в экономном (каноническом) ее представлении является четырехуровневой и двухблочной при $\mathbf{A} = \mathbf{D}$ и $\mathbf{B} = \mathbf{C}$. Благодаря инверсии знака одного блока она содержит элементы $\{1, -1, b, -b\}$.

По отношению к матрице Адамара это менее «плоские» конструкции. Инверсия знака при $-b$ у них не порождает единицу, поэтому преобразование, благодаря которому матрица Адамара может рассматриваться как целочисленная матрица, для двух остальных граней гиперобъекта не проходит. Это объекты иррациональные, которые следует находить итерациями.

Ничто не мешает не удалять, а добавлять кайму к матрице. Матрица Ферма $\mathbf{F}$ может трактоваться как четвертая проекция гиперобъекта, которая требует увеличения числа уровней до трех: кроме единицы и $-b$ появляется уровень элементов каймы s [12]. Однако сделать это просто можно только по отношению к регулярным матрицам Адамара, которые характеризуются дополнительным инвариантом — суммы строк и столбцов одинаковы. К выделенным в $n + 1$ порядкам 5, 17, 256 и т. п. относятся числа Ферма, что привлекает к ним особое внимание.

Можно расширяться и дальше, но это будут многоуровневые матрицы, которые менее универсальны и менее интересны, поскольку состав стабилизирующих их свойства инвариантов исчерпывается описанными выше.

Изменение определения матрицы. В основу определения матриц с элементами 1 и $-b$ можно положить функцию уровня: для матриц Адамара $b = 1$, для матриц Мерсенна $b = \frac{\lambda}{\lambda + \sqrt{\lambda}}\alpha$ и для матриц Эйлера $b = \frac{\lambda}{\lambda + \sqrt{2\lambda}}\alpha$, — связав столбцы условием ортогональности. Но уравнение ортогональности не связано непосредственно с экстремальными свойствами этих матриц, и желательно определить их через итерационный алгоритм оптимизации детерминанта.

Алгоритм Прокруста

Детерминант любой ортогональной в смысле $\underline{\mathbf{A}}^{\mathrm{T}}\underline{\mathbf{A}} = \mathbf{I}$ матрицы порядка n равен единице. Это является удобной точкой отсчета. Разделив ее на максимальный по абсолютному значению элемент μ матрицы $\underline{\mathbf{A}}$, получаем квазиортогональную матрицу $\mathbf{A}$ такую, что $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \omega\mathbf{I}$, где $\omega \le 1$ — некоторый весовой коэффициент $\omega = 1/\mu^2$ [12].

Теорема. Чем меньше μ, тем выше детерминант $\det(\mathbf{A}) = \omega^{n/2} = 1/\mu^n$ матрицы, приведенной к форме с единицей в качестве максимального элемента.

Эта теорема напрямую следует из выражения для ω и элементарным следствием дает хорошо известную в теории матриц Адамара $\mathbf{H}^{\mathrm{T}}\mathbf{H} = n\mathbf{I}$ границу сверху $\det(\mathbf{A}) \le n^{n/2}$, достижимую *только* на порядках 1, 2 и кратных четырем. На прочих порядках приходится увеличивать число модульных уровней. При этом матрицы Мерсенна и Эйлера, будучи образованными от столь мощной по детерминанту основы, не утрачивают экстремальных свойств, однако приобретают специфику.

Приступим к описанию алгоритма их нахождения, целиком следующего из их определения как квазиортогональных матриц, характеризуемых максимальным элементом μ матрицы $\underline{\mathbf{A}}$.

Название алгоритма напрямую связано с мифом, согласно которому тиран Прокруст, накормив гостей, укладывал их на кровать. Выступающие за пределы кровати ноги он отрубал, а короткие — вытягивал, стремясь придать им эстетичные формы. Вкратце изложенное соответствует описанию алгоритма оптимизации детерминанта, подаренному античной традицией доводить все до совершенства.

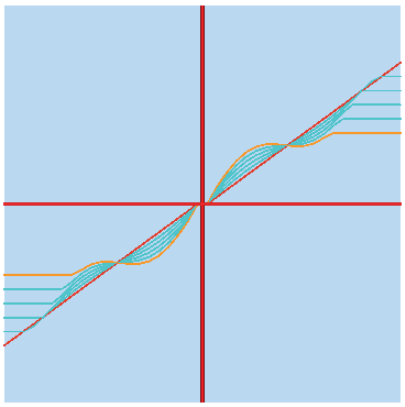
У квазиортогональной матрицы максимума детерминанта $\mathbf{A}$ максимальный по абсолютной величине элемент μ минимален. Это типичная минимаксная задача. Алгоритм Прокруста для матриц сводится к усечению максимального элемента. Первую фазу алгоритма назовем сжатием. При слишком примитивном прокрустовом сжатии матрица теряет ортогональность, но ее несложно восстановить известной процедурой Грама — Шмидта [18], которая вызывает растяжение. Возникает двухэтапный процесс, сжатие плюс растяжение, который итерациями действительно ведет матрицу к оптимуму.

Основа алгоритма Прокруста — нелинейный блок насыщения. Как мы его настроим, так он и будет работать. В простейшем случае это только насыщение, но можно добавить и вытягивание слишком малых элементов — ограничивать большие элементы и увеличивать малые. Уровень насыщения верхнего порога $p \le 1$ и искажение уровней малых элементов можно регулировать изменением профиля кривой, делая порог и искажения сильнее в начале итераций (рис. 4).

Повысить качество алгоритма можно еще несколькими способами, например перестановкой наиболее изменяемых столбцов на первое место или выбором начальной матрицы. Однако это уже второстепенные детали численного метода [19, 20]. Для двухуровневых матриц 1, $-b$ малых элементов (в итоговом решении) не бывает, но и задирать слишком сильно насыщение (снизу) до значения b нельзя, поскольку важна возможность изменять знаки элементов матрицы в процессе итераций. Это сказывается на эффективности поиска локальных экстремумов.

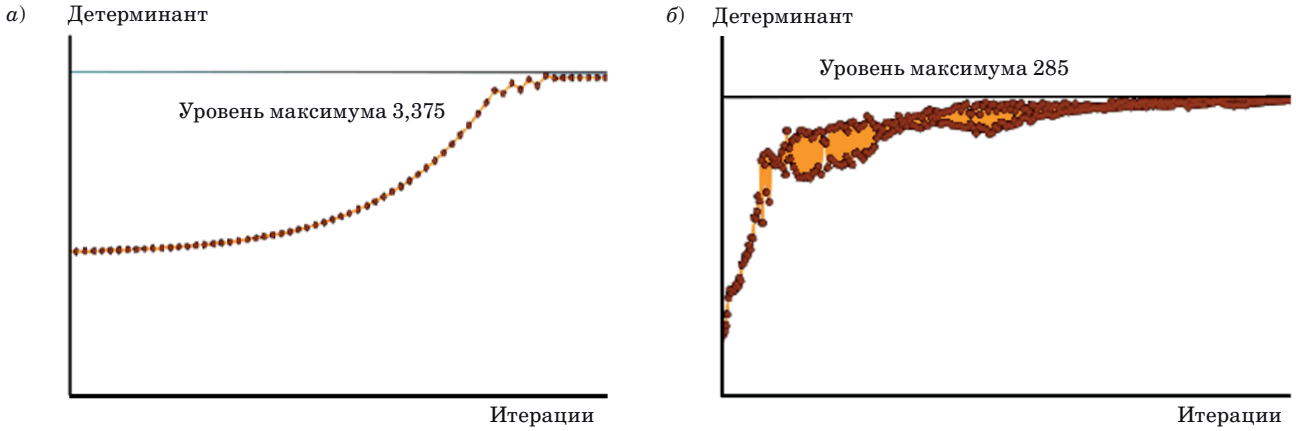
Матрица Адамара является естественным продуктом алгоритма Прокруста. У нее все элементы одинаковы по абсолютной величине, так что они определенно побывали в «прокрустовой кровати».

Сам по себе нелинейный блок статический, но в итерациях с порогом p, постепенно стремящим-



■ ***Рис. 4.*** Диаграмма пропускания нелинейного блока насыщения

■ ***Fig. 4.*** Transmission diagram of a nonlinear saturation block



■ ***Рис. 5.*** Кривые роста детерминантов матриц порядков 3 (*а*) и 7 (*б*)
■ ***Fig. 5.*** Curves of determinant growth for matrices of 3rd (*а*) and 7th (*б*) orders

ся к максимуму (к единице) — это динамическая система [19], хорошо известная в теории автоматического управления. Например, таков контур астатического регулирования системы с интегратором в цепи обратной связи по ошибке регулирования. Точка статического равновесия (неподвижная точка) приходится на искомую матрицу.

Ортогональная матрица с изменениями элементов, ортогонализацией и нормализацией столбцов вынужденно движется к экстремуму. Приложение теоремы о неподвижной точке отражения ограничено здесь наличием локальных экстремумов, нежелательных препятствий на порядках, кратных четырем, но результативных для матриц нечетных порядков. Матрица нечетного порядка не может достичь успокоения в форме с одинаковыми уровнями элементов — ее попросту нет. Элементы имеют два неравных значения или более двух при большем детерминанте. Однако это мало что меняет в работе алгоритма (рис. 5, *а* и *б*).

Условия останова алгоритма Прокруста

Для понимания условий останова алгоритма Прокруста следует учесть, что это алгоритм поиска не элементов, а орнамента. Как и в условиях теоремы Тарского [1], нарушим целочисленные значения элементов матрицы Адамара в любую сторону — к нулю или к бесконечности, не меняя знака. Значит ли это, что мы утратили эту самую матрицу Адамара? Судя по инвариантам $k = n/2$ и $\lambda = n/4$, — нет.

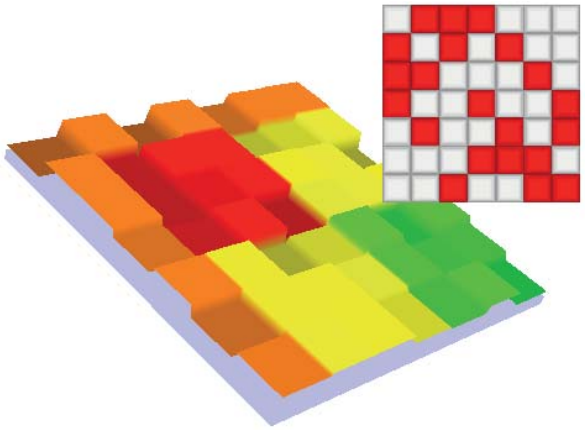
Отсюда следуют два вывода. Во-первых, найти исходную матрицу Адамара не составит труда. Для этого достаточно подтянуть элементы к значениям 1 и –1. Во-вторых, если некоторый итерационный процесс оптимизации детерминанта со сколь угодно большими ошибками выдаст такую матрицу до своего завершения, матрица Адамара будет найдена. Это процесс грубый, терпимый к большим неточностям. Мы ищем орнамент, а не матрицу. Инварианты у нее орнаментальные, описывающие узор, т. е. некий геометрически дискретный объект.

Пример. Матрицы Мерсенна являются двухуровневыми, с элементами 1 и $-b$. Допустим, в процессе итераций алгоритма Прокруста получается первая из приведенных ниже матриц порядка 7:

$$\begin{pmatrix} b & 1 & 1 & b & 1 & 1 & b \\ 1 & b & b & 1 & -1 & -1 & -b \\ 1 & b & -1 & -b & -1 & b & 1 \\ b & 1 & -b & -1 & 1 & -b & -1 \\ 1 & -1 & -1 & 1 & b & b & -b \\ b & -b & 1 & -1 & -b & 1 & -1 \\ 1 & -1 & b & -b & b & -1 & 1 \end{pmatrix};$$

$$\begin{pmatrix} -b & 1 & 1 & -b & 1 & 1 & -b \\ 1 & -b & -b & 1 & 1 & 1 & -b \\ 1 & -b & 1 & -b & 1 & -b & 1 \\ -b & 1 & -b & 1 & 1 & -b & 1 \\ 1 & 1 & 1 & 1 & -b & -b & -b \\ -b & -b & 1 & 1 & -b & 1 & 1 \\ 1 & 1 & -b & -b & -b & 1 & 1 \end{pmatrix}.$$

Уровень достигнут, достигнуты и инварианты: по $k = 4$ элемента 1 с положительным или отрицательным знаком в каждой строке и столбце и по $\lambda = 2$ соседствующих таких же элемента — в любой паре строк или столбцов. Так как инварианты соблюдены, мы можем не заниматься эквивалентными преобразованиями и заме-



■ ***Рис. 6.*** Матрица итога итераций и она же — решение
■ ***Fig. 6.*** Final iteration matrix and it is as the solution

нить все элементы ±1 на 1 и $\pm b$ на $-b$. При этом безразлично, как считать такие инварианты. Можно ориентироваться на знаки элементов, а можно на модули уровней (чего нет у матриц Адамара). Это означает, что останов алгоритма можно выполнить, не дожидаясь схождения величин элементов к заведомо недостижимым точно иррациональным значениям, которые мы знаем заранее. Можно селектировать их величины по порогу $(1 + b)/2$ на большие и малые, а по достижении нужных инвариантов привести итог к нормальному виду с бинарной каймой (рис. 6).

Функции поиска четко разделены. За вычисление уровня отвечает алгоритм вычисления корня квадратного с его более чем двухтысячелетней историей. Орнамент, а не матрицу с иррациональными элементами ищет алгоритм Прокруста. Он тянет матрицы к глобальному или к локальным экстремумам.

Поиск локальных и условных экстремумов

Оптимальные по детерминанту матрицы отличаются характером своих экстремумов, это могут быть:

— глобальный (абсолютный) экстремум, как у матриц Адамара;

— локальный экстремум, отличающийся от глобального меньшим значением максимума;

— условный экстремум седловых точек, когда свобода вариации аргумента ограничена условием.

Задачи поиска корней уравнений дистанцируются от задач оптимизации тем, что положение корней не обязательно увязывать с максимумом какой-либо функции. Итерационный алгоритм рассматривают как систематически применяемое правило переноса (отображения) все новых и новых точек приближения к корню, а сам корень считается найденным, если находится в «неподвижной точке» такого отображения. Если точка только одна, схождение к ней обеспечивается из любых начальных условий.

Алгоритм Прокруста имеет регулируемый шаг по величине насыщения. При малом шаге на его итерациях не происходит изменение знаков элементов. Алгоритм ищет условный, а не абсолютный или локальный экстремум. Это удобно для поиска матриц в седловых точках, например матриц Ферма, когда ненужное направление изменения орнамента отсекается знаками. Орнамент таких матриц плавным изменением уровней (с изменением знаков) можно перевести в орнамент матриц глобального экстремума, уравнения перехода рассмотрим ниже.

Связь иррационального и целочисленного инвариантов

Матриц Адамара, а значит, и матриц Мерсенна бесконечно много. Согласно теореме Тарского о бесконечном числе математических объектов выносится определенное суждение на основании конечного числа манипуляций.

Допустим, оптимизацией детерминанта найдены несколько ортогональных в смысле $\mathbf{M}^{\mathrm{T}}\mathbf{M} = \omega\mathbf{I}$ матриц с элементами 1, $-b$ разных порядков $m = n - 1$, где $\omega \le 1$ – некоторый весовой коэффициент. Для того чтобы идентифицировать функцию $b = b(m) = x_1(m)b^2 + x_2(m)b + x_3(m) = 0$, много матриц не потребуется. Параметры $x_1(m)$, $x_2(m)$, $x_3(m)$ — заранее неизвестные линейные рациональные функции от порядка m.

Они линейны, поскольку сводятся к подсчету числа элементов b^2, b^1, b^0 в скалярных взаимных произведениях столбцов $\mathbf{M}$, причем точное значение b не требуется. Уравнение $(m - 3)b^2 - 2(m + 1)b + (m + 1) = 0$ отвечает характеристическому уравнению матриц Мерсенна $(\lambda - 1)b^2 - 2\lambda ab + \lambda a^2 = 0$ при $a = 1$. Учитывая $\lambda = (m + 1)/4$, можно найти $b = \dfrac{\lambda}{\lambda + \sqrt{\lambda}} = \dfrac{m+1}{m + 2\sqrt{m+1} + 1}$, идентифицируя этот класс всего по нескольким его представителям.

Идентификация параметров на основании данных эксперимента давно используется в адаптивных системах. Для описания динамической системы, выполняющей маневр на отрезке времени любой протяженности, нет необходимости анализировать его весь. Достаточно по конечному числу точек найти параметры модели системы в форме дифференциального уравнения. Это касается установления эквива-

лентности между законами Ньютона и законами Кеплера описания движения планет. Никто не сверяет всю орбиту поточечно с решением дифференциального уравнения тяготения.

Если изучается класс бесконечного числа ортогональных матриц по конечному числу их представителей, обнаруженных алгоритмом Прокруста, можно установить аналитически точный вид функции уровня для всех них [20]. Далее важно то, что функция $b = b(m)$ монотонна, не содержит разрывов при всех значениях m, свидетельствующих об отсутствии решения. Уровень b помогает идентифицировать важнейший орнаментальный инвариант $\lambda = b^2/(1-b)^2$ через квадрат отношения отрезков, на которые $b < 1$ делит уровень $a = 1$. Следовательно, если из опыта экспериментов с алгоритмом Прокруста удалось установить значение b (даже не очень точно), можно оценить перспективу нахождения λ, т. е. того самого инварианта, который при поиске матриц Адамара перебором до завершения перебора аналитически установить невозможно. Это замечательное утверждение связывает два мира (дискретный и континуальный): целочисленный орнаментальный инвариант и иррациональный уровень.

Замкнутые орнаменты

Греческий математик и философ Прокл Диадох, сторонник чрезмерного лаконизма, разрабатывал концепцию числа как моста между двумя началами — умом и чувственным восприятием. Вспомнили мы о нем вот почему.

Рассмотрим матрицу Эйлера, которая в экономном своем варианте с блоками $\mathbf{A} = \mathbf{D}$ и $\mathbf{B} = \mathbf{C}$ может строиться на основе всего одной матрицы Мерсенна $\mathbf{M}$. Поскольку матрицы Мерсенна регулярны (суммы столбцов и строк совпадают), составная матрица Эйлера порядка $2m$ допускает назначение двух разных значений плеч $\mathbf{A} = \mathbf{M}(b_1)$, $\mathbf{B} = \mathbf{M}(b_2)$, причем можно выделить уравновешенное

$$b_1 = b_2 = \frac{\lambda}{\lambda + \sqrt{\lambda}} = \frac{m+1}{m + 2\sqrt{m+1} + 1}$$

решение $\mathbf{M} = \mathbf{A} = \mathbf{B} = \mathbf{C} = \mathbf{D}$ и экстремальное по детерминанту решение при $b_1 = 1$:

$$b_2 = \frac{\lambda - 1}{\lambda + \sqrt{2\lambda - 1}} = \frac{m-3}{m + 2\sqrt{2m+2} + 1}.$$

Первое решение назовем матрицей Эйлера, а второе — матрицей Прокла; усилив преобладание единичных по абсолютному значению элементов диагонального блока, поднимем значение детерминанта. Как и качели, они имеют два положения — уравновешенное и когда один край качелей вверху. В таких качелях важно то, что орнамент у обеих матриц общий. Они принадлежат континуальному множеству матриц, которые плавным координированным изменением параметров плеч переходят, без изменения рисунка из знаков, в состояние с бóльшим детерминантом.

Это означает, что алгоритм Прокруста, если его не останавливать, остановится на второй матрице, но находит-то он их обе. Таким образом, найден орнамент бесконечного множества матриц, одна из которых оставляет ему возможность быть в виде большого детерминанта.

Определение: Орнамент называется замкнутым, если любое малое изменение параметров узора приводит к понижению детерминанта.

Примеры замкнутых орнаментов: матрицы Адамара, Белевича, Мерсенна, Прокла. Стартовая матрица Прокла при $m = 3$ эквивалентна конференц-матрице с нулем на диагонали, что согласуется с ролью этой трехуровневой матрицы, замещающей матрицы Адамара на четных не достижимых ими порядков.

Примеры незамкнутых орнаментов: матрицы Эйлера как основа матриц Мерсенна и матрицы Одина как основа конференц-матриц [21]. Это параметрически изменяемые узоры. Алгоритм Прокруста к их поиску применим с рядом оговорок о поиске замкнутых орнаментов, к инвариантам которых имеют отношение нестабильные матрицы.

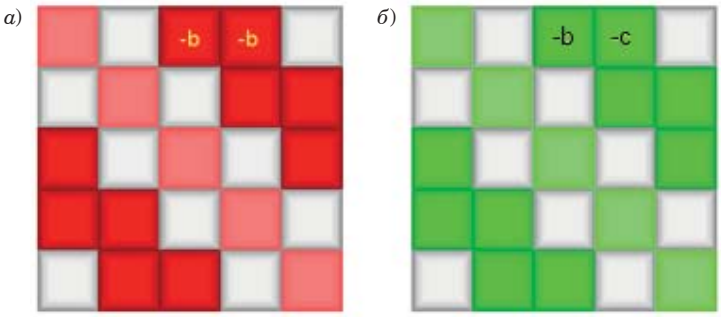
Устойчивость орнаментам первых трех матриц Адамара, Белевича и Мерсенна придает явно прослеживаемая в их конструкции кайма. Матрица Прокла зажата иначе — фиксированием половины ее параметров в –1, остальные находятся в блоках вне диагонали. Поскольку орнамент матрицы Мерсенна замкнут, то замкнут и орнамент матрицы Прокла.

Различие матриц Адамара и конференц-матриц

В тридцатых годах Пэли [10] обратил внимание, что если инвариант $k = n/2$ таков, что $p = k - 1$ разложимо на сумму двух квадратов, то симметричные четырехблочные матрицы Адамара можно строить на основе одного симметричного ортогонального блока с нулевой диагональю, варьируя знак диагонали:

$$\begin{pmatrix} \mathbf{C}+\mathbf{I} & \mathbf{C}-\mathbf{I} \\ \mathbf{C}-\mathbf{I} & -(\mathbf{C}+\mathbf{I}) \end{pmatrix}.$$

Особенно легко эта конструкция строится для простых p (или степеней простого числа), когда первая строка циклического блока $\mathbf{C}$ строится несложными алгоритмами конечных полей $GF(p)$.



■ ***Рис. 7.*** Портреты матриц Одина (*а*) и Модина (*б*)
■ ***Fig. 7.*** Portraits of Odin (*а*) and Modin (*б*) matrices

Позднее В. Белевич предложил рассматривать этот блок в качестве самостоятельной ортогональной матрицы, назвав ее конференц-матрицей в силу приложения ее к задачам телефонии. Конференц-матрица во многом напоминает матрицу Адамара и считается ее прямым продолжением на недостижимые матрицами Адамара четные порядки. После нормализации количество положительных и отрицательных элементов (вне каймы) симметричной матрицы Белевича совпадает в строках и столбцах.

Возможны построения аналогов матриц Мерсенна и Эйлера, названные матрицами Одина и Тени [20].

Ортогонализация матрицы Одина порядка $m = 4t - 3$ (и Тени порядка $m - 1$) невозможна без отрыва диагонального элемента от нуля $d = \frac{1}{1+\sqrt{m}}$, уровень $b = 1 - 2d$ компенсирует этот отрыв. Стартовая матрица Одина пятого порядка (рис. 7, *а*) не является матрицей локального максимума детерминанта. Доказать это несложно, разделив пару элементов первой строки $-b$ на $-b$ и $-c$. Чтобы не путаться, назовем вторую матрицу матрицей Модина (рис. 7, *б*).

По орнаменту обе приведенные матрицы являются матрицами «под-Белевичами». Орнаменты матриц Одина не замкнуты, при $m = 5$ диагональ $d = \frac{1}{1+\sqrt{m}} + \delta$ меняется за счет добавки δ, изменяющиеся в противоположные стороны параметры $b = \frac{1}{d+1} - c$, $c = \frac{1+\sqrt{8d^3+16d^2+4d-3}}{2(d+1)}$ играют ту же роль, что и плечи матриц Прокла.

Иными словами, детерминант матрицы Модина можно повысить в сравнении с детерминантом матрицы Одина, разорвав тождество $b = c$ сколь угодно малым изменением параметра d. Это означает, что матрицы Одина — седловые точки, а не точки локального оптимума детерминанта. У матриц Одина есть возможность при помощи варьируемой диагонали, позволяющей матрице измениться, повышать детерминант.

Это показывает существенное отличие матриц Мерсенна от матриц Одина. Элементы диагонали насыщены в единицы, так что «кредита доверия» для повышения детерминанта нет. Значение d максимально высоко.

Границы детерминантов критских матриц

Критские матрицы — это матрицы семейства Адамара $\mathbf{A}^T\mathbf{A} = \omega\mathbf{I}$ с небольшим числом уровней. Число уровней критских матиц нечетных порядков растет линейно (почти линейно) до критического порядка 13, где структура матриц резко усложняется возникновением почти хаотических матриц [12]. Детерминанты $\det(\mathbf{A}) = \omega^{n/2}$ и веса ω первых экстремальных по детерминанту матриц хорошо известны, это позволяет оценить то, насколько мало экстремумы локально оптимальных матриц Мерсенна уступают глобальным экстремумам при сравнении. Для этого используем тройки значений (порядок матрицы; вес матрицы Мерсенна; вес матрицы глобального экстремума): (3; 2,25; 2,25), (7; 5,03; 5,08), (11; 8,01; 8,5), (15; 11,11; 11,99).

Первая из этих матриц, приведенная к единице по максимальным значениям элементов, хорошо известная ортогональная матрица поворота на три угла Эйлера $\alpha = -\arcsin(2/3)$, $\beta = \gamma = \pi - \arctan(2)$:

$$\begin{pmatrix} \cos(\alpha)\cos(\gamma) & \cos(\alpha)\sin(\gamma) & -\sin(\alpha) \\ \sin(\alpha)\sin(\beta)\cos(\gamma) - \cos(\beta)\sin(\gamma) & \sin(\alpha)\sin(\beta)\sin(\gamma) + \cos(\beta)\cos(\gamma) & \cos(\alpha)\sin(\beta) \\ \sin(\alpha)\cos(\beta)\cos(\gamma) + \sin(\beta)\sin(\gamma) & \sin(\alpha)\cos(\beta)\sin(\gamma) - \sin(\beta)\cos(\gamma) & \cos(\alpha)\cos(\beta) \end{pmatrix}.$$

Функцию трех переменных невозможно представить в виде обычного графика поверхности с пиком экстремума, но можно показать детерминант диаметром шарика. Максимум в центре соответствует повороту, порождающему изображенную матрицу Мерсенна. Вариация углов в окрестности этой точки может только уменьшить сферы вокруг центральной точки (рис. 8, *а*).

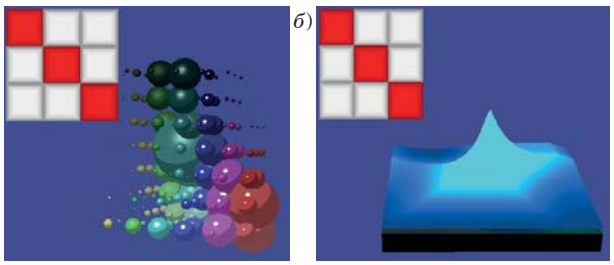
Традиционный рисунок максимума экстремума тоже можно увидеть, задавая развертку двумя углами, когда третий угол выбирается из условия максимума (рис. 8, *б*). На множестве матриц седьмого и более высоких порядков матрицы глобального **A** и локального **M** экстремумов не совпадают. Профиль экстремальной кривой можно изучать, ортогонализуя осредненную матрицу $\mathbf{A}t + \mathbf{M}(1-t)$, где t — вещественный параметр, регулирующий уход от матрицы локального в сторону глобального детерминанта (рис. 9).

Хороший рисунок в состоянии многое показать. Если детерминант любой критской матрицы $\omega^{n/2}$ поделить на оценку Адамара сверху $n^{n/2}$, то такой относительный детерминант не будет превышать единицу. Приведенный детерминант — это корень n-й степени из относительного детерминанта $\sqrt{\frac{\omega}{n}}$. Диаграмма с приведенными детерминантами экстремальных матриц, нанесенная с шагом 4, показывает их прижатыми к верхней границе, причем сжатие нарастает. Максимальные по возможному детерминанту матрицы Адамара находятся на верхней единичной полочке (рис. 10).

Порядки $4t$, $4t-1$, $4t-2$, $4t-3$, определяющие основные типы матриц семейства, отвечают еще трем границам. Более жесткую границу сверху дают оценки детерминантов конференц-матриц дополнительных четных порядков. На рис. 10 представлена граница детерминантов матриц Мерсенна (чуть ниже можно добавить границу для матриц Одина, основ матриц Адамара и конференц-матриц). Точки между ними являются значениями норм многоуровневых матриц максимума детерминанта. Портреты матриц и гистограммы их уровней представлены на рис. 11.

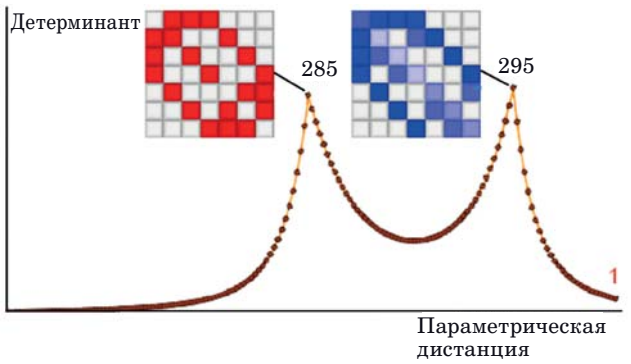
На порядке 22, где нет конференц-матрицы, есть шестиуровневая бициклическая и взвешенная матрицы. Детерминант последней дает точку, не дотягивающую до оптимистичной границы детерминантов конференц-матриц. Экстремальные многоуровневые матрицы приспосабливаются к порядку увеличением числа уровней и неограниченным нарастанием сложности их структур.

Как видно, граница детерминантов матриц Мерсенна — это наиболее удачное и относительно несложное приближение снизу. Так как матрицы Адамара и матрицы Мерсенна представляют объект с одним и тем же орнаментом (пренебрегая каймой), то получается, что критские матрицы зажаты между этими двумя проекциями гиперобъекта.



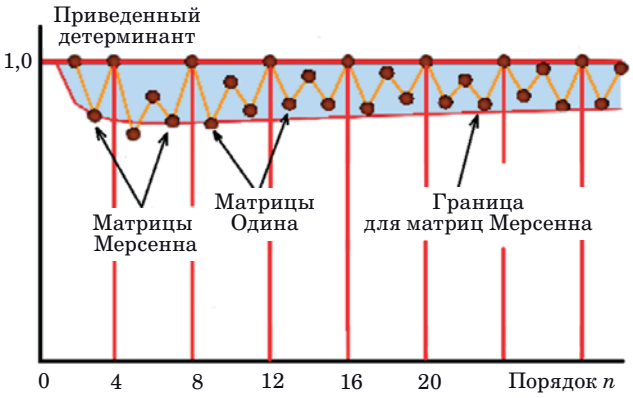
■ ***Рис. 8.*** Представление максимума детерминанта в 3D (*а*) и 2D (*б*) вариантах

■ ***Fig. 8.*** Maximum determinant in 3D (*а*) and 2D (*б*) versions



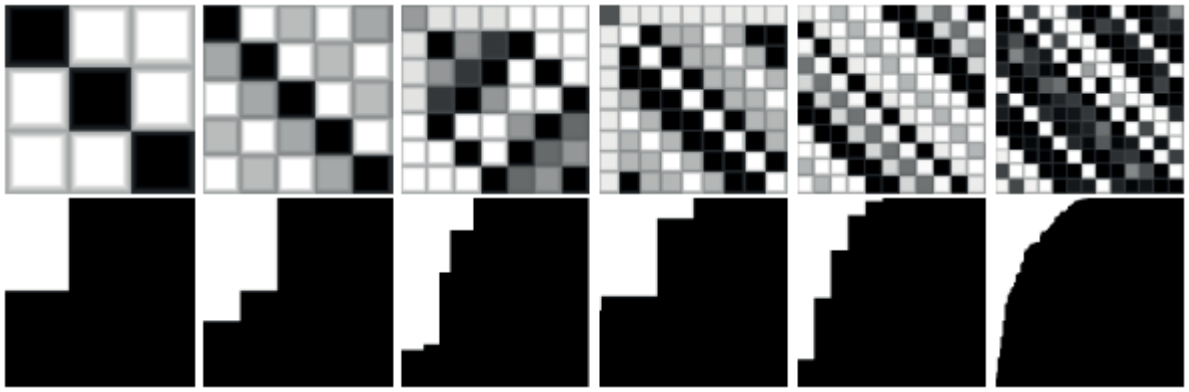
■ ***Рис. 9.*** Локальный и глобальный экстремумы детерминанта

■ ***Fig. 9.*** Local and global extremes of determinant



■ ***Рис. 10.*** Диаграмма максимальных детерминантов

■ ***Fig. 10.*** Maximum determinants diagram



■ ***Рис. 11.*** Портреты и диаграммы уровней элементов оптимальных матриц
■ ***Fig. 11.*** Portraits and element level diagrams of optimal matrices

Связь гиперобъекта с теоремой Дженнифер Себерри

Матрицы Адамара основной последовательности порядков Сильвестра отвечают числам Мерсенна $2^s - 1$, дающим размеры вложенных в них матриц, которые существуют вне зависимости от того, простые это числа, степени простых чисел или составные числа [22, 23].

Дженнифер Себерри доказала следующее [2].

Теорема. Минимальное расстояние $s > 0$ между нечетными простыми числами p и $2^s p - 1$ (размеров потенциальных основ конференц-матриц **C** или матриц Адамара конструкции Пэли) конечно и не превышает удвоенного логарифма от p.

Два — это максимальный коэффициент, его можно снижать, добавляя к логарифму константу смещения. Это называется оценкой асимптотики существования матриц Адамара, поскольку последующими удвоениями порядка можно получить неограниченное количество этих матриц. Для составных нечетных чисел оценки не меняются, так как матрицу Адамара можно получить кронекеровым произведением [24] матриц меньшего порядка, построенных для сомножителей по той же схеме.

Поскольку числа $2^s p - 1$ являются обобщением чисел Мерсенна, естественно предположить, что от расстояния s зависит не столько само существование матриц Адамара, сколько существование их в специфической форме Пэли. Кроме матриц, которые строятся на квадратичных вычетах, их можно найти в форме двояко-симметричных бициклов с парной каймой — описанных выше матриц Эйлера. За потерю возможности использовать поле есть чем платить: из симметричной и кососимметричной клеток универсального бицикла гарантированно остается одна.

Симметрии матриц принято описывать дициклическими группами или, как у коциклических матриц, таблицей умножения группы после дихотомии элементов группы на описывающие 1 или –1. Эта ясная точка зрения не разрешает вопрос существования, поскольку дихотомия неоднозначна и столь же трудна, как попытка найти матрицу Адамара прямым перебором. Тем не менее теорема Дж. Себерри об асимптотическом существовании матриц Адамара свидетельствует о том, что для выделяемого ею анклава матриц Адамара составной характер размера их основ не имеет никакого значения.

Заключение

В противопоставлении матриц Адамара и неортогональных матриц максимума детерминанта содержится доля противоречия. Выходит, что кроме экстремальных по детерминанту матриц Адамара есть еще какие-то экстремальные матрицы. Все это плохо согласуется между собой и будит желание разобраться глубже в проблеме. Матрицы Мерсенна, которые оказываются двойниками матриц Адамара, не утрачивают свойство быть экстремальными, и их находят алгоритмы поиска неподвижной точки отображения.

Экстремальные матрицы на любом разрешенном для них порядке идентифицируемы алгоритмом Прокруста — нет принципиальных препятствий найти экстремум, даже если этот экстремум локален и приблизиться к нему поиском области притяжения сложно. Алгоритм находит матрицы стартовых порядков различной природы, которые теоретически были открыты в разное время разными методами. Задача поиска раскладывается на субалгоритм вычисления иррационального уровня и субалгоритм поиска орнамента. Первый — это алгоритм Ньютона

для скалярного случая, а второй реализует поиск орнамента.

Благодаря модификации профиля нелинейного блока насыщения алгоритм может искать как локальные экстремумы, так и седловые точки — условные экстремумы, поскольку при малой амплитуде изменений, вносимых отображением, знаки элементов матрицы не меняются. Удержание знаков орнамента стабилизирует итерации на такой «невыгодной» точке, как седловая, из которой есть путь по увеличению значения детерминанта. Так, например, алгоритмом Прокруста были найдены матрицы Ферма, ошибочно классифицируемые изначально как матрицы локального экстремума именно потому, что алгоритм их ищет устойчиво и не теряет при отклонениях их параметров. Аналогичные «грубые» алгоритмы сыграли большую роль при поиске нулей дзета-функции, находить точные значения которой сложно.

Критские матрицы тесно связаны с числовой системой; всем особым числам: золотому сечению, числам Ферма, Мерсенна, числам-близнецам — отвечают соответствующие матрицы. На глубокое различие нечетных порядков $4t - 1$ и $4t - 3$ впервые обратили внимание Ферма и Эйлер в рамках так называемой Рождественской теоремы Ферма. Как видно, у матриц тоже есть это различие, выражающееся в том, что все матрицы Мерсенна порядков $4t - 1$ существуют без исключения, на что указывает наличие у них идентифицируемого экстремума.

Если фиксировать элементы уровнями, мостика между матрицей Мерсенна и матрицей глобального экстремума нет при том, что они близки по орнаменту. Поэтому, строя график изменения детерминанта матриц порядков $4t - 1$, нам приходится не просто усреднять, а еще и ортогонализировать промежуточные матрицы. Ничего подобного нет у седловых точек матриц порядков $4t - 3$, поскольку у них путь наверх заведомо облегчен структурой матрицы, будь это матрица Одина или матрица Ферма. Математический аппарат конечных полей и групп начала прошлого века ускоряет процесс поиска матриц. Эти инструменты, превосходные в своей эффективности на некоторых выделенных простотой порядках, уступают в мощности вычислений в поле вещественных чисел. Уже в скалярном случае полноценное рассмотрение задач алгебраической геометрии невозможно без привлечения итерационных процедур, с помощью которых было сформировано само понятие иррационального числа.

Финансовая поддержка

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

Благодарности

Авторы выражают искреннюю благодарность за многолетнюю помощь и поддержку профессору Д. Джоковичу. За помощь в технической работе с рукописью благодарят Т. В. Балонину.

Литература

1. **Матиясевич Ю. В.** Алгоритм Тарского. *Компьютерные инструменты в образовании*, 2008, № 6, с. 4–14.
2. **Jennifer S., Yamada M.** *Hadamard matrices: Constructions using number theory and linear algebra*. Wiley, 2020. 384 p.
3. **Colbourn C. J., Dinitz J. H.** *Handbook of Combinatorial Designs*. Second ed. Chapman and Hall/CRC, 2007. 967 p.
4. **Silvester J. J.** Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers. *Philosophical Magazine*, 1867, no. 34, pp. 461–475.
5. **Hadamard J.** Résolution d'une question relative aux déterminants. *Bulletin des Sciences Mathématiques*, 1893, vol. 17, pp. 240–246.
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. **Djocovic D. Z., Kotsireas I. S.** Periodic Golay Pairs of Length 72. In: *Algebraic Design Theory and Hadamard Matrices*. C. J. Colbourn (ed). Springer, 2015. Pp. 83–92.
8. **Ito N.** On Hadamard groups IV. *Journal of Algebra*, 2000, no. 234, pp. 651–663.
9. **Schmidt B.** Williamson matrices and a conjecture of Ito's. *Designs, Codes and Cryptography*, 1999, no. 17, pp. 61–68.
10. **Paley R. E. A. C.** On orthogonal matrices. *Journal of Mathematics and Physics*, 1933, vol. 12, pp. 311–320.
11. **Williamson J.** Hadamard's determinant theorem and the sum of four squares. *Duke Math. J.*, 1944, vol. 11, pp. 65–81.
12. **Балонин Н. А., Сергеев М. Б.** *Специальные матрицы: псевдообратные, ортогональные, адамаровы и критские*. СПб.: Политехника, 2019. 196 с. doi:10.25960/7325-1155-0

13. **Сергеев А. М., Куртяник Д. В., Тарашкевичус К. Ф.** Матричный портрет как основа дискретного текстильного орнамента. *Известия высших учебных заведений. Технология легкой промышленности*, 2019, т. 44, № 2, с. 102–107.
14. **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
15. **Гаусс К. Ф.** *Труды по теории чисел*. М.: Изд-во АН СССР, 1959. 978 с.
16. **Liouville J.** Nouveaux théorémes concernant les nombres triangulaires. *Journal de Mathématiques Pures et Appliquées*, 1863, no. 8, pp. 73–84.
17. **Awyzio G., Seberry J.** On Good Matrices and Skew Hadamard Matrices. In: *Algebraic Design Theory and Hadamard Matrices*: Springer Proceedings in Mathematics & Statistics/Ch. Colbourn (eds). Springer, Cham., 2015. Vol. 133. https://doi.org/10.1007/978-3-319-17729-8_2
18. **Pursell L. and Trimble S. Y.** Gram — Schmidt orthogonalization by Gauss elimination. *The American Mathematical Monthly*, 1991, vol. 98, no. 6, pp. 544–549.
19. **Балонин Н. А., Сергеев М. Б.** О значении матриц начального приближения в алгоритме поиска обобщенных взвешенных матриц глобального и локального максимума детерминанта. *Информационно-управляющие системы*, 2015, № 6, с. 2–9. doi:10.15217/issn1684-8853.2015.6.2
20. **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
21. **Belevitch V.** Theory of 2n-terminal networks with application to conference telephony. *Electrical Communication*, 1950, vol. 27, no. 3, pp. 231–244.
22. **Сергеев А. М.** Простые числа и симметрии квазиортогональных циклических матриц Мерсенна. *Математические методы и модели в высокотехнологичном производстве*: тезисы докл. I Междунар. форума, Санкт-Петербург, 10–11 ноября 2021 г. СПб., 2021, с. 14–15.
23. **Сергеев А. М.** Об одном подходе к вычислению квазиортогональных циклических матриц с симметриями как основы кодов. *Телекоммуникации*, 2022, № 9, с. 28–33. doi: 10.31044/1684-2588-2022-0-9-28-33
24. **Van Loan C.** The ubiquitous Kronecker product. *Journal of Computational and Applied Mathematics*, 2000, vol. 123, iss. 1-2, pp. 85–100. doi:10.1016/S0377-0427(00)00393-9

UDC 519.614
doi:10.31799/1684-8853-2023-1-2-16
EDN: KOMNBV

Solvable and unsolvable problems. Using Procrustes analysis algorithm for obtaining a family of Hadamard matrices

N. A. Balonin[a], Dr. Sc., Tech., Professor, orcid.org/0000-0001-7338-4920, korbendfs@mail.ru
J. Seberry[b], Dr. Sc., Tech., Honorary Professor, orcid.org/0000-0002-9558-4293
M. B. Sergeev[a], Dr. Sc., Tech., Professor, orcid.org/0000-0002-3845-9277
[a]Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation
[b]Department of Computing and Information Techology, University of Wollongong, NSW 2522, Australia

Introduction: The development of the Hadamard matrix theory encountered an obstacle caused not so much by the nature of the integer problem as by the artificial limitation of the solution of quadratic equations applying exhaustive search algorithms. Ignoring the direct path and rejecting irrationality led to the opinion that the hypothesis of the existence of Hadamard matrices was unprovable. **Purpose:** To prove the solvability of the Hadamard problem by orthogonal matrices via identifying their stable connection with matrices containing irrational elements. **Results:** We show that irrationality manifests itself in the quadratic norm of the columns of the Hadamard matrix of the second order. We consider the transfer of iterative algorithms for calculating roots to the matrix. To minimize the maximum absolute value element of the orthogonal matrix we propose the Procrustes analysis algorithm. Since Hadamard matrices are determined by invariants of smaller-order matrices embedded in their structure, the algorithm turns out to be a universal basis for finding them together. We consider the hypothesis of the existence of Hadamard matrices in the operational domain of iterative algorithms determined over the field of real numbers that give advantage over the tools in the form of finite fields and groups. **Practical relevance:** Orthogonal sequences obtained from rows (columns) of Hadamard matrices, and high-order Hadamard matrices themselves are of great practical importance for problems of noise-correcting coding, compression, masking and image processing.

Keywords — Hadamard matrices, conference matrices, Cretan matrices, Procrustes analysis algorithm, finite fields, matrix symmetries.

For citation: Balonin N. A., Seberry J., Sergeev M. B. Solvable and unsolvable problems. Using Procrustes analysis algorithm for obtaining a family of Hadamard matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 2–16 (In Russian). doi:10.31799/1684-8853-2023-1-2-16, EDN: KOMNBV

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. Matiyasevich Y. V. Tarski's algorithm. *Komp'yuternye instrumenty v obrazovanii*, 2008, no. 6, pp. 4–14 (In Russian).
2. Jennifer S., Yamada M. *Hadamard matrices: Constructions using number theory and linear algebra*. Wiley, 2020. 384 p.
3. Colbourn C. J., Dinitz J. H. *Handbook of Combinatorial Designs*. Second ed. Chapman and Hall/CRC, 2007. 967 p.
4. Silvester J. J. Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers. *Philosophical Magazine*, 1867, no. 34, pp. 461–475.
5. Hadamard J. Résolution d'une question relative aux déterminants. *Bulletin des Sciences Mathématiques*, 1893, vol. 17, pp. 240–246 (In French).
6. Balonin N. A., Sergeev A. M., Sinitshina 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. Djocovic D. Z., Kotsireas I. S. *Periodic Golay Pairs of Length 72*. In: *Algebraic Design Theory and Hadamard Matrices*. C. J. Colbourn (ed). Springer, 2015. Pp. 83–92.
8. Ito N. On Hadamard groups IV. *Journal of Algebra*, 2000, no. 234, pp. 651–663.
9. Schmidt B. Williamson matrices and a conjecture of Ito's. *Designs, Codes and Cryptography*, 1999, no. 17, pp. 61–68.
10. Paley R. E. A. C. On orthogonal matrices. *Journal of Mathematics and Physics*, 1933, vol. 12, pp. 311–320.
11. Williamson J. Hadamard's determinant theorem and the sum of four squares. *Duke Math. J.*, 1944, vol. 11, pp. 65–81.
12. Balonin N. A., Sergeev M. B. *Special'nye matricy: psevdoobratnye, ortogonal'nye, adamarovy i kritskie* [Special matrices: pseudo-return, orthogonal, Hadamardian and Cretan]. Saint-Petersburg, Politekhnika Publ., 2019. 196 p. (In Russian). doi:10.25960/7325-1155-0
13. Sergeev A. M., Kurtyanik D. V., Tarashkevichus C. A. Matrix portrait as the basis of discrete textile ornament. *Izvestiya vysshih uchebnyh zavedenij. Tekhnologiya legkoj promyshlennosti*, 2019, vol. 44, no. 2, pp. 102–107 (In Russian).
14. Balonin N. A., Sergeev M. B., Seberry J., Sinitshina 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
15. Gauss J. K. *Trudy po teorii chisel* [Proceedings on number theory]. Moscow, AN SSSR Publ., 1959. 978 p. (In Russian).
16. Liouville J. Nouveaux théorémes concernant les nombres triangulaires. *Journal de Mathématiques Pures et Appliquées*, 1863, no. 8, pp. 73–84 (In French).
17. Awyzio G., Seberry J. *On Good Matrices and Skew Hadamard Matrices*. In: *Algebraic Design Theory and Hadamard Matrices*: Springer Proceedings in Mathematics & Statistics. Ch. Colbourn (eds). Springer, Cham., 2015. Vol. 133. https://doi.org/10.1007/978-3-319-17729-8_2
18. Pursell L. and Trimble S. Y. Gram — Schmidt orthogonalization by Gauss elimination. *The American Mathematical Monthly*, 1991, vol. 98, no. 6, pp. 544–549.
19. Balonin N. A., Sergeev M. B. Initial approximation matrices in search for generalized weighted matrices of global or local maximum determinant. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 6, pp. 2–9 (In Russian). doi:10.15217/issn1684-8853.2015.6.2
20. Balonin N. A., Sergeev A. M. 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
21. Belevitch V. Theory of 2n-terminal networks with application to conference telephony. *Electrical Communication*, 1950, vol. 27, no. 3, pp. 231–244.
22. Sergeev A. M. Prime numbers and symmetries of quasi-orthogonal cyclic Mersenne matrices. *Tezisy dokladov I Mezhdunarodnogo foruma "Matematicheskie metody i modeli v vysokotekhnologichnom proizvodstve"* [Proc. of the I Int. Forum "Mathematical Methods and Models in High-Tech Production"], Saint-Petersburg, 2021, pp. 14–15 (In Russian).
23. Sergeev A. M. On one approach to calculation of quasi-orthogonal cyclic matrices with symmetries as basis of codes. *Telecommunications*, 2022, no. 9, pp. 28–33 (In Russian). doi:10.31044/1684-2588-2022-0-9-28-33
24. Van Loan C. The ubiquitous Kronecker product. *Journal of Computational and Applied Mathematics*, 2000, vol. 123, iss. 1-2, pp. 85–100. doi:10.1016/S0377-0427(00)00393-9

UDC 004.05
doi:10.31799/1684-8853-2023-1-17-28
EDN: ORVZMP

Articles

Advanced metric analysis tool for Java source code

V. V. Burakov[a], *Dr. Sc., Tech., Associate Professor, orcid.org/0000-0002-0158-8681, burakov@compmechlab.com*
A. I. Borovkov[b], *PhD, Tech., Professor, orcid.org/0000-0003-3177-0959*
[a]*CompMechLab LLC, bld. 2a, 21, Gzhatskaya St., 195220, Saint-Petersburg, Russian Federation*
[b]*Peter the Great St. Petersburg Polytechnic University, 29, Politekhnicheskaia St., 195251, Saint-Petersburg, Russian Federation*

Introduction: *Despite considerable efforts made by numerous researchers and developers, procedures for software quality assessment are still not sufficiently formalized and automated.* ***Purpose:*** *To develop a specialized software tool to quantify the structural properties of Java code.* ***Results:*** *We have developed MetricsTree, a software tool that calculates 61 established object-oriented metrics and is one of the largest sets among similar tools. MetricsTree is integrated into the IDE to ensure the fastest possible information delivery, contains unique visualization tools to increase the efficiency of metrics analysis, and implements a metrics profile mechanism to select classes based on a set of metrics values.* ***Practical relevance:*** *As a result of applying MetricsTree to automate quality assurance processes in the development of Peter the Great St.Petersburg Polytechnic University's flagship CML-Bench system (a platform for developing digital twins), the average number of software defects detected by external means decreased by 34% during the year.*

Keywords *— software quality, software metrics, object-oriented metrics, Java metrics, software metrics tool, metrics analysis, software metrics visualization, code smell.*

For citation: Burakov V. V., Borovkov A. I. Advanced metric analysis tool for Java source code. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 17–28. doi:10.31799/1684-8853-2023-1-17-28, EDN: ORVZMP

Introduction

The digital platform for development and application of CML-Bench for digital twins has been created and continues to be maintained by specialists at the STI Centre "New Manufacturing Technologies" at Peter the Great St. Petersburg Polytechnic University (https://cml-bench.ru). Digital twins [1], which our system manages, are families of complex multidisciplinary mathematical and computer models with a high level of adequacy to real materials, objects, structures, machines, devices, technical systems, physico-mechanical, technological and production processes [2, 3]. CML-Bench is a system for numerical design, mathematical modelling and computer-aided engineering control that belongs to the SPDRM (Simulation, Process, Data & Recourses Management) class of systems, providing standardized computations and processing of results, coordination of distributed engineering and design teams, greater transparency of design and engineering processes, optimized development of digital twins through an accumulated knowledge base and intelligent supercomputer management.

CML-Bench has been developed since 2014, including more than 600,000 lines of code; it comprises a set of services whose server side is implemented in Java. CML-Bench is implemented at enterprises from the high-tech industries with stringent requirements to code quality.

As initial steps, we adopted a regulation that we built for quality assurance, describing the relevant procedures and tools. Initially, we used SonarQube (https://www.sonarqube.org) and PMD (https://pmd.github.io) as basic tools.

Problem statement

The importance of the tasks solved by software implemented in all areas of human activity, the possible damage from the exploitation of low-quality software is very high, so it is very important to formalize all components of software quality assessment. A set of specialized mathematical models has been developed to solve this problem [4–6]. One way to formalize the quality assessment process is based on the calculation and interpretation of object-oriented values of source code metrics. Researches on source code metric analysis started in the 80-s, at present several tens of metrics types are used, the correlation of their values with the qualitative state of software is scientifically proved, and there are several tens of software applications, designed for calculation of Java code metrics. We investigated this software as part of selecting a calculation tool to implement in our quality assurance processes. After determining the functionality requirements needed for our processes, we identified weaknesses in the existing metrics calculation software that made implementation difficult. We concluded that the existing products primarily lacked the following important features:

— support for some types of metrics;

— possibility to calculate new metrics with minimum labor intensity;
— rapid data delivery;
— integration into IDE;
— possibility to investigate samples of classes corresponding to a given set of metrics values.

As a solution to these problems, we developed a plugin for the IntelliJ IDEA IDE, which we called MetricsTree.

Practical implementation

MetricsTree is a plugin for IntelliJ IDEA (https://plugins.jetbrains.com/plugin/13959-metricstree) that is compatible with the Ultimate, Community, Educational, and Android Studio editions. Plugin development for IDEA consists in correlating plugin classes with special extension points provided by this IDE to extend functionality. As a result, classes become available to the elements of the IDE from the methods of the plugin; the IDE then determines from the extension points where to delegate the corresponding calls in response to user actions, events of the plugin or the IDE itself. The Java code model generated by IDEA and used by MetricsTree is the PSI (Program Structure Interface) tree, which is similar to the AST (Abstract Syntax Tree) in this IDE. This makes it possible to feed the model of program elements of the source Java file to MetricsTree input as efficiently as it is implemented in IDEA itself.

MetricsTree accesses Java code elements in the form of a PSI-tree, performs their calculation and collects the value of a particular metric. Two hierarchies of classes, for metrics and for code elements, are used to model the subject domain. Hierarchies of classes have been developed that are responsible for sequential traversal of PSI-tree nodes and calculate them according to the rules defined by the metrics, which are traditionally represented as the Visitor pattern for this kind of problems.

The source code of MetricsTree (https://github.com/b333vv/metricstree) is freely distributed under the Apache 2.0 license and is structured, clean, simple and straightforward to the extent that there is no need for more detailed descriptions of its technical implementation in this article.

Supported metrics

We analyzed the most cited research on object-oriented metrics, some of them were related to coupling metric analysis [7–10], the other parts is with maintenance [11], reusability [12, 13] and complexity [14, 15] and chose a set of metrics to implement in our plugin. The set turned out to be one of the most significant among the peers. In total, MetricsTree supports 61 metrics: 18 on project level, 11 on package level, 22 on class level, 11 on method level. The full list of supported metrics can be found in Table A (Appendix) and on the plugin's GitHub page (https://github.com/b333vv/metricstree).

Along with including a large number of supported metrics, the software design of the plugin is intended for minimizing the amount of rework required to add new types of metrics. All countable Java code elements are available (which the IDE is responsible for), and there are entities for representing metrics. Thus, all that remains to implement a new basic metric is to define an heir in the Visitor hierarchy, describing in it the rules for counting Java code elements of this new metric. If a derivative metric is added, it is necessary to make an heir in the metrics hierarchy, describing in it the formula for complexing the base and derivative metrics to get the value of the new metric.

Table 1 compares MetricsTree with its closest counterparts in terms of the main families of supported metrics (open-source tools are highlighted in gray).

Along with the calculation of metrics values, MetricsTree implements functions for setting reference metrics values. These values were derived based on the analysis of studies of metrics correlation with software attributes. At the same time, to adapt to the peculiarities of development, the values of the benchmark intervals can be changed. To increase the speed of assimilation of information by developers, MetricsTree uses color indication to indicate the extent to which metrics fall within the benchmark interval.

Key features

Trees

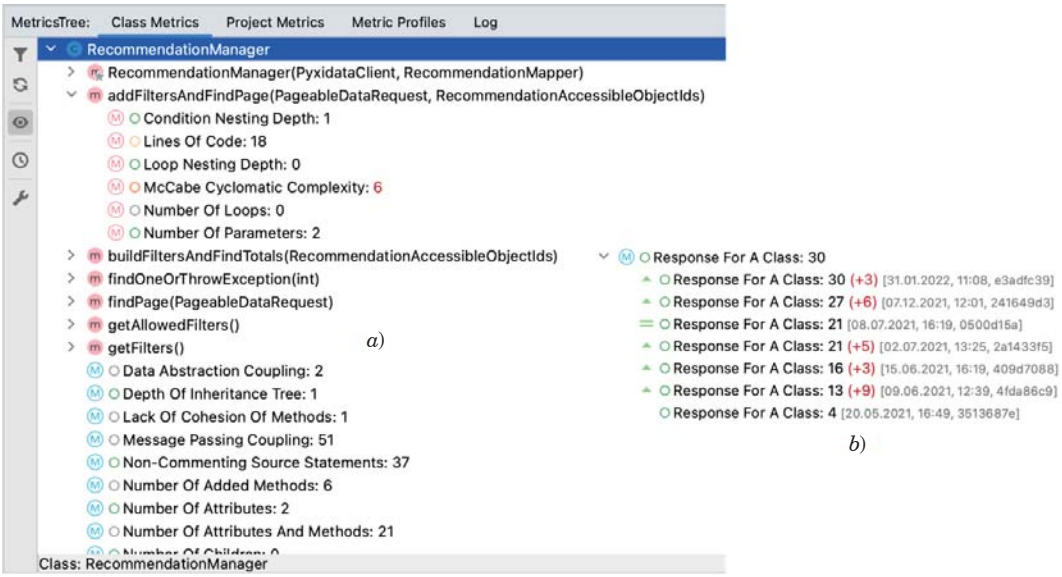
Trees grouping metrics by class and method level are used to display the hierarchy of class metrics (Fig. 1, *a*). The grouping levels for mapping the application metrics are application, package, class, and method (Fig. 2, *a*, *b*). In addition, the project metrics tree groups metrics by metrics authors (e.g., Chidamber — Kemerer and Robert Martin's metrics, etc.). For each metric, a choice of 5 colors is used: green (if the metric value is within the control limits), yellow (if the metric value is slightly out of bounds), red (if the metric value is out of bounds), bright red (if the metric value is significantly out of bounds). The function for getting the evolution of the metric value of each class during development is also implemented (Fig. 1, *b*).

Treemaps

Treemap charts are used to view data in a hierarchical view, using a grid of rectangles with sizes

■ ***Table 1.*** Support for families of metrics by metric analysis tools

App/Metric Set	MOOD [16]	QMOOD [17]	Robert C. Martin [18]	Chidamber – Kemerer [19]	Lorenz – Kidd [20]	Li – Henry [21]	Lanza – Marinescu [22]	Holstead [23]	MI [24]	Statistic
CAST's AIP (doc.castsoftware.com/display/TECHNOS)			+			+				+
CKJM (www.spinellis.gr/sw/ckjm)				+						
CMT++/CMTJava (www.verifysoft.com/en_cmtx.html)								+	+	+
CodeMR (www.codemr.co.uk)			+	+			+			+
Halstead Metrics Tool (sourceforge.net/projects/halsteadmetricstool)								+		
JHawk (www.virtualmachinery.com/jhawkprod.htm)			+	+		+		+	+	
MetricsReloaded (github.com/BasLeijdekkers/Metrics-Reloaded)	+		+	+	+	+		+		+
MetricsTree	+	+	+	+	+	+	+			+
PMD							+			+
SonarQube						+			+	+
Understand (emenda.com/scitools-understand)			+	+		+				+

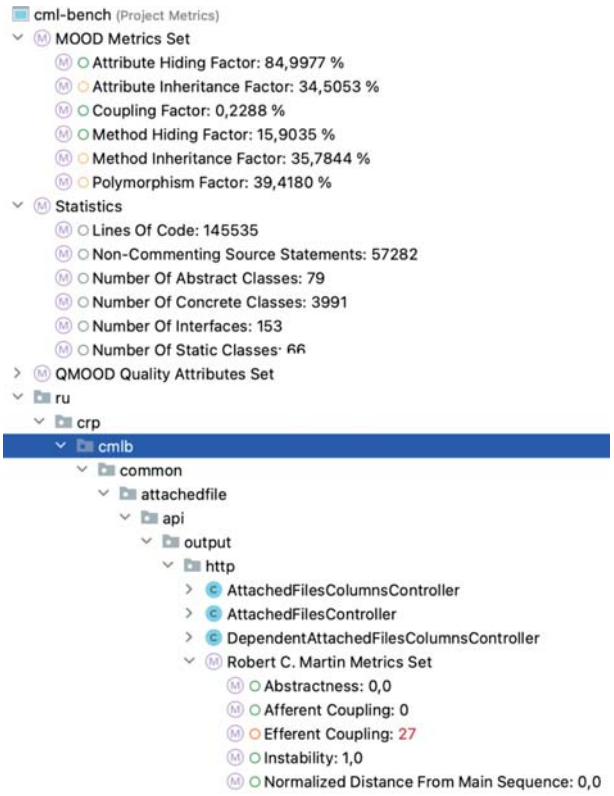


■ ***Fig. 1.*** Class metrics tree (*a*) and metric evolution tree using the RFC metric as an example (*b*)

proportional to the size of the class, which is defined by the Non-Commenting Source Statements metric. The color of each rectangle (5 shades from red to green) reflects the degree to which the value of the metric selected in the list on the left corresponds to the reference range. Clicking on the list in the left part of the window (Fig. 3) displays the Treemap chart in the middle, clicking on the rectangle in the middle representing the class displays all the metrics of this class on the right.

Charts

Several types of charts are used to improve the efficiency of metric analysis of Java source code. Some of them clearly show the general metric picture reflecting the state of source code quality



■ ***Fig. 2.*** Project metrics tree

(Figs. 4–6). Another part visualizes different types of relationships between metric values and metric profiles and serves for more in-depth metric analysis of structural properties of Java code (Figs. 7–9).

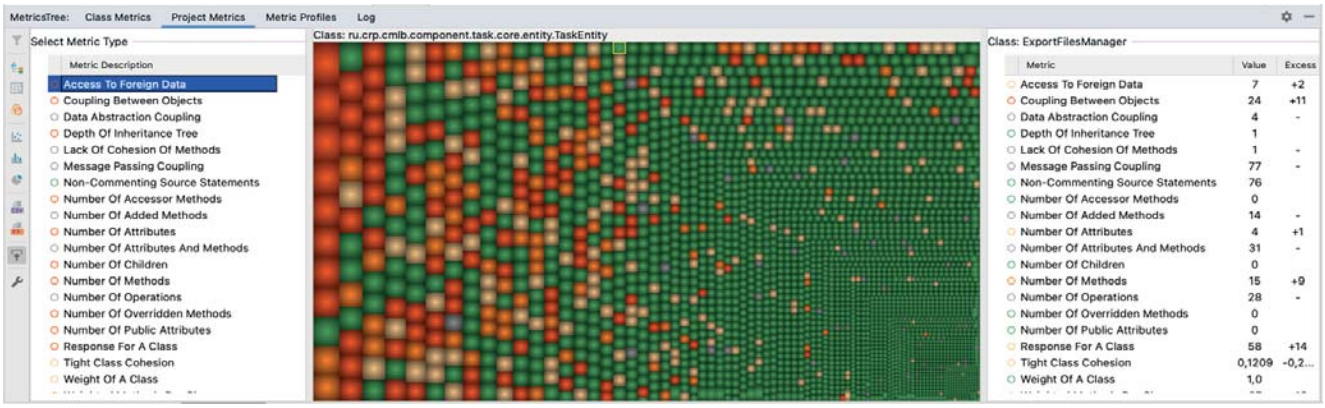
Synergy with the IDE

Once a class is selected in the project tree, its source code is opened in the editor, its metrics are simultaneously calculated and displayed as a tree and a table (Fig. 10, *a*–*d*). The developers do not have to take any additional steps to get the metrics values for the class processed at the moment. Embedding the information about the metrics into the IDE generates a synergy effect, serving to increase the degree of validity and effectiveness of the decisions made by the developers by promptly obtaining the metrics values.

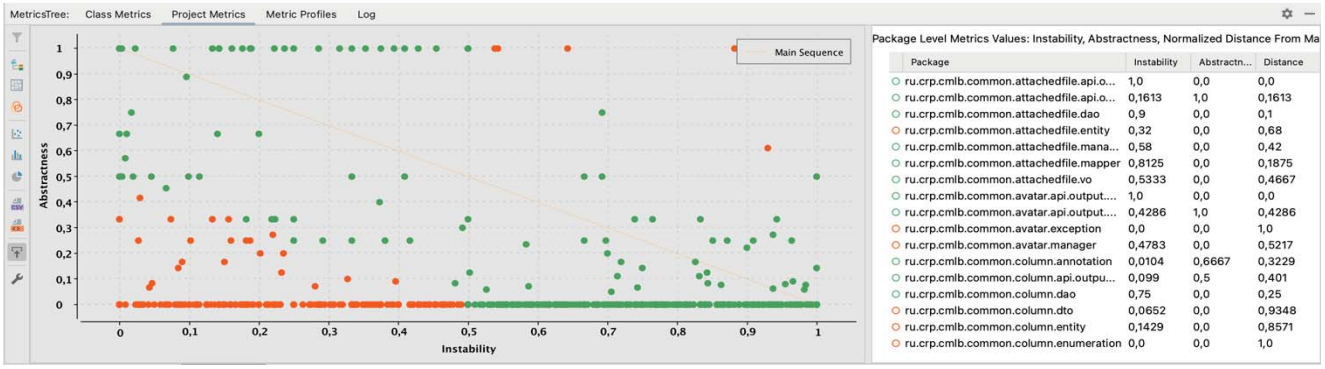
MetricsTree provides an option to control the composition of the metrics displayed in the tree, as well as the reference thresholds for base and derived metrics values.

Metric profiles

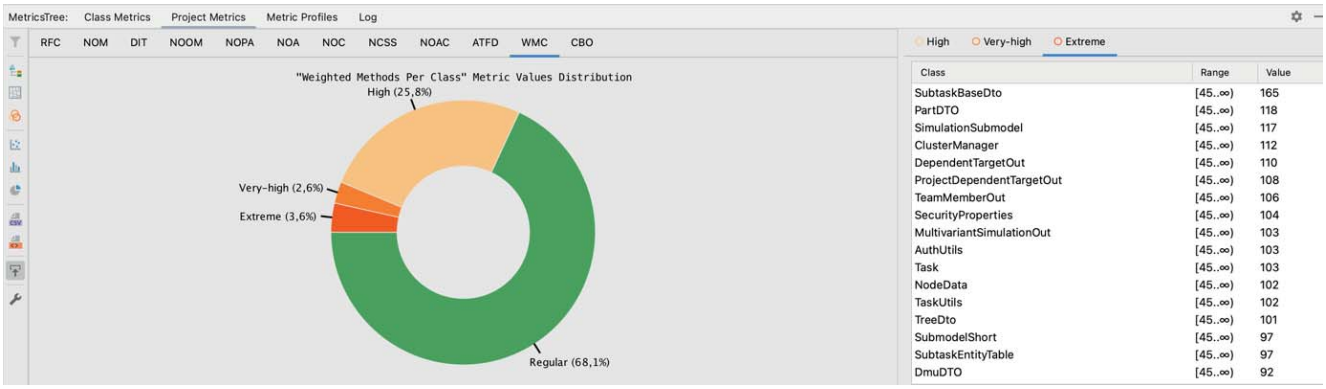
The metric profiles tool was originally inspired by research into the possibility of detecting code smells using metrics [24]. We have extended the interpretation of this tool by calling it the metric



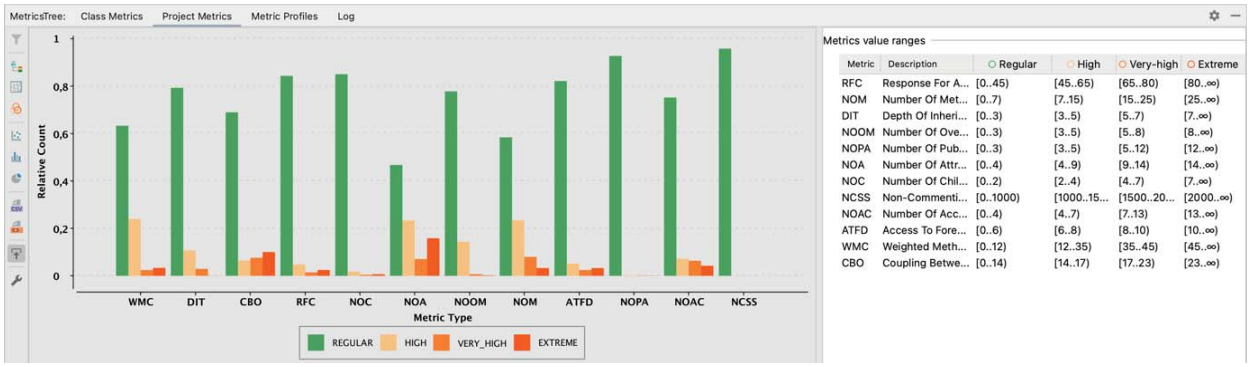
■ ***Fig. 3.*** Treemap with class-level metrics



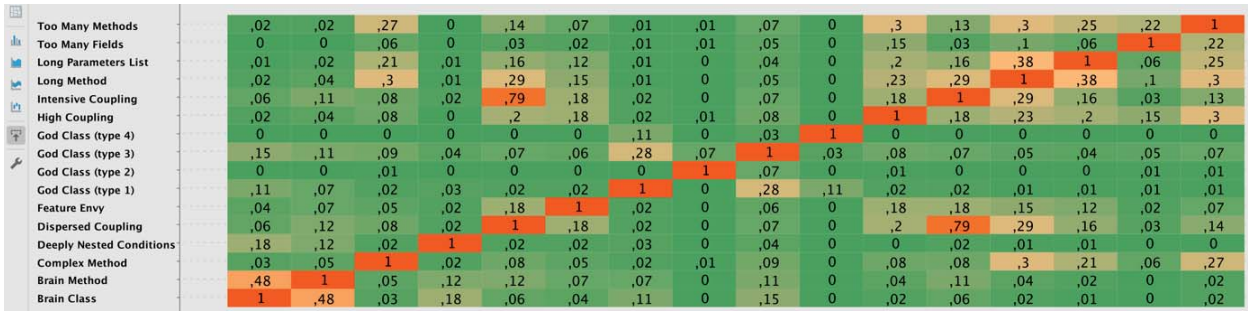
■ ***Fig. 4.*** Abstractness vs. instability distribution



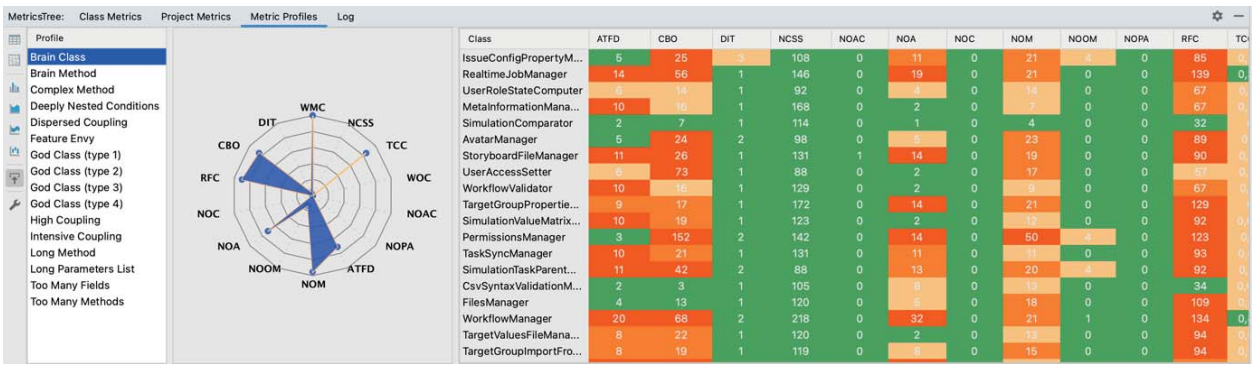
■ ***Fig. 5.*** Pie chart for metrics distribution by type



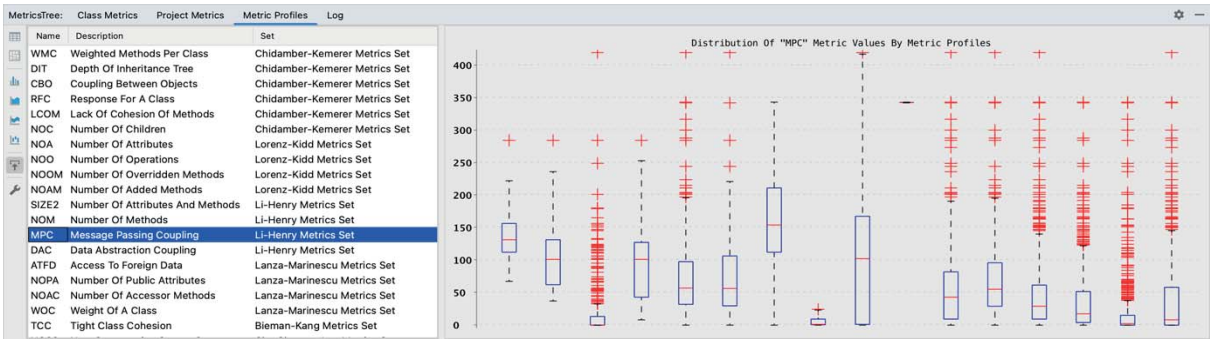
■ ***Fig. 6.*** Metrics distribution



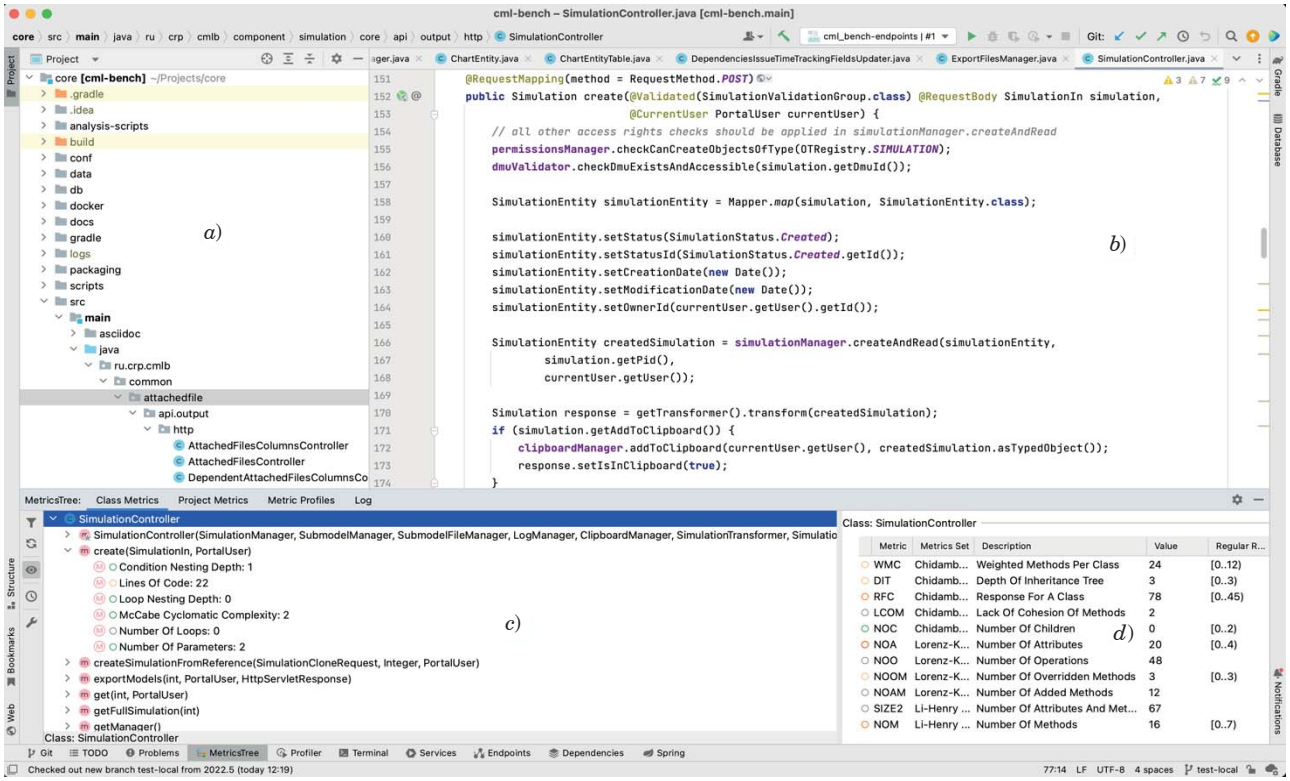
■ ***Fig. 7.*** Correlation between metric profiles



■ ***Fig. 8.*** Distribution of metric values across metric profiles



■ ***Fig. 9.*** Correlation between metric values and metric profiles using the MPC metric as an example



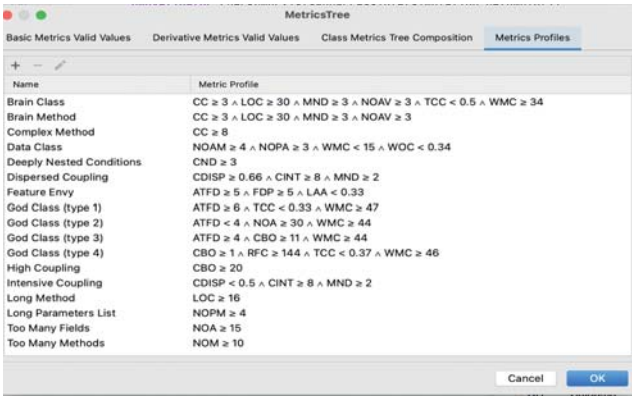
■ ***Fig. 10.*** Main window in IntelliJ IDEA with embedded MetricsTree plugin: *a* — selecting a class in the project tree; *b* — displaying source code in the editor; *c* —results of metrics calculation in the tree; *d* — results of the metrics calculation in the table

profiles formation tool, allowing to identify classes with given sets of metric values by constructing appropriate samples. The conjunct of metric value ranges serves as an analytical specification of the metric profile (Fig. 11).

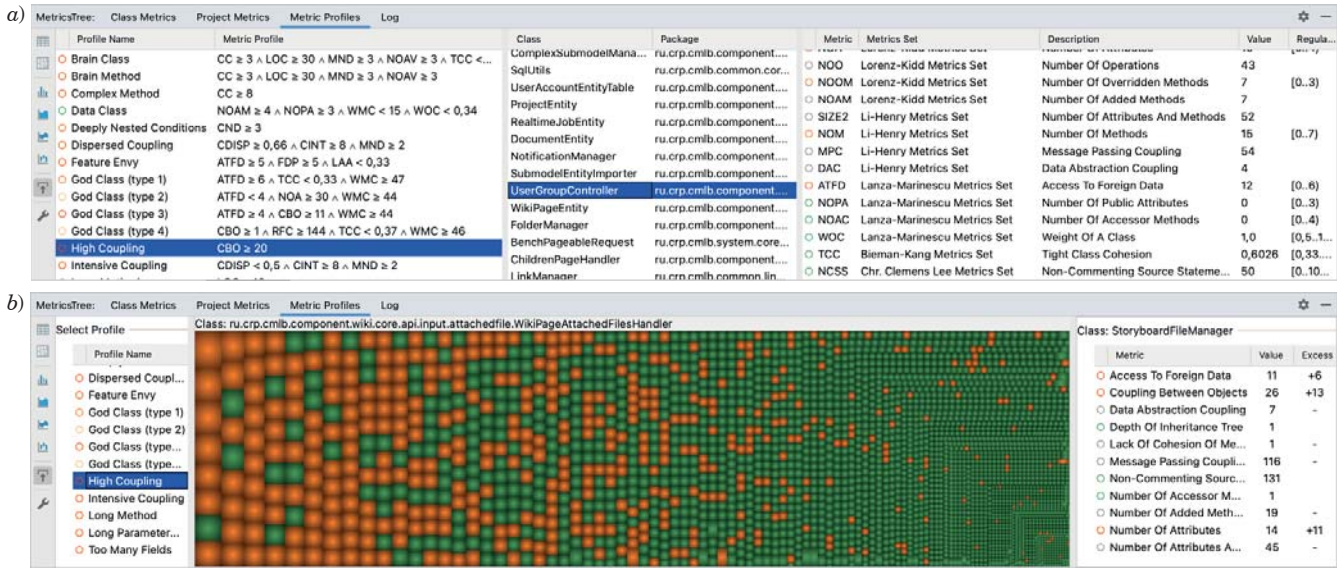
Once defined, metric profiles allow generating subsets of classes with certain metric values, which can be visualized as a table (Fig. 12, *a*) or a Treemap (Fig. 12, *b*).

Comparison with counterparts in terms of functionality

Table 2 presents a comparison of the selected tools for metric analysis within the selected set of



■ ***Fig. 11.*** Dialog box for setting the metric thresholds



■ ***Fig. 12.*** Class distribution by metric profile: *a* — table view; *b* — Treemap view

■ ***Table 2.*** Comparison of metrics calculation tools by functionality

App/Feature	Automatic metric calculation	Metric evolution monitoring	Tabular view	Tree view	Diagram view	Treemap view	Metric profiles	Metric profiles setting	Metric values control	Metric reference value setting
CAST's AIP			+		+				+	
CKJM			+							
CMT++/CMTJava			+						+	
CodeMR			+		+	+	+	+	+	
Halstead Metrics Tool			+							
JHawk			+		+					
MetricsReloaded			+		+				+	
MetricsTree	+	+	+	+	+	+	+	+	+	+
PMD			+				+			
SonarQube			+		+				+	+
Understand			+							

properties. Many of the metrics analysis tools compared have useful features not considered in the presented comparison. The rows with open-source tools are highlighted in gray.

Effects of use

Over the past year, while still using SonarQube and PMD, we have integrated MetricsTree into our quality assurance processes. MetricsTree is implemented as a multifunctional tool, the main aspects of its implementation can be summarized as the following three groups.

Working on the task

The main element of the MetricsTree that accompanies a programmer's daily work in the IDE is the tree of class metrics (see Fig. 10, *c*). It displays the values of metrics at the class and method level, drawing the developer's attention to the class metrics whose values fall outside the control limits by means of color-coding. There are five color codes, which change depending on the degree of deviation from

the reference values. The developer can use the user interface to control the thresholds of the metric value ranges. To help the developer focus on the important tasks, the UI has a feature to control the composition of the metrics in the tree. Developers do not need to take any steps to manage the metric values, as normal work with Java classes in IDEA in the project tree and editor is accompanied by quick calculation and display of metrics values for the class considered at the moment. Based on feedback from our developers, we can conclude that adding real-time metrics information to the typical IDE window structure helps them make the right decisions when working on a certain class or its method in a timelier manner.

Final checks on the task

After completing the task in the local environment, the developer performs a quality check of the new or modified functionality before implementing the changes in the main branch of code. The developer uses the metric profiles tool to look for classes with code smells or defects. If the developer discovers classes for which they are responsible in the generated sample, they fix the defects found before putting the results of the work into the common code base.

Prerelease activities

After the set of tasks making up the new CML-Bench release is completed, several specialized measures are taken as part of our quality assurance regulations.

First, analysis of code smells or defects is performed based on metrics profiles, similar to the one described in the previous section, but project-wide.

The second important component is the analysis of project quality using diagrams. We evaluate the evolution of QMOOD metrics on a diagram, controlling for possible quality degradation associated with refinements. Other diagrams are used to investigate different aspects of the project metrics state at a deeper level: distribution of metrics values, dependencies between different metrics, dependencies between metrics and metrics profiles, etc.

Third, we use an unloading of all project metrics as an XML file and compare it to a similar unloading associated with the previous releases. Assessing the evolution of metrics values, we look for system elements whose quality has deteriorated due to changes made when the current release was prepared, analyze the reasons for the deterioration, and determine ways to fix them.

As noted above, after we developed and started to use MetricsTree, we continued to use SonarQube and PMD. In the year that MetricsTree was adopted to support quality assurance processes, our three CML-Bench development teams created 5 new services and refined 11 existing services, while writing about 72,000 lines of new and modified Java source code. At the same time, the number of code defects identified by SonarQube and PMD during this period decreased by 34%. The improvement in quality metrics from release to release of CML-Bench reinforces our confidence in the effectiveness of automating quality assurance processes with our MetricsTree tool and justifies the importance of its use in the development process.

Conclusion

The plugin we developed for IDE IntelliJ IDEA is designed for metric analysis of structural properties of Java code and has a number of key features:

— is based on rigorous mathematical models for quality and measurement of software quality control theory;

— aims to minimize the cost of adding new types of metrics;

— synergistically extends the space generated by the IDE to quickly add information about quantitative properties of the software entities developed;

— supports different ways to visualize metrics information, i.e., trees, Treemaps, charts of different types to enhance analysis of quantitative properties of Java code;

— implements a mechanism for describing metric profiles and forming class subsets with appropriate combinations of metric values.

MetricsTree has been successfully used for a year to ensure the high quality of the advanced CML-Bench platform for developing and implementing digital twins we are working on.

Our immediate plans include:

— conducting an in-depth study to validate the feasibility of the metric profile mechanism;

— conducting a comprehensive study of the impact of various MetricsTree features used on software quality assurance processes;

— applying MetricsTree to multiple open-source Java projects to increase the generality and validity of the conclusions;

— evolution towards taking into account the other components of software development.

As short-term directions for MetricsTree, we will focus on engineering a tool for formulating analytical expressions for metric profiles. Furthermore, we plan to implement a number of additional metrics:

— Halstead Metrics [23];

— Maintainability Index (MI) [24];

— Decoupling Level (DL) [25];

— Propagation Cost (PC) [26].

In addition, the MetricsTree repository on Github currently contains about 30 Issues created by plugin users from different countries; we intend to analyze the feasibility of implementing these in the near future.

Appendix

■ ***Table A.*** List of metrics calculated by MetricsTree

Level	Family (Author)	Abbreviation	Description	Source
Project	Chr. Clemens Lee	NCSS	Non-Commenting Source Statements	www2.informatik.hu-berlin.de/swt/intkoop/jcse/tools/JavaNCSS%20-%20A%20Source%20Measurement%20Suite%20for%20Java.html
Package	Chr. Clemens Lee	NCSS	Non-Commenting Source Statements	
Class	Chr. Clemens Lee	NCSS	Non-Commenting Source Statements	
Project	–	LOC	Lines of Code	–
Package	–	LOC	Lines of Code	–
Class	–	LOC	Lines of Code	–
Method	–	LOC	Lines of Code	–
Project	–	NOC	Number of Concrete Classes	–
Package	–	NOC	Number of Concrete Classes	–
Project	–	NOA	Number of Abstract Classes	–
Package	–	NOA	Number of Abstract Classes	–
Project	–	NOSC	Number of Static Classes	–
Package	–	NOSC	Number of Static Classes	–
Project	–	NOI	Number of Interfaces	–
Package	–	NOI	Number of Interfaces	–
Project	MOOD	MHF	Method Hiding Factor	[16]
Project	MOOD	AHF	Attribute Hiding Factor	
Project	MOOD	MIF	Method Inheritance Factor	
Project	MOOD	AIF	Attribute Inheritance Factor	
Project	MOOD	PF	Polymorphism Factor	
Project	MOOD	CF	Coupling Factor	
Project	QMOOD	–	Reusability	[17]
Project	QMOOD	–	Flexibility	
Project	QMOOD	–	Understandability	
Project	QMOOD	–	Functionality	
Project	QMOOD	–	Extendibility	
Project	QMOOD	–	Effectiveness	
Package	Robert C. Martin	Ce	Efferent Coupling	[18]
Package	Robert C. Martin	Ca	Afferent Coupling	
Package	Robert C. Martin	I	Instability	
Package	Robert C. Martin	A	Abstractness	
Package	Robert C. Martin	D	Normalized Distance from Main Sequence	
Class	Chidamber — Kemerer	WMC	Weighted methods per class	[19]
Class	Chidamber — Kemerer	DIT	Depth of Inheritance Tree	
Class	Chidamber — Kemerer	NOC	Number of Children	
Class	Chidamber — Kemerer	CBO	Coupling between object classes	

■ *Ending of Table A*

Level	Family (Author)	Abbreviation	Description	Source
Class	Chidamber — Kemerer	RFC	Response for a Class	
Class	Chidamber — Kemerer	LCOM	Lack of cohesion in methods	
Class	Lorenz – Kidd	NOA	Number of Attributes	[20]
Class	Lorenz – Kidd	NOO	Number of Operations	
Class	Lorenz – Kidd	NOAM	Number of Added Methods	
Class	Lorenz – Kidd	NOOM	Number of Overridden Methods	
Class	Li – Henry	SIZE2	Number of Attributes and Methods	[21]
Class	Li – Henry	MPC	Message Passing Coupling	
Class	Li – Henry	DAC	Data Abstraction Coupling	
Class	Li – Henry	NOM	Number of Methods	
Class	Lanza – Marinescu	ATFD	Access to Foreign Data	[22]
Class	Lanza – Marinescu	NOPA	Number of Public Attributes	
Class	Lanza – Marinescu	–	Number of Accessor Methods	
Class	Lanza – Marinescu	WOC	Weight of a Class	
Class	Bieman – Kang	TCC	Tight Class Cohesion	[27]
Method	McCabe	CC	McCabe Cyclomatic Complexity	–
Method	–	–	Maximum Nesting Depth	–
Method	–	–	Loop Nesting Depth	–
Method	–	–	Condition Nesting Depth	–
Method	–	–	Number of Loops	–
Method	–	LAA	Locality of Attribute Accesses	–
Method	–	FDP	Foreign Data Providers	–
Method	–	NOAV	Number of Accessed Variables	–
Method	–	CINT	Coupling Intensity	–
Method	–	CDISP	Coupling Dispersion	–

References

1. State Standard R 57700.37-2021. *Digital Twins of Products. General Provisions*. Moscow, Rossijskij institut standartizacii Publ., 2021. 11 p. (In Russian).
2. Borovkov A. I., Rozhdestvenskiy O. I., Pavlova E., Glazunov A. I., and Savichev K. Key barriers of digital transformation of the high-technology manufacturing: An evaluation method. *Sustainability*, 2021, vol. 13, no. 20: 11153. doi:10.3390/su132011153
3. Borovkov A. I., Gamzikova A. A., Kukushkin K. V., Ryabov Yu. A. *Cifrovye dvojniki v vysokotekhnologichnoj promyshlennosti* [Digital Twins in the High-Technology Manufacturing Industry]. Politekhnicheskij universitet Publ., Saint-Petersburg, 2019. 62 p. (In Russian).
4. Burakov V. V. *Upravlenie kachestvom programmny`kh sredstv* [Software Quality Management]. Saint-Petersburg, GUAP Publ., 2009. 287 p. (In Russian).
5. Burakov V. V. The use of simulation modeling to improve software quality. *Trudy devyatoj vserossijskoj nauchno-prakticheskoj konferencii po imitacionnomu modelirovaniyu i ego primeneniyu v nauke i promyshlennosti "Imitacionnoe modelirovanie. Teoriya i praktika" IMMOD-2019* [Proc. of the IX All-Russian Scientific and Practical Conf. on Simulation Modelling and its Applications in Science and Industry. "Simulation Modelling. Theory and Practice" (IMMOD-2019)], 2019, pp. 368–374 (In Russian).
6. Burakov V. V., Kulakov A. Yu., Cherniy A. N. Software maintenance evaluation. *Informatization and Communication*, 2019, no. 4, pp. 14–21 (In Russian). doi:10.34219/2078-8320-2019-10-4-14-21
7. Kumar N. R., Viji C., and Duraisamy S. Measuring cohesion and coupling in object oriented system using Java reflection. *ARPN Journal of Engineering and Applied Sciences*, 2015, vol. 10, no. 7, pp. 3096–3101.

8. Ankush Vesra, Rahul. A study of various static and dynamic metrics for open source software. *International Journal of Computer Applications*, 2015, vol. 122, no. 10, pp. 17–19. doi:10.5120/21736-4927

9. Nicolaescu A., Lichter H., Xu Yi. Evolution of object oriented coupling metrics: A sampling of 25 years of research. *Software Architecture and Metrics (SAM)*, 2015 IEEE/ACM 2nd Intern. Workshop, May 2015. doi:10.1109/SAM 2015.14

10. Schnoor H., Hasselbring W. Toward measuring software coupling via weighted dynamic metrics. *40th Intern. Conf. on Software Engineering: Common Proc.*, ACM/IEEE, 2018, pp. 342–343. doi:10.1145/3183440.3195000

11. Tkachuk M., Nagornyi K., and Gamzayer R. Models, methods and tools for effectiveness estimation of post object oriented technologies in software maintenance. *ICTERI 2015*, CCIS 594, Springer International Publisher Switzerland, 2016, pp. 20–37.

12. Padhy N., Satapathy S., and Singh R. P. State-of-the-art object oriented metrics and its reusability: A decade review. *Smart Computing and Informatics*, 2018, pp. 431–441. doi:10.1007/978-981-10-5544-7_42

13. Padhy N., Satapathy S., Singh R. P., and Sethlani J. A systematic literature review of an object oriented metrics components: Case study for evaluation of reusability criteria. *Intern. Conf. on Advanced Studies in Engineering and Sciences*, 2017, pp. 49–61.

14. Mao C., Xu Ch. Entropy based dynamic complexity metrics for service oriented systems. *24th Asia-Pacific Software Engineering Conf. Workshops*, IEEE, 2017, pp. 90–97. doi:10.1109/APSECW.2017.14

15. Aswini S., Yazhini M. An assessment framework of routing complexities using LOC metrics. *Intern. Conf. on Innovations in Power and Advanced Computing Technologies*, IEEE, 2017, pp. 1–6. doi:10.1109/IPACT.2017.8245022

16. Brito e Abreu F. and Carapuça R. Object-oriented software engineering: Measuring and controlling the development process. *4th Intern. Conf. on Software Quality*, Mc Lean, VA, USA, 1994, pp. 1–8.

17. Bansiya Ja., and Davis C. G. A hierarchical model for object-oriented design quality assessment. *IEEE Transactions on Software Engineering*, 2002, vol. 28, iss. 1, pp. 4–17.

18. Martin R. C. *Agile Software Development: Principles, Patterns, and Practices*. Alant Apt Series. Prentice Hall, Upper Saddle River, NJ, USA, 2002. 552 p.

19. Chidamber S. R., and Kemerer C. F. A metrics suite for object oriented design. *IEEE Transactions on Software Engineering*, 1994, vol. 20 (6), pp. 476–493.

20. Lorenz M., Kidd J. *Object Oriented Software Metrics*. Pearson, 2008. 160 p.

21. Li W., and Henry S. Object-oriented metrics that predict maintainability. *Journal of Systems and Software*, 1993, vol. 23, iss. 2, pp. 111–122.

22. Lanza M., Marinescu R. *Object-oriented metrics in practice*: *Using software metrics to characterize, evaluate, and improve the design of object-oriented systems*. Springer, 2006. 207 p. doi:10.1007/3-540-39538-5

23. Chhillar R. S., and Gahlot S. An evolution of software metrics: A review. *Proc. of the Intern. Conf. on Advances in Image Processing, ICAIP 2017*, New York, NY, USA, 2017, pp. 139–143. doi:10.1145/3133264.3133297

24. Coleman D., Ash D., Lowther B., and Oman P. Using metrics to evaluate software system maintainability. *Computer*, 1994, vol. 27, no. 8, pp. 44–49. doi:10.1109/2.303623

25. Mo R., Cai Ya., Kazman R., Xiao Lu, Feng Q. Decoupling level: A new metric for architectural maintenance complexity. *Proc. of the Intern. Conf. on Software Engineering*, Austin, TX, 2016, pp. 499–510.

26. MacCormack A., Rusnak J., Baldwin C. Y. Exploring the structure of complex software designs: An empirical study of open source and proprietary code. *Management Science*, 2006, vol. 52, iss. 7, pp. 1015–1030.

27. Bieman J. M., Kang B. Cohesion and reuse in an object-oriented system. *Proc. of the 1995 Symp. on Software Reusability*, Seattle, Washington, United States, 1995, pp. 259–262.

УДК 004.05
doi:10.31799/1684-8853-2023-1-17-28
EDN: ORVZMP

Средство для углубленного метрического анализа исходного кода на Java

В. В. Бураков[а], доктор техн. наук, доцент, orcid.org/0000-0002-0158-8681, burakov@compmechlab.com
А. И. Боровков[б], канд. техн. наук, профессор, orcid.org/0000-0003-3177-0959
[а]ООО Лаборатория «Вычислительная механика», Гжатская ул., 21, к. 2а, Санкт-Петербург, 195220, РФ
[б]Санкт-Петербургский политехнический университет Петра Великого, Политехническая ул., 29, Санкт-Петербург, 195251, РФ

Введение: несмотря на значительные усилия многочисленных исследователей и разработчиков, процедуры оценки качества программного обеспечения все еще нуждаются в формализации и автоматизации. **Цель:** разработать специализированное программное средство, предназначенное для количественной оценки структурных свойств Java-кода. **Результаты:** разработано программное средство MetricsTree, которое рассчитывает 61 устоявшуюся объектно-ориентированную метрику (это один из самых больших наборов среди аналогичных инструментов). MetricsTree интегрировано в IDE для обеспечения максимально быстрой доставки инфор-

мации, содержит уникальные средства визуализации для повышения эффективности анализа метрик, а также реализует механизм профилей метрик для выбора классов на основе набора значений метрик. **Практическая значимость:** в результате применения MetricsTree для автоматизации процессов обеспечения качества при разработке флагманской системы Санкт-Петербургского государственного политехнического университета Петра Великого CML-Bench (платформы для разработки и применения цифровых двойников) в течение года среднее количество выявленных программных дефектов внешними средствами сократилось на 34 %.

Ключевые слова — качество программного обеспечения, метрики программного обеспечения, объектно-ориентированные метрики, метрики исходного кода на Java, программные средства расчета метрики, метрический анализ, визуализация программного обеспечения, дефекты исходного кода.

Для цитирования: Burakov V. V., Borovkov A. I. Advanced metric analysis tool for Java source code. *Информационно-управляющие системы*, 2023, № 1, с. 17–28. doi:10.31799/1684-8853-2023-1-17-28, EDN: ORVZMP

For citation: Burakov V. V., Borovkov A. I. Advanced metric analysis tool for Java source code. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 17–28. doi:10.31799/1684-8853-2023-1-17-28, EDN: ORVZMP

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

UDC 003.26
doi:10.31799/1684-8853-2023-1-29-40
EDN: KSCBTZ

Articles

Post-quantum algebraic signature algorithms with a hidden group

A. A. Moldovyan[a], *Dr. Sc., Tech., Professor, Chief Researcher, orcid.org/0000-0001-5480-6016*
N. A. Moldovyan[a], *Dr. Sc., Tech., Professor, Chief Researcher, orcid.org/0000-0002-4483-5048, nmold@mail.ru*
[a]*St. Petersburg Federal Research Center of the RAS, 39, 14th Line, 199178, Saint-Petersburg, Russian Federation*

Introduction: *The development of post-quantum standards on digital signature algorithms is one of the current challenges faced by the global cryptographic community. Recently, two types of algebraic signature schemes with a hidden group have been proposed, in which the finite non-commutative associative algebras set over the field GF(p) are used as an algebraic support. The design of that type of signature algorithms on the algebras set over the finite fields of Characteristic two represent significant interest for improving the performance and reducing the hardware implementation cost.* ***Purpose:*** *To develop post-quantum algebraic signature algorithms in which computations in a finite field of Characteristic two are used.* ***Results:*** *Several 4-dimensional finite non-commutative algebras set over the GF(2^z) fields are proposed as algebraic support of the signature schemes with a hidden group. We suggest some recommendations for choosing the value of the extension degree z. In particular cases the value of z represents a Mersenne degree. Compared with the signature algorithms which are based on the hidden logarithm problem, the algebraic signature algorithms based on the computational complexity of solving systems of many quadratic equations with many variables are considered to be a preferable type of cryptoschemes with a hidden group. We have introduced new practical signature algorithms with a hidden group. In two of the developed algorithms the signature (e,* ***S****) represents an integer e and a 4-dimensional vector* ***S*** *and is verified with vector equations with three and four entries of the signature element* ***S****.* ***Practical relevance:*** *Like other known signature schemes with a hidden group, the proposed two schemes have sufficiently small size of signature and public key. Due to comparatively small hardware implementation cost and high performance, the introduced candidates for post-quantum signature algorithms represent practical interest and are attractive as a potential prototype of a post-quantum digital signature standard.*

Keywords *— post-quantum cryptoschemes, computer security, digital signature, discrete logarithm problem, multivariate cryptography, finite non-commutative algebras, associative algebras, cyclic groups, multidimensional cyclicity.*

For citation: Moldovyan A. A., Moldovyan N. A. Post-quantum algebraic signature algorithms with a hidden group. *Informatsionno-upravliaiushchiesistemy* [Information and Control Systems], 2023, no. 1, pp. 29–40. doi:10.31799/1684-8853-2023-1-29-40, EDN: KSCBTZ

Introduction

The predicted emergence of quantum computers in practice in the near future and the availability of polynomial in time quantum algorithms for solving the discrete logarithm problem and the factorization problem [1–3] determine the high degree of relevance of the development of post-quantum public-key cryptographic schemes, which are resistant to quantum attacks (attacks with using ordinary and quantum computers). Post-quantum signature algorithms are to be based on hard problems different from discrete logarithm and factorization problems.

In particular, the quantum computer is not effective for finding solutions of systems of many quadratic equations with many unknowns and computational difficulty of this problem underlies the resistance of the multivariate signature algorithms [4–6]. There are known signature schemes on algebras [7, 8], on algebraic lattices [9], on codes [10, 11], and on hash functions [12]. A certain disadvantage of the known post-quantum signature schemes is a large size of public key and signature. In order to reduce the total size of the signature and the key, the signature schemes with a hidden group are proposed, in which finite non-commutative associative algebras (FNAA) are used as an algebraic support [13, 14]. One can distinguish two types of algorithms with a hidden group, which differ in the type of the used computationally difficult problem:

1) algorithms, security of which is based on the computational difficulty of the hidden discrete logarithm problem (HDLP) [13, 15];

2) algorithms, security of which is based on the computational difficulty of finding a solution of a system of many quadratic equations with many unknowns [14, 16].

A hidden group represents a subset of elements of some m-dimensional FNAA, which composes a commutative group. In the algorithms of the both types, the elements of the public key are computed as a masked (secret) element $\mathbf{H}$ of the hidden group. The masking is performed, for example, as the left and the right multiplications of the m-dimensional invertible vector $\mathbf{H}$ by some secret invertible vectors $\mathbf{A}$ and $\mathbf{B}$ which satisfy the following conditions $\mathbf{BA} \neq \mathbf{AB}$, $\mathbf{HA} \neq \mathbf{AH}$, $\mathbf{HB} \neq \mathbf{BH}$.

The FNAAs defined over a ground finite field $GF(p)$ with prime $p = 2q + 1$, where q is also a

prime, are used as algebraic supports of the known signature algorithms with a hidden group [7, 13]. To improve the performance and reduce the hardware implementation cost, development of the post-quantum algebraic signature algorithms on FNAAs set over finite fields of characteristic two, i. e. over the fields $GF(2^z)$, represents significant interest.

In this paper, three different 4-dimensional FNAAs, including the algebras defined by a sparse basis vector multiplication tables (BVMTs), set over the $GF(2^z)$ fields are used as algebraic support of the proposed three new algebraic signature algorithms with a hidden group: i) one HDLP-based algorithm and ii) two algorithms with a hidden group, which are based on computational difficulty of solving a system of many quadratic equations with many unknowns. Compared with the former one, the latter are considered as more preferable candidates for post-quantum signature schemes. Recommendations for choosing the value of the extension degree z of the $GF(2^z)$ field are suggested for each of two types of the signature algorithms with a hidden group.

Four-dimensional FNAA used as algebraic support

Brief explanation of the notion of FNAA is provided in [16]: "A vector space of dimension m, which is set over a finite field $GF(p)$ or $GF(2^z)$, with additionally defined vector multiplication operation (that possesses the property of distributivity at the left and at the right relatively the addition operation) is called m-dimensional algebra [16]. A vector $\mathbf{A}$ can be represented in the following two forms: i) as an ordered set of its coordinates: $\mathbf{A} = (a_0, a_1, \ldots, a_{m-1})$ and ii) as a sum of its components: $\mathbf{A} = a_0\mathbf{e}_0 + a_1\mathbf{e}_1 + \ldots + a_{m-1}\mathbf{e}_{m-1}$, where $\mathbf{e}_i$ ($i = 0, 1, \ldots, m-1$) are basis vectors. If the defined multiplication operation is non-commutative and associative, then one gets m-dimensional FNAA. Usually, the product of the vectors $\mathbf{A} = \sum_{i=0}^{m-1} a_i\mathbf{e}_i$ and $\mathbf{B} = \sum_{j=0}^{m-1} b_j\mathbf{e}_j$ is defined by the formula $\mathbf{AB} = \sum_{i=0}^{m-1}\sum_{j=0}^{m-1} a_i b_j \mathbf{e}_i\mathbf{e}_j$, where the values a_i and b_i are multiplied as the field elements and every the product of two formal basis vectors is to be replaced by an one-component vector indicated in a cell at the intersection of the i-th row and j-th column of so called BVMT".

Usually, to perform one multiplication operation in some 4-dimensional algebra (see, for example, Table 1 [8]) one need to execute 16 multiplications and 12 additions in the field $GF(p)$ or $GF(2^z)$. However, computational complexity of this operation can be reduced, using sparse BVMTs (see, for example, Tables 2 [7] and 3 [16]).

■ ***Table 1.*** Multiplication of basis vectors ($\lambda\sigma \neq 1$, $\lambda \neq 0$, and $\sigma \neq 0$) in the 4-dimensional FNAA [8]

·	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_0$	$\lambda\mathbf{e}_0$	$\lambda\mathbf{e}_1$	$\mathbf{e}_0$	$\mathbf{e}_1$
$\mathbf{e}_1$	$\mathbf{e}_0$	$\mathbf{e}_1$	$\sigma\mathbf{e}_0$	$\sigma\mathbf{e}_1$
$\mathbf{e}_2$	$\lambda\mathbf{e}_2$	$\lambda\mathbf{e}_3$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_3$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\sigma\mathbf{e}_2$	$\sigma\mathbf{e}_3$

■ ***Table 2.*** Sparse BVMT ($\lambda \neq 0$) defining the 4-dimensional FNAA with global two-sided unit $\mathbf{E} = (1, 1, 0, 0)$ [7]

·	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_0$	$\mathbf{e}_0$	0	0	$\mathbf{e}_3$
$\mathbf{e}_1$	0	$\mathbf{e}_1$	$\mathbf{e}_2$	0
$\mathbf{e}_2$	$\mathbf{e}_2$	0	0	$\lambda\mathbf{e}_1$
$\mathbf{e}_3$	0	$\mathbf{e}_3$	$\lambda\mathbf{e}_0$	0

■ ***Table 3.*** Sparse BVMT ($\lambda \neq 0$) defining the 4-dimensional FNAA with global two-sided unit $\mathbf{E} = (0, 1, 1, 0)$ [16]

·	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_0$	0	0	$\mathbf{e}_0$	$\lambda\mathbf{e}_1$
$\mathbf{e}_1$	$\mathbf{e}_0$	$\mathbf{e}_1$	0	0
$\mathbf{e}_2$	0	0	$\mathbf{e}_2$	$\mathbf{e}_3$
$\mathbf{e}_3$	$\lambda\mathbf{e}_2$	$\mathbf{e}_3$	0	0

In addition to a faster multiplication operation, the 4-dimensional FNAAs defined by the sparse BVMTs are attractive to the fact that their detailed structure (in terms of decomposition into a set of commutative subalgebras) is known for the case of defining the algebras over the fields $GF(p)$ with arbitrary odd characteristics p. Besides, using the technique by [7, 8], one can show that, in the case of defining the algebras over the fields $GF(2^z)$, where $z > 0$, the 4-dimensional FNAAs set by Tables 1 and 2, possess the following common properties:

1) the 4-dimensional FNAA contains $2^{2z} + 2^z + 1$ of commutative subalgebras of the order 2^{2z}, every pair of which intersecting exactly in the set of scalar vectors $\{\mathbf{L}: \mathbf{L} = h\mathbf{E}, h = 0, 1, \ldots, 2^z - 1\}$, where $\mathbf{E}$ is the global two-sided unit;

2) the order of multiplicative group Γ of the algebra is equal to

$$\Omega = 2^z(2^{2z} - 1)(2^z - 1); \qquad (1)$$

3) the group Γ contains sufficiently large number ($> 2^z$) of commutative subgroups Γ_1 possessing

2-dimensional cyclicity (i. e., a minimum generator system of the subgroup Γ_1 contains two vectors of the same order) and having order equal to

$$\Omega_1 = (2^z - 1)^2; \quad (2)$$

4) the group Γ contains sufficiently large number ($> 2^z$) of commutative cyclic subgroups Γ_2 of the order

$$\Omega_2 = 2^{2z} - 1 = (2^z - 1)(2^z + 1); \quad (3)$$

5) the group Γ contains commutative cyclic subgroups Γ_3 of the order

$$\Omega_3 = 2^z(2^z - 1). \quad (4)$$

The condition of invertibility of some vector **A** in the FNAA set by Table 2 over a field $GF(p)$ [7] is also valid in the cased of defining the FNAA over the $GF(2^z)$ fields:

$$a_0a_1 \neq \lambda a_2a_3. \quad (5)$$

Similarly, we have the following invertibility condition for the FNAA set by Table 3 over the $GF(2^z)$ fields [16]:

$$a_1a_2 \neq \lambda a_0a_3. \quad (6)$$

For the 4-dimensional FNAA set by Table 1 over the $GF(2^z)$ fields (commutative groups of the Γ_1, Γ_2, and Γ_3 types are also contained in this algebra), from [8] one gets the invertibility condition

$$a_1a_2 \neq a_0a_3 \quad (7)$$

and the following formula for the two-sided global unit **E** depending on the structural constants λ and σ (that can be selected arbitrarily, but satisfying the conditions $\lambda\sigma \neq 1$, $\lambda \neq 0$, and $\sigma \neq 0$):

$$\mathbf{E} = \left(\frac{\sigma}{\sigma\lambda - 1}, \frac{1}{1 - \sigma\lambda}, \frac{1}{1 - \sigma\lambda}, \frac{\lambda}{\sigma\lambda - 1} \right). \quad (8)$$

To execute the exponentiation operation in FNAAs, i. e. for calculating the value $\mathbf{R} = \mathbf{W}^k$ (**W** is a vector; k is a non-negative integer), we propose the following modification of the fast-exponentiation algorithm, which is free of using the **E** value:

INPUT: **W** and $k > 0$.

1. Set $\mathbf{V} \leftarrow \mathbf{W}$, and $n \leftarrow k$.
2. If $n \bmod 2 = 1$, then go to step 4.
3. $\mathbf{V} \leftarrow \mathbf{V}^2$, $n \leftarrow n$ div 2, and go to step 2.
4. $\mathbf{R} \leftarrow \mathbf{V}$, $\mathbf{V} \leftarrow \mathbf{V}^2$, $n \leftarrow n$ div 2.
5. If $n = 0$, then STOP.
6. If $n \bmod 2 = 1$, then go to step 8.
7. $\mathbf{V} \leftarrow \mathbf{V}^2$, $n \leftarrow n$ div 2, and go to step 6.
8. $\mathbf{R} \leftarrow \mathbf{RV}$, $\mathbf{V} \leftarrow \mathbf{V}^2$, $n \leftarrow n$ div 2, and go to step 5.

OUTPUT: $\mathbf{R} = \mathbf{W}^k$.

Development of the algebraic signature algorithms with a hidden group, which are based on computational difficulty of the HDLP, is connected with the requirement of existence of a large-size prime factor of the order of the hidden group. Taking into account that the said algorithms use hidden groups which are subgroups of the commutative groups of the Γ_1 and Γ_2 types, one can recommend the values of z shown in Tables 4 and 5. The values z = 61, 89, 107, 127, 521, and 607 are Mersenne degrees that define prime values of $2^z - 1$.

Development of the algorithms with a hidden group, which are based on computational difficulty of solving a system of many quadratic equations with many unknowns, is free of the requirement of existence of a large-size prime factor of the order of the hidden group. Security of the algorithms of this type depends on the size of the order of the hidden group and is not dependent on the factorization of the order. However, to provide a higher performance the hidden group order should be free of small-size factors (for example, less than 20 bits). If the order of a group of the Γ_1-type is free of the said factors,

■ ***Table 4.*** The case of using the Γ_1-type and Γ_2-type commutative groups (or their subgroups) as a hidden group in the HDLP-based signature algorithms

Degree z	Number of prime factors of the value $2^z - 1$ (their size in bits)	Degree z	Number of prime factors of the value $2^z - 1$ (their size in bits)
61	1 (61)	281	2 (17 and 265)
89	1 (89)	373	2 (25 and 349)
107	1 (107)	421	2 (50 and 372)
127	1 (127)	457	2 (28 and 430)
131	2 (9 and 123)	521	1 (521)
197	2 (13 and 185)	607	1 (607)

■ ***Table 5.*** Additional values of z for the case of using the Γ_2-type commutative groups (or their subgroups) as a hidden group in the HDLP-based signature algorithms

Degree z	Number of prime factors of the value $2^z + 1$ (their size in bits)	Degree z	Number of prime factors of the value $2^z + 1$ (their size in bits)
101	1 (100)	311	2 (16 and 294)
127	1 (126)	313	1 (312)
179	2 (36 and 142)	347	1 (346)
199	1 (198)	433	2 (22 and 410)
229	2 (25 and 204)	–	–

■ ***Table 6.*** The case of using a hidden group of the Γ_1-type

Degree z	Number of β-bit prime ($\beta \geq 30$) factors of the value $2^z - 1$ (their size in bits)	Degree z	Number of β-bit prime ($\beta \geq 26$) factors of the value $2^z - 1$ (their size in bits)
89	Mersenne degree	257	3 (49, 80, and 129) [17]
101	2 (43 and 59) [17]	271	2 (34 and 238) [17]
103	2 (39 and 63) [17]	293	2 (86 and 208) [17]
107	Mersenne degree	307	4 (31, 42, 68, and 166)
109	2 (30 and 80) [17]	331	3 (44, 50, and 238) [17]
127	Mersenne degree	347	2 (74 and 274) [17]
137	2 (65 and 73) [17]	379	2 (38 and 342) [17]
139	2 (43 and 97) [17]	389	3 (26, 33, and 332) [17]
149	2 (67 and 83) [17]	421	2 (50 and 372)
173	3 (41, 56, and 78) [17]	433	4 (65, 80, 83, and 208)
199	2 (38 and 162) [17]	503	4 (52, 64, 71, and 318)

■ ***Table 7.*** The case of using a hidden group of the Γ_2-type

Degree z	Number of β-bit prime ($\beta \geq 36$) factors of the value $2^z + 1$ (their size in bits)	Degree z	Number of β-bit prime ($\beta \geq 22$) factors of the value $2^z + 1$ (their size in bits)
101	1 (100)	307	4 (31, 42, 68, and 166)
127	1 (126)	347	1 (346)
179	2 (36 and 142)	379	3 (44, 100, and 235)
199	1 (198)	389	4 (40, 51, 52, and 246)
257	3 (46, 69, and 142)	433	2 (22 and 410)
271	2 (45 and 231)	503	4 (52, 64, 71, and 318)

then each of them can be used as the hidden group of the designed signature scheme.

The order Ω_2 of the Γ_2 group contains factor 3 [see formula (3)]. If no other small-size factors are contained in $\Omega_2 = (2^z - 1)(2^z + 1)$, then the subgroup of the order $\Omega_2/3$ can be used as a hidden group. The values of z suitable for development of signature algorithms with a hidden group of the Γ_1-type, based on difficulty of solving a system of quadratic equations, are shown in Table 6 [17] (Γ_2-type – in Table 7). In the case of the hidden group of the Γ_2-type, one should use the z values that determine the absence of short divisors for the values $2^z - 1$ and $2^z + 1$ (except the two-bit divisor 3 for the second value).

When developing the signature schemes with a hidden group, it is assumed to use algorithms for generating a basis (minimum generator system) of the hidden group. For example, you can use the following algorithms.

Algorithm 1: generating a basis of the Γ_1-type group.

1. Using the invertibility condition [see formulas (5)–(7)], generate a random invertible vector $\mathbf{V}$ of the order $q = 2^z - 1$.

2. If the vector $\mathbf{V}$ is contained in the set of scalar vectors, i. e., if $\mathbf{V} = \sigma\mathbf{E}$ for some value $\sigma \in GF(2^z)$, then go to step 1.

3. Generate a random integer k ($0 < k < q$) and a random binary polynomial $\beta \in GF(2^z)$ of the order $2^z - 1$.

4. Compute the vector $\mathbf{H} = \beta\mathbf{V}^k$.

5. Output the pair of vectors $\mathbf{H}$ and $\mathbf{G} = \mathbf{V}$ as a basis $<\mathbf{G}, \mathbf{H}>$ of a random Γ_1-type group.

This algorithm works correctly, since the vectors of the order $2^z - 1$ in the groups of the Γ_2 and Γ_3-types are scalar vectors.

Algorithm 2: generating a basis of the Γ_2-type group.

1. Using the invertibility condition [see formulas (5)–(7)], generate a random invertible vector $\mathbf{V}$ of order the $q = (2^z - 1)(2^z + 1)$.

2. Output the vector $\mathbf{V}$ as a generator (basis $<\mathbf{V}>$) of a random Γ_2-type group.

This algorithm works correctly, since the Γ_1-type and Γ_3-type groups do not contain vectors of the order $(2^z - 1)(2^z + 1)$. Evidently, the vector $\mathbf{J} = \mathbf{V}^3$ is a generator of a commutative cyclic group Γ_2' of the order $q/3$, which is a subgroup of the Γ_2-type group generated by the vector $\mathbf{V}$.

The described 4-dimensional FNAAs are used as algebraic carrier of three new signature algorithms with a hidden group. Evidently the said FNAAs (set over $GF(2^z)$) could be used to update the known algorithms of such type, for example, described in [7, 13] (for the first type of the signature algorithms with a hidden group) and in [15] (for the second type of the signature algorithms with a hidden group). However, the authors prefer to illustrate existence of variety of possibilities, when designing algorithms with a hidden group.

A signature scheme based on HDLP

In this section it is introduced a HDLP-based signature algorithm (the first signature scheme) that illustrates the first type of the algebraic signa-

ture schemes with a hidden group. The development of various types of HDLP-based algorithms and methods for setting a hidden group formed the prerequisites on the basis of which the second type of signature algorithms with a hidden group was born. The reader can easily see the similar construction elements in the two types of the algorithms introduced in this paper (see also the next section).

Suppose a 4-dimensional FNAA is set by Table 2 over the field $GF(2^z)$, where $z = 521$ and $q = 2^z - 1$ is a prime number. Using a group of the Γ_1-type (set by some basis $<\mathbf{G}, \mathbf{H}>$), you can generate a public key in the form of three vectors **U**, **Y**, and **Z** as follows:

1. Generate two random invertible vectors **A** and **B** of the order $\omega \geq p - 1$, satisfying the conditions $\mathbf{AB} \neq \mathbf{BA}$, $\mathbf{AG} \neq \mathbf{GA}$, $\mathbf{BG} \neq \mathbf{GB}$.

2. Generate two random integers $x < q$ and $u < q$. Then calculate the first element **U** of the public key: $\mathbf{U} = \mathbf{AG}^x\mathbf{H}^u\mathbf{B}^{-1}$.

3. Calculate the second element **Y** of the public key: $\mathbf{Y} = \mathbf{BGB}^{-1}$.

4. Calculate the third element **Z** of the public key: $\mathbf{Z} = \mathbf{BHA}^{-1}$.

The pair of numbers (x, u) and the vectors **G**, **H**, **A**, and **B** compose a secret key (having size ≈ 1173 bytes) and are used for generating a signature to some electronic document M. The size of the public key represented by the triple of vectors (**U**, **Y**, **Z**) is equal to ≈ 782 bytes.

Algorithm for generating a signature.

1. Generate a random natural integer $k < q$ and calculate the vector $\mathbf{K} = \mathbf{G}^k$.

2. Generate a random natural integer $t < q$ and calculate the vector $\mathbf{R} = \mathbf{AKH}^t\mathbf{A}^{-1}$.

3. Using a specified 521-bit hash function f, calculate the first signature element e as a hash-function value from the document M to which the vector **R** is concatenated: $e = f(M, \mathbf{R})$.

4. Compute the second signature element s:

$$s = \sqrt{\frac{1}{e}\left(k - \frac{tx}{u+1}\right)} \bmod q.$$

5. If the value under the root is a quadratic non-residue modulo q, then go to step 2.

6. Compute the third signature element d:

$$d = \left(\frac{t}{s(u+1)} - 1\right) \bmod q.$$

This algorithm outputs a 196-byte signature in the form of a triple of 521-bit integers (e, s, d). Computational difficulty of the signature generation algorithm is defined mainly by exponentiation operations performed at steps 1 and 2. It is easy to see that on the average three exponentiations in the FNAA used as algebraic support (≈ 18 432 multiplications in $GF(2^{521})$) are executed to generate one signature.

Algorithm for verifying a signature.

1. Calculate the vector

$$\mathbf{R}' = \left(\mathbf{UY}^{es}\mathbf{Z}(\mathbf{UZ})^d\right)^s.$$

2. Calculate the hash-function value from the document M to which the vector $\mathbf{R}'$ is concatenated: $e' = f(M, \mathbf{R}')$.

3. If $e' = e$, then the signature is accepted as a genuine one. Otherwise the signature is rejected.

Computational difficulty of the signature verification procedure can be estimate as three exponentiations in the 4-dimensional FNAA used as algebraic support (≈ 18 432 multiplications in $GF(2^{521})$).

Correctness proof of the described signature scheme is as follows (see formulas used at steps 4 and 6 of the signature generation algorithm):

$$\mathbf{R}' = \left(\mathbf{UY}^{es}\mathbf{Z}(\mathbf{UZ})^d\right)^s = \left(\mathbf{AG}^x\mathbf{H}^u\mathbf{B}^{-1}\mathbf{BG}^{es} \times \right.$$

$$\left. \times \mathbf{B}^{-1}\mathbf{BHA}^{-1}\left(\mathbf{AG}^x\mathbf{H}^u\mathbf{B}^{-1}\mathbf{BHA}^{-1}\right)^d\right)^s =$$

$$= \left(\mathbf{AG}^{x+es}\mathbf{H}^{u+1}\mathbf{A}^{-1}\mathbf{AG}^{xd}\mathbf{H}^{d(u+1)}\mathbf{A}^{-1}\right)^s =$$

$$= \left(\mathbf{AG}^{x+es+xd}\mathbf{H}^{u+1+d(u+1)}\mathbf{A}^{-1}\right)^s =$$

$$= \mathbf{AG}^{xs(d+1)+es^2}\mathbf{H}^{sd(u+1)+s(u+1)}\mathbf{A}^{-1} =$$

$$= \mathbf{AG}^{xs\frac{t}{s(u+1)}+e\frac{1}{e}\left(k-\frac{tx}{u+1}\right)}\mathbf{H}^{\left(\frac{t}{u+1}-s\right)(u+1)+s(u+1)} \times$$

$$\times \mathbf{A}^{-1} = \mathbf{AG}^k\mathbf{H}^t\mathbf{A}^{-1} = \mathbf{R} \Rightarrow$$

$$\Rightarrow f(M, \mathbf{R}') = f(M, \mathbf{R}) \Rightarrow e' = e.$$

A critical point of the consideration of the HDLP-based signature algorithms as candidates for post-quantum cryptoschemes is potential possibility of using algebraic methods to reduce the HDLP to ordinary DLP. Therefore, the second-type algebraic signature schemes with a hidden group, which are based on computational difficulty of solving a system of many quadratic equations with many unknowns (the problem for solving of which the quantum computer is not efficient), can be estimated as a more preferable candidates for post-quantum signature schemes.

Signature schemes based on difficulty of solving a system of many quadratic equations

The second proposed signature scheme is described as follows. Suppose the 4-dimensional FNAA is set by Table 3 over the field $GF(2^z)$, where $z = 257$.

Then, generating a random secret basis <**G**, **H**> of a group of the Γ_1-type one can generate a public key in the form of six vectors ($\mathbf{Y}_1$, $\mathbf{Z}_1$, $\mathbf{Y}_2$, $\mathbf{Z}_2$, $\mathbf{Y}_3$, $\mathbf{Z}_3$, **T**) as follows.

Public-key generation algorithm.

1. Using the invertibility condition (6), generate at random invertible vectors **A**, **B**, **D**, and **F** satisfying the following non-equalities: $\mathbf{AB} \neq \mathbf{BA}$, $\mathbf{AD} \neq \mathbf{DA}$, $\mathbf{AF} \neq \mathbf{FA}$, $\mathbf{AG} \neq \mathbf{GA}$, $\mathbf{BD} \neq \mathbf{DB}$, $\mathbf{BF} \neq \mathbf{FB}$, $\mathbf{BG} \neq \mathbf{GB}$, $\mathbf{DF} \neq \mathbf{FD}$, $\mathbf{DG} \neq \mathbf{GD}$, and $\mathbf{FG} \neq \mathbf{GF}$.

2. Calculate the vectors $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$, $\mathbf{D}^{-1}$, and $\mathbf{F}^{-1}$.

3. Generate non-negative integers $x < q$ and $w < q$, where $q = 2^z - 1$ is a 256-bit number that is product of three primes having the size 49, 80, and 129 bits (see Table 6). Then compute the public key ($\mathbf{Y}_1$, $\mathbf{Z}_1$, $\mathbf{Y}_2$, $\mathbf{Z}_2$, $\mathbf{Y}_3$, $\mathbf{Z}_3$, **T**) by formulas

$$\mathbf{Y}_1 = \mathbf{AGB};\ \mathbf{Z}_1 = \mathbf{DHA}^{-1};$$

$$\mathbf{Y}_2 = \mathbf{FH}^x\mathbf{B};\ \mathbf{Z}_2 = \mathbf{DH}^w\mathbf{GF}^{-1};$$

$$\mathbf{Y}_3 = \mathbf{AG}^w\mathbf{B};\ \mathbf{Z}_3 = \mathbf{DHGF}^{-1};\ \mathbf{T} = \mathbf{DHG}^x\mathbf{B}. \quad (9)$$

The secret key (with total size ≈ 833 bytes) represents two integers x, w and six vectors **G**, **H**, **A**, **B**, **D**, and **F**. The size of public key is equal to ≈ 900 bytes. Computation of a signature to some electronic document M is performed, using the following algorithm.

Signature generation algorithm.

1. Generate at random two natural numbers k ($k < q$) and t ($t < q$). Then calculate the vector

$$\mathbf{R} = \mathbf{AG}^k\mathbf{H}^t\mathbf{F}^{-1}. \quad (10)$$

2. Using a specified $2z$-bit hash function f, calculate the first signature element e as a hash-function value from the document M to which the vector **R** is concatenated: $e = e_1||e_2 = f(M, \mathbf{R})$, where the hash-value e is represented as concatenation of two z-bit integers e_1 and e_2.

3. If the integers $2e_1 + e_2 + 1$ and q are not mutually prime, then go to step 1. Otherwise, calculate the natural numbers n and d:

$$n = \frac{k - e_1 - xe_1 - e_2 - w - 1}{2e_1 + e_2 + 1} \bmod q; \quad (11)$$

$$d = \frac{t - 2e_1 - xe_2 - we_2 - 1}{2e_1 + e_2 + 1} \bmod q. \quad (12)$$

4. Calculate the second signature element in the form of the vector **S**:

$$\mathbf{S} = \mathbf{B}^{-1}\mathbf{G}^n\mathbf{H}^d\mathbf{D}^{-1}. \quad (13)$$

Since the integer q contains three factors of sufficiently large size (≥ 49 bits), the probability of repeating the first step of the algorithm is negligible. Therefore, the computational complexity of this algorithm is determined mainly by 4 exponentiations in the used FNAA (≈ $48z$ = 12 336 multiplications in $GF(2^z)$). The size of the signature (e, **S**) is equal to $6z$ bits (≈ 193 bytes). Verification of the signature is performed, using the public key ($\mathbf{Y}_1$, $\mathbf{Z}_1$, $\mathbf{Y}_2$, $\mathbf{Z}_2$, $\mathbf{Y}_3$, $\mathbf{Z}_3$, **T**) and the following procedure.

Signature verification algorithm.

1. Compute the vector $\mathbf{R}'$ by the following formula with four entries of the signature element **S**:

$$\mathbf{R}' = \left(\mathbf{Y}_1\mathbf{STSZ}_1\right)^{e_1}\mathbf{Y}_3\mathbf{SZ}_3\left(\mathbf{Y}_2\mathbf{SZ}_2\right)^{e_2}. \quad (14)$$

2. Calculate the hash-value e' from the document to which the vector $\mathbf{R}'$ is concatenated: $e' = f(M, \mathbf{R}')$.

3. If $e' = e$, then the signature is genuine. Otherwise the signature is rejected.

The computational complexity of the signature verification algorithm is determined mainly by 2 exponentiations in the used FNAA (≈ $24z$ = 6168 multiplications in $GF(2^z)$).

Correctness of the signature scheme is proven as follows.

Signature scheme correctness proof.

Compute the vectors

$$\mathbf{J}_1 = \left(\mathbf{Y}_1\mathbf{STSZ}_1\right)^{e_1} = \left(\mathbf{AGBB}^{-1}\mathbf{G}^n\mathbf{H}^d\mathbf{D}^{-1} \times \right.$$

$$\left. \times\ \mathbf{DG}^x\mathbf{HBB}^{-1}\mathbf{G}^n\mathbf{H}^d\mathbf{D}^{-1}\mathbf{DHA}^{-1}\right)^{e_1} =$$

$$= \left(\mathbf{AGG}^n\mathbf{H}^d\mathbf{G}^x\mathbf{HG}^n\mathbf{H}^d\mathbf{HA}^{-1}\right)^{e_1} =$$

$$= \left(\mathbf{AG}^{2n+x+1}\mathbf{H}^{2d+2}\mathbf{A}^{-1}\right)^{e_1} =$$

$$= \mathbf{AG}^{2ne_1+xe_1+e_1}\mathbf{H}^{2de_1+2e_1}\mathbf{A}^{-1};$$

$$\mathbf{J}_2 = \mathbf{Y}_3\mathbf{SZ}_3 = \mathbf{AG}^w\mathbf{BB}^{-1}\mathbf{G}^n\mathbf{H}^d\mathbf{D}^{-1}\mathbf{DHGF}^{-1} =$$

$$= \mathbf{AG}^{n+w+1}\mathbf{H}^{d+1}\mathbf{F}^{-1};$$

$$\mathbf{J}_3 = \left(\mathbf{Y}_2\mathbf{SZ}_2\right)^{e_2} =$$

$$= \left(\mathbf{FH}^x\mathbf{BB}^{-1}\mathbf{G}^n\mathbf{H}^d\mathbf{D}^{-1}\mathbf{DH}^w\mathbf{GF}^{-1}\right)^{e_2} =$$

$$= \left(\mathbf{FG}^{n+1}\mathbf{H}^{d+x+w}\mathbf{F}^{-1}\right)^{e_2} =$$

$$= \mathbf{FG}^{ne_2+e_2}\mathbf{H}^{de_2+xe_2+we_2}\mathbf{F}^{-1}.$$

Then compute the vector $\mathbf{R}'$:

$$\mathbf{R}' = \mathbf{J}_1\mathbf{J}_2\mathbf{J}_3 = \mathbf{AG}^{2ne_1+xe_1+e_1}\mathbf{H}^{2de_1+2e_1}\mathbf{A}^{-1} \times$$

$$\times\ \mathbf{AG}^{n+w+1}\mathbf{H}^{d+1}\mathbf{F}^{-1} \times \mathbf{FG}^{ne_2+e_2}\mathbf{H}^{de_2+xe_2+we_2}\mathbf{F}^{-1} =$$

$$= \mathbf{AG}^{2ne_1+xe_1+e_1+n+w+1+ne_2+e_2} \times$$

$$\times\ \mathbf{H}^{2de_1+2e_1+d+1+de_2+xe_2+we_2}\mathbf{F}^{-1} =$$

$$= \mathbf{AG}^{n(2e_1+e_2+1)+e_1+xe_1+e_2+w+1} \times$$

$$\times\ \mathbf{H}^{d(2e_1+e_2+1)+2e_1+xe_2+we_2+1}\mathbf{F}^{-1}.$$

Taking into account the formulas (11) and (12) we get:

$$\mathbf{R}' = \mathbf{A}\mathbf{G}^k\mathbf{H}^t\mathbf{F}^{-1} = \mathbf{R} \Rightarrow$$
$$\Rightarrow f\left(M \,\|\, \mathbf{R}'\right) = f\left(M \,\|\, \mathbf{R}\right) \Rightarrow e' = e.$$

The final equality means validity of the input signature.

Security of the described signature scheme is based on computational difficulty of solving the system of 13 vector quadratic equations with the following 11 unknowns: **A**, **B**, **D**, **F**, **G**, **H**, $\mathbf{H}' = \mathbf{H}^x$, $\mathbf{H}'' = \mathbf{H}^w\mathbf{G}$, $\mathbf{G}' = \mathbf{G}^w$, $\mathbf{G}'' = \mathbf{G}\mathbf{H}$, and $\mathbf{G}''' = \mathbf{G}^x\mathbf{H}$, which are determined by the formulas (9) and the pair-wise permutability of the unknowns **G**, **H**, **H**′, **H**″, **G**′, **G**″, and **G**‴: $\mathbf{G}\mathbf{H} = \mathbf{H}\mathbf{G}$, $\mathbf{G}\mathbf{H}' = \mathbf{H}'\mathbf{G}$, $\mathbf{G}\mathbf{H}'' = \mathbf{H}''\mathbf{G}$, $\mathbf{G}\mathbf{G}' = \mathbf{G}''\mathbf{G}$, $\mathbf{G}\mathbf{G}'' = \mathbf{G}'\mathbf{G}$, and $\mathbf{G}\mathbf{G}''' = \mathbf{G}'''\mathbf{G}$. Using Table 3, the latter system reduces to a system of 52 quadratic equations (with 44 unknowns) in the field $GF(2^z)$.

A remarkable feature of the algebraic algorithms with a hidden group is the multiple entries of the signature element **S** in the vector verification equation set over a non-commutative algebra. This provides resistance to forging signature attacks base on using the value **S** as a fitting parameter. In the algorithm describe above we have four entries of the vector **S**. The number η of entries should satisfy the condition $\eta \geq 2$. The next digital signature scheme uses the value $\eta = 3$.

The third developed signature scheme is described as follows. Suppose the 4-dimensional FNAA is set by Table 1 over the field $GF(2^z)$, where $z = 199$ (see Tables 6 and 7). Then, generating a random secret basis <**G**> of a cyclic group of the Γ_2'-type (subgroup of a Γ_2-type group), which has order $q = \Omega_2/3 = 3^{-1}(2^z - 1)(2^z + 1)$, one can generate a public key in the form of seven vectors ($\mathbf{Y}_1$, $\mathbf{Z}_1$, $\mathbf{U}_1$, $\mathbf{Y}_2$, $\mathbf{Z}_2$, $\mathbf{U}_2$, **V**) as follows.

Public-key generation algorithm.

1. Using the invertibility condition (8), generate at random invertible vectors **A**, **B**, **D**, and **F** satisfying the following non-equalities: $\mathbf{AB} \neq \mathbf{BA}$, $\mathbf{AD} \neq \mathbf{DA}$, $\mathbf{AF} \neq \mathbf{FA}$, $\mathbf{AG} \neq \mathbf{GA}$, $\mathbf{BD} \neq \mathbf{DB}$, $\mathbf{BF} \neq \mathbf{FB}$, $\mathbf{BG} \neq \mathbf{GB}$, $\mathbf{FD} \neq \mathbf{DF}$, and $\mathbf{GF} \neq \mathbf{FG}$.

2. Calculate the vectors $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$, $\mathbf{D}^{-1}$, and $\mathbf{F}^{-1}$.

3. Calculate the vector $\mathbf{J} = \mathbf{G}^{q\left(2^z-1\right)^{-1}}$ of the order $q' = 3^{-1}(2^z + 1)$ and the vector $\mathbf{I} = \mathbf{G}^{3q\left(2^z+1\right)^{-1}}$ of the order $q'' = 2^z - 1$.

4. Generate at random non-negative integers x ($x < q'$) and w ($w < q''$), where q' is a 198-bit prime number and q'' is a product of two primes having the size 38 and 162 bits (see Tables 6 and 7). Then compute the public key ($\mathbf{Y}_1$, $\mathbf{Z}_1$, $\mathbf{U}_1$, $\mathbf{Y}_2$, $\mathbf{Z}_2$, $\mathbf{U}_2$, **V**) by formulas

$$\mathbf{Y}_1 = \mathbf{B}^{-1}\mathbf{J}\mathbf{A}^{-1};\ \mathbf{Z}_1 = \mathbf{B}^{-1}\mathbf{I}\mathbf{B};\ \mathbf{U}_1 = \mathbf{B}^{-1}\mathbf{J}^x\mathbf{F}^{-1};$$
$$\mathbf{Y}_2 = \mathbf{D}\mathbf{J}\mathbf{I}\mathbf{A}^{-1};\ \mathbf{Z}_2 = \mathbf{F}\mathbf{J}^w\mathbf{I}\,\mathbf{D}^{-1};$$
$$\mathbf{U}_2 = \mathbf{D}\mathbf{J}\mathbf{I}^x\mathbf{A}^{-1};\ \mathbf{V} = \mathbf{B}^{-1}\mathbf{I}^w\mathbf{D}^{-1}. \qquad (15)$$

The secret key (with total size ≈ 650 bytes) represents two integers x, w and six vectors **J**, **I**, **A**, **B**, **D**, and **F**. The size of public key is equal to ≈ 700 bytes. Computation of a signature to some electronic document M is performed, using the following algorithm.

Signature generation algorithm.

1. Generate at random two natural numbers k ($k < q'$) and t ($t < q''$). Then calculate the vector

$$\mathbf{R} = \mathbf{F}\mathbf{J}^k\mathbf{I}^t\mathbf{F}^{-1}. \qquad (16)$$

2. Using a specified $3z$-bit hash function f, calculate the first signature element e as a hash-function value from the document M to which the vector **R** is concatenated: $e = e_1||e_2||e_3 = f(M, \mathbf{R})$, where the hash-value e is represented as concatenation of tree z-bit integers e_1, e_2, and e_3.

3. If the integer $e_1e_2e_3 + e_2e_3 + e_3$ is not mutually prime with q' or with q'', then go to step 1. Otherwise, calculate the natural numbers n and d:

$$n = \left(\frac{k - we_3 - xe_3}{e_1e_2e_3 + e_2e_3 + e_3} - 1\right) \bmod q'; \qquad (17)$$

$$d = \left(\frac{t - we_2e_3 - xe_3}{e_1e_2e_3 + e_2e_3 + e_3} - 1\right) \bmod q''. \qquad (18)$$

4. Calculate the second signature element in the form of the vector **S**:

$$\mathbf{S} = \mathbf{A}\mathbf{J}^n\mathbf{I}^d\mathbf{B}. \qquad (19)$$

Since the integer q' is prime and q'' contains two factors of sufficiently large size (38 and 162 bits), the probability of repeating the first step of the algorithm is negligible. Therefore, the computational complexity of this algorithm is determined mainly by 4 exponentiations in the used FNAA (≈ $96z$ = 19 104 multiplications in $GF(2^z)$). The size of the signature (e, **S**) is equal to ≈ $7z$ bits (≈ 175 bytes). Verification of the signature is performed, using the public key ($\mathbf{Y}_1$, $\mathbf{Z}_1$, $\mathbf{U}_1$, $\mathbf{Y}_2$, $\mathbf{Z}_2$, $\mathbf{U}_2$, **V**) and the following procedure.

Signature verification algorithm.

1. Compute the vector **R**′ by the following formula with three entries of the signature element **S**:

$$\mathbf{R}' = \left[\mathbf{Z}_2\left(\mathbf{Y}_2\mathbf{S}\left(\mathbf{Y}_1\mathbf{S}\mathbf{Z}_1\right)^{e_1}\mathbf{V}\right)^{e_2}\mathbf{U}_2\mathbf{S}\mathbf{U}_1\right]^{e_3}. \qquad (20)$$

2. Calculate the hash-value e' from the document to which the vector **R**′ is concatenated: $e' = f(M, \mathbf{R}')$.

3. If $e' = e$, then the signature is genuine. Otherwise the signature is rejected.

The computational complexity of the signature verification algorithm is determined mainly by 3 exponentiations in the used FNAA ($\approx 72z$ = = 14 328 multiplications in $GF(2^z)$).

Correctness of the latter signature scheme is proven as follows.

Signature scheme correctness proof.

Calculate the values $\mathbf{X}_1$ and $\mathbf{X}_2$:

$$\mathbf{X}_1 = \left(\mathbf{Y}_1\mathbf{S}\mathbf{Z}_1\right)^{e_1} = \left(\mathbf{B}^{-1}\mathbf{J}\mathbf{A}^{-1}\mathbf{A}\mathbf{J}^n\mathbf{I}^d\mathbf{B}\mathbf{B}^{-1}\mathbf{I}\mathbf{B}\right)^{e_1} =$$

$$= \mathbf{B}^{-1}\mathbf{J}^{e_1 n+e_1}\mathbf{I}^{e_1 d+e_1}\mathbf{B};$$

$$\mathbf{X}_2 = \left(\mathbf{Y}_2\mathbf{S}\mathbf{X}_1\mathbf{V}\right)^{e_2} =$$

$$= \left(\mathbf{D}\mathbf{J}\mathbf{I}\mathbf{A}^{-1}\mathbf{A}\mathbf{J}^n\mathbf{I}^d\mathbf{B}\mathbf{B}^{-1}\mathbf{J}^{e_1 n+e_1}\mathbf{I}^{e_1 d+e_1}\mathbf{B}\mathbf{B}^{-1}\mathbf{I}^w\mathbf{D}^{-1}\right)^{e_2} =$$

$$= \left(\mathbf{D}\mathbf{J}^{n(e_1+1)+e_1+1}\mathbf{I}^{d(e_1+1)+e_1+w+1}\mathbf{D}^{-1}\right)^{e_2} =$$

$$= \mathbf{D}\mathbf{J}^{n(e_1e_2+e_2)+e_1e_2+e_2}\mathbf{I}^{d(e_1e_2+e_2)+e_1e_2+we_2+e_2}\mathbf{D}^{-1}.$$

Then compute the vector $\mathbf{R}'$:

$$\mathbf{R}' = \left[\mathbf{Z}_2\mathbf{X}_2\mathbf{U}_2\mathbf{S}\mathbf{U}_1\right]^{e_3} =$$

$$= \begin{pmatrix} \mathbf{F}\mathbf{J}^w\mathbf{I}\mathbf{D}^{-1}\mathbf{D}\mathbf{J}^{n(e_1e_2+e_2)+e_1e_2+e_2} \times \\ \times\, \mathbf{I}^{d(e_1e_2+e_2)+e_1e_2+we_2+e_2} \times \\ \times\, \mathbf{D}^{-1}\mathbf{D}\mathbf{J}\mathbf{I}^x\mathbf{A}^{-1}\mathbf{A}\mathbf{J}^n\mathbf{I}^d\mathbf{B}\mathbf{B}^{-1}\mathbf{J}^x\mathbf{F}^{-1} \end{pmatrix}^{e_3} =$$

$$= \begin{pmatrix} \mathbf{F}\mathbf{J}^{n(e_1e_2+e_2+1)+e_1e_2+e_2+w+x+1} \times \\ \times\, \mathbf{I}^{d(e_1e_2+e_2+1)+e_1e_2+we_2+e_2+x+1}\mathbf{F}^{-1} \end{pmatrix}^{e_3} =$$

$$= \mathbf{F}\mathbf{J}^{n(e_1e_2e_3+e_2e_3+e_3)+e_1e_2e_3+e_2e_3+we_3+xe_3+e_3} \times$$

$$\times\, \mathbf{I}^{d(e_1e_2e_3+e_2e_3+e_3)+e_1e_2e_3+we_2e_3+e_2e_3+xe_3+e_3}\mathbf{F}^{-1}.$$

Taking into account the formulas (17) and (18) we get:

$$\mathbf{R}' = \mathbf{A}\mathbf{J}^k\mathbf{I}^t\mathbf{A}^{-1} = \mathbf{R} \Rightarrow f\left(M, \mathbf{R}'\right) = f\left(M, \mathbf{R}\right) \Rightarrow e' = e,$$

where the latter equality proves the correct performance of the signature scheme.

Discussion

In this paper, the first developed signature algorithm, based on HDLP, is considered as an illustration of signature schemes attributed to the first type of the algebraic signature algorithms with a hidden group. Comparison with the second-type algorithms shows that in the both cases the main operations used to generate the public key, to generate a signature, and to verify the signature are exponentiation operations. However, the signature algorithms of the second type have principal difference, namely, they are based on computational complexity of finding a solution of a system of many quadratic equations with many unknowns. To solve the latter problem, the quantum computer is not efficient [18]. This fact is used in the area of multivariate cryptography that is one of the directions in the development of post-quantum public-key cryptographic algorithms. The multivariate cryptography was initiated by the paper [19] in 1988.

Over the past 30 years of the research in the field of multivariate cryptography many multivariate signature algorithms are currently known. A merit of the multivariate signature schemes is small size of the signature. Unfortunately, their significant drawback is a very large size of the public key. The latter is associated with a specific method for developing the multivariate signature algorithms, including generation of the public key as a set of quadratic (or cubic) polynomials that describe a trapdoor one-way mapping of vectors of large dimensions (from 30 to 200), given over a finite field of sufficiently small order (from 2^2 to 2^{16}).

At present the cryptographic community has well worked out the basic methods for cryptanalysis of the multivariate-cryptography algorithms. The following two types of attacks are distinguished [18]: i) direct attacks based on the algorithms for solving systems of many power (quadratic in many cases) equations with many unknowns and ii) structural attacks that use the structural features of the cryptoscheme design.

Because of significantly different design of the signature algorithms with a hidden group and the multivariate-cryptography algorithms the structural attacks developed for cryptanalysis of the latter are hardly applicably to the former and novel types of structural attacks are to be developed. Therefore, for preliminary security estimation of the second and third proposed algebraic signature algorithms the known direct attacks can be considered. The most effective direct attack is the use of algorithms for solving systems of many power equations, based on the calculation of the Gröbner basis [20, 21]. Table 8 computed on the base of the results of the papers [20, 21] can be used to estimate security W of the introduced algebraic algorithms with a hidden group to the direct attack.

Security of the second introduced signature scheme (algorithm with the value $\eta = 4$) is based on difficulty of solving the system of 12 vector quadratic equations with 11 unknowns $\mathbf{A}$, $\mathbf{B}$, $\mathbf{D}$, $\mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$, $\mathbf{G}' = \mathbf{G}^w$, $\mathbf{H}' = \mathbf{H}^x$, $\mathbf{H}'' = \mathbf{H}^w$, $\mathbf{G}''' = \mathbf{G}\mathbf{H}$, and $\mathbf{G}'' = \mathbf{G}^w\mathbf{H}$, which are determined by the formulas (9) and the pair-wise permutability of the unknowns $\mathbf{G}$, $\mathbf{H}$, $\mathbf{G}'$, $\mathbf{H}'$, $\mathbf{H}''$, $\mathbf{G}''$, and $\mathbf{G}'''$: $\mathbf{G}\mathbf{H} = \mathbf{H}\mathbf{G}$, $\mathbf{G}\mathbf{G}' = \mathbf{G}'\mathbf{G}$, $\mathbf{G}\mathbf{H}' = \mathbf{H}'\mathbf{G}$, $\mathbf{G}\mathbf{H}'' = \mathbf{H}''\mathbf{G}$, and $\mathbf{G}\mathbf{G}'' = \mathbf{G}''\mathbf{G}$. Using

Table 3, the latter system reduces to a system of $\mu = 48$ quadratic equations (with $\delta = 44$ unknowns) in the field $GF(2^{257})$.

Security of the third developed signature scheme (algorithm with the value $\eta = 3$) is based on difficulty of solving the system of 13 vector quadratic equations with the unknowns $\mathbf{A}$, $\mathbf{B}$, $\mathbf{J}$, $\mathbf{I}$, $\mathbf{J}' = \mathbf{J}^x$, $\mathbf{I}' = \mathbf{I}^w$, which are determined by the formulas (15) and the pair-wise permutability of the following 11 unknowns $\mathbf{J}$, $\mathbf{I}$, $\mathbf{J}'$, and $\mathbf{H}'$: $\mathbf{JI} = \mathbf{IJ}$, $\mathbf{JI}' = \mathbf{I}'\mathbf{J}$, $\mathbf{JJ}' = \mathbf{J}'\mathbf{J}$. Using Table 1, the latter system reduces to a system of 52 quadratic equations (with 44 unknowns) in the field $GF(2^{199})$.

Thus, one can take the number of equations μ equal to the number of the unknowns δ, and use the recommended minimum values of μ presented in the Table 8 for different values of the order of the field $GF(n)$ in which the system of quadratic equation is given. Since the system of quadratic equations related to the proposed signature algorithms is set in the fields $GF(2^z)$, where $2^z \gg 256$, one can use the values μ that relates to the case $n = 256$. In this case, we get overstated requirements for the minimum value, however, this overestimation can be considered insignificant due to relatively weak dependence on the value n. For the second and third proposed signature algorithms we get the value $W > 2^{128}$.

Since the value μ is proportional to the FNAA dimension, one can propose an evident way to improve the value W that is using six-dimensional and eight-dimensional FNAAs (set over the fields $GF(2^z)$ with smaller values of z) as algebraic support of the proposed signature algorithms, however, this way is connected with the study of the decomposition of the said FNAAs into the set of commutative subalgebras or to provide another method for justifying existence of sufficiently large number of commutative groups of a certain type. Potentially, using the 8-dimensional FNAAs as algebraic support of the second and third proposed signature algorithms for each of latter one gets the values $\mu = 104$, $\delta = 88$ and $W > 2^{192}$.

In the developed signature scheme with $\eta = 4$ the vectors $\mathbf{G}'$, $\mathbf{H}'$, $\mathbf{H}''$, $\mathbf{G}''$, and $\mathbf{G}'''$ are computed as $\mathbf{G}' = \mathbf{G}^w$, $\mathbf{H}' = \mathbf{H}^x$, $\mathbf{H}'' = \mathbf{H}^w$, $\mathbf{G}'' = \mathbf{G}^w\mathbf{H}$, and $\mathbf{G}''' = \mathbf{GH}$, correspondingly. This technique improves the performance of the signature generation algorithm. Actually, when generating a signature, you can select at random the vectors $\mathbf{G}'$, $\mathbf{H}'$, $\mathbf{H}''$, $\mathbf{G}''$, and $\mathbf{G}'''$ from the hidden group and use an alternative signature generation algorithm with many additional exponentiation operations (the reader can easily compose such algorithm), while the signature verification algorithm retains its original form. The analogous remark is valid for the algorithm with $\eta = 4$ entries of the $\mathbf{S}$ signature element in the signature verification equation. The noted remark clearly shows that the exponentiation operations are used as a part of the mechanism for calculating the signature element $\mathbf{S}$ that satisfies the verification equation with its multiple occurrences (entries) in the latter.

Table 9 shows a rough comparison of the developed post-quantum signature algorithms with the algorithms selected as finalists of the NIST world competition on the development of the post-quantum public-key algorithms [22]. Table 10 (where W denotes security to direct attack, which is estimated using Table 8) shows a rough comparison of the introduced signature algorithms based on computational difficulty of solving a system of quadratic equations with some known multivariate signature algorithms. The post-quantum algorithms introduced in this article have a significant advantage in the sizes of the signature and

■ ***Table 8.*** Minimum number of equations providing a given security level to the direct attack for different values of the order of the field $GF(n)$ in the case $\mu = \delta$ [18]

n	W				
	2^{80}	2^{100}	2^{128}	2^{192}	2^{256}
16	30	39	51	80	110
31	28	36	48	75	103
256	26	33	43	68	93

■ ***Table 9.*** Comparison with some known digital signature algorithms

Signature scheme	Signature size, bytes	Public key size, bytes	Signature generation rate, arb. un.	Signature verification rate, arb. un.
Falcon [23]	1280	1793	50	25
CRYSTALS-Dilithium [24]	2701	1472	15	2
Rainbow [25] (3 different versions)	66… 204	> 150 000 … > 1 900 000	–	–
The first proposed (HDLP-based)	196	782	25	25
The second proposed ($\eta = 4$)	193	900	150	300
The third proposed ($\eta = 3$)	175	700	150	200

■ ***Table 10.*** Comparison with some known digital signature algorithms

Signature algorithm	Signature size, bytes	Public key size, bytes	# quadratic equations μ (unknowns δ)	Order of the field over which the quadratic equations are set	W
[5]	–	–	27 (27)	2^{16}	$\approx 2^{80}$
Rainbow [26]	33	16 065	27 (33)	2^{8}	$\approx 2^{80}$
QUARTZ [6]	16	72 704	100 (107)	2^{4}	$> 2^{192}$
Rainbow [25] (3 different versions)	66… 204	> 150 000 … > 1 900 000	64 (96)… 128 (204)	2^{4}, 31, 2^{8}	2^{128}… 2^{256}
With a hidden group [16] η = 2	160	512	28 (28)	$> 2^{256}$	$\approx 2^{80}$
The second proposed (η = 4)	193	900	52 (44)	2^{257}	$> 2^{128}$
The third proposed (η = 3)	175	700	52 (44)	2^{199}	$> 2^{128}$

public key. Besides, the developed algebraic algorithms based on computational difficulty of solving a system of quadratic equations have significantly higher performance than finalists Falcon [23] and CRYSTALS-Dilithium [24]. However, a detailed security estimation of the introduced signature algorithms are to be performed as an independent research work.

The signature schemes with a hidden group, which are based on computational difficulty of solving a system of many quadratic equations, suite well for using the 6-dimensional and 8-dimensional FNAAs as algebraic support. The latter allows to define the FNAAs over the fields $GF(2^z)$ with comparatively small values of z. For composing the BVMTs defining the FNAAs of such dimensions, you can use the unified methods by [27, 28]. Using the FNAAs with a large set of global single-sided units (see, for example, [29]) as algebraic support of the signature algorithms with a hidden group also represent an item of a future study.

It should be noted that in passing to using FNAAs with a higher dimension value m (in order to get a higher security to the direct attack) as an algebraic support, we have the possibility to define algebras over the fields $GF(2^z)$ with lower degrees of z (for example, $z = 101$ and $z = 128$; see Tables 6 and 7). For a fixed value m, a decrease in the value of z has little effect on the resistance to direct attacks, however, we assume that this will lead to a significant decrease in the resistance to potential structural attacks. For this reason, sufficiently large values of z are used in the developed signature algorithms on the four-dimensional FNAAs.

The results of this study complement the results of the papers [14, 16] and give grounds to consider signature algorithms with a hidden group as candidates for practical post-quantum cryptoschemes with small signature size. The latter is a motive for the cryptographic community to pay attention to the issue of considering structural attacks on signature algorithms of the type considered.

Conclusion

Within the framework of the methods [13, 16], new post-quantum algebraic signature algorithms with a hidden group has been developed, using 4-dimensional FNAAs, defined over finite fields of characteristic two, as algebraic support. It is shown that there are quite ample opportunities to choose suitable fields $GF(2^z)$ with different degrees of extension. The use of FNAAs, set over the fields $GF(2^z)$, as algebraic support of post-quantum signature algorithms with a hidden group is an essential moment for improving the performance and reducing the hardware implementation cost compared to the case of using FNAAs defined over the ground finite fields $GF(p)$.

An additional increase in performance can be achieved by using 6-dimensional and 8-dimensional FNAAs defined over the fields $GF(2^z)$ with the value of z from 80 to 150 as an algebraic support, including the case of defining FNAAs by sparse BVMTs. However, this is the subject of an independent study, which includes the study of the structure of such FNAAs and developing new forms of the signature verification equations.

References

1. Shor P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on quantum computer. *SIAM Journal of Computing*, 1997, vol. 26, pp. 1484–1509.
2. Ekert A., Jozsa R. Quantum computation and Shor's factoring algorithm. *Reviews of Modern Physics*, 1996, vol. 68, pp. 733–752.
3. Smolin J. A., Smith G., Vargo A. Oversimplifying quantum factoring. *Nature*, 2013, vol. 499, no. 7457, pp. 163–165.
4. Ding J., Schmidt D. Rainbow, a new multivariable polynomial signature scheme. *Conf. on Applied Cryptography and Network Security — ACNS 2005*, Springer Lecture Notes in Computer Science, 2005, vol. 3531, pp. 164–175.
5. Shuaiting Q., Wenbao H., Yifa Li, Luyao J. Construction of extended multivariate public key cryptosystems. *International Journal of Network Security*, 2016, vol. 18, no. 1, pp. 60–67.
6. Jintai D., Dieter S. *Multivariable Public Key Cryptosystems*. 2004. Available at: https://eprint.iacr.org/2004/350.pdf (accessed 09 March 2022).
7. Moldovyan D. N. A practical digital signature scheme based on the hidden logarithm problem. *Computer Science Journal of Moldova*, 2021, vol. 29, no. 2(86), pp. 206–226.
8. Moldovyan D. N. New form of the hidden logarithm problem and its algebraic support. *Bulletin of Academy of Sciences of Moldova. Mathematics*, 2020, no. 2 (93), pp. 3–10.
9. Ducas L., Kiltz E., Lepoint T., Lyubashevsky V., Schwabe P., Seiler G., Stehlé D. *CRYSTALS-Dilithium: A Lattice-Based Digital Signature Scheme*. Available at: https://eprint.iacr.org/2017/633.pdf (accessed 09 March 2022).
10. Alamelou Q., Blazy O., Cauchie S., Gaborit Ph. A code-based group signature scheme. *Designs, Codes and Cryptography*, 2017, vol. 82, no. 1–2, pp. 469–493. doi:10.1007/s10623-016-0276-6
11. Kosolapov Y. V., Turchenko O. Y. On the construction of a semantically secure modification of the McEliece cryptosystem. *Prikl. Diskr. Mat.*, 2019, no. 45, pp. 33–43. doi:10.17223/20710410/45/4
12. Dahmen E., Okeya K., Takagi T., Vuillaume C. Digital signatures out of second-preimage resistant hash functions. *Proc. of the Second Intern. Workshop on Post-Quantum Cryptography, PQCrypto 2008*, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2008, vol. 5299, pp. 109–123. Available at: http://dblp.uni-trier.de/db/conf/pqcrypto/pqcrypto2008.html#DahmenOTV08 (accessed 09 March 2022).
13. Moldovyan N. A., Moldovyan A. A. Candidate for practical post-quantum signature scheme. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2020, vol. 16, iss. 4, pp. 455–461. doi:10.21638/11701/spbu10.2020.410
14. Moldovyan D. N., Moldovyan A. A., Moldovyan N. A. A new concept for designing post-quantum digital signature algorithms on non-commutative algebras. *Voprosy kiberbezopasnosti*, 2022, no. 1(47), pp. 18–25 (In Russian). doi:10.21681/2311-3456-2022-1-18-25
15. Moldovyan D. N., Moldovyan A. A., Moldovyan N. A. A novel method for development of post-quantum digital signature schemes. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 6, pp. 21–29. doi:10.31799/1684-8853-2020-6-21-29
16. Moldovyan A. A., Moldovyan D. N., Moldovyan N. A. A novel method for developing post-quantum digital signature algorithms on non-commutative associative algebras. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 1, pp. 44–53. doi:10.31799/1684-8853-2022-1-44-53
17. Moldovyan A. A., Moldovyan N. A. Signature algorithms on finite non-commutative algebras over fields of characteristic two. *Voprosy kiberbezopasnosti*, 2022, no. 3(49), pp. 58–68 (In Russian). doi:10.21681/2311-3456-2022-3-58-68
18. Ding J., Petzoldt A. Current state of multivariate cryptography. *IEEE Security and Privacy Magazine*, 2017, vol. 15, no. 4, pp. 28–36.
19. Matsumoto T., Imai H. Public quadratic polynomial-tuples for efficient signature verification and message-encryption. *Proc. of Conf. Advances in Cryptology – Eurocrypt'88*, Lecture Notes in Computer Science, Springer Berlin Heidelberg, 1988, vol. 330, pp. 419–453. https://doi.org/10.1007/3-540-45961-8_39
20. Faugére J.-C. A new efficient algorithm for computing Gröbner basis (*F*4). *J. Pure Appl. Algebra*, 1999, vol. 139, no. 1–3, pp. 61–88.
21. Faugére J.-C. A new efficient algorithm for computing Gröbner basis without reduction to zero (*F*5). *Proc. of the Intern. Symp. on Symbolic and Algebraic Computation*, 2002, pp. 75–83. doi:10.1145/780506.780516
22. Moody D., Alagic G., Apon D., Cooper D., Dang Q., Kelsey J., Liu Y., Miller C., Peralta R., Perlner R., Robinson A., Smith-Tone D., and Alperin-Sheriff J. *Status Report on the Second Round of the NIST Post-Quantum Cryptography Standardization Process, NIST Interagency/Internal Report (NISTIR)*. National Institute of Standards and Technology, Gaithersburg, MD, 2020. https://doi.org/10.6028/NIST.IR.8309. Available at: https://csrc.nist.rip/external/nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8309.pdf (accessed 09 March 2022).
23. *Fast-Fourier lattice-based compact signatures over NTRU*. Available at: https://falcon-sign.info/ (accessed 09 March 2022).
24. Ducas L., Kiltz E., Lepoint T., Lyubashevsky V., Schwabe P., Seiler G., Stehlé D. *CRYSTALS-Dilithium: A Lattice-Based Digital Signature Scheme*. Available at: https://eprint.iacr.org/2017/633.pdf https://pq-crystals.org/dilithium/index.shtml (accessed 09 March 2022).

25. *Rainbow Signature. One of three NIST Post-quantum Signature Finalists*. 2021. Available at: https://www.pqcrainbow.org/ (accessed 09 March 2022).
26. Ding J., Schmidt D. Rainbow, a new multivariable polynomial signature scheme. *Proc. of Conf. on Applied Cryptography and Network Security — ACNS 2005*, Springer Lecture Notes in Computer Science, 2005, vol. 3531, pp. 164–175.
27. Moldovyan D. N. A unified method for setting finite non-commutative associative algebras and their properties. *Quasigroups and Related Systems*, 2019, vol. 27, no. 2, pp. 293–308.
28. Moldovyan N. A. Unified method for defining finite associative algebras of arbitrary even dimensions. *Quasigroups and Related Systems*, 2018, vol. 26, no. 2, pp. 263–270.
29. Moldovyan D. N. Post-quantum public key-agreement scheme based on a new form of the hidden logarithm problem. *Computer Science Journal of Moldova*, 2019, vol. 27, no. 1(79), pp. 56–72.

УДК 003.26
doi:10.31799/1684-8853-2023-1-29-40
EDN: KSCBTZ

Постквантовые алгебраические алгоритмы цифровой подписи со скрытой группой

А. А. Молдовян[a], доктор техн. наук, главный научный сотрудник, orcid.org/0000-0001-5480-6016
Н. А. Молдовян[a], доктор техн. наук, главный научный сотрудник, orcid.org/0000-0002-4483-5048, nmold@mail.ru
[a]Санкт-Петербургский Федеральный исследовательский центр РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

Введение: разработка постквантовых стандартов на алгоритмы цифровой подписи является одним из современных вызовов для мирового криптографического сообщества. Недавно предложены два типа алгебраических схем подписи со скрытой группой, в которых конечные некоммутативные ассоциативные алгебры над полем $GF(p)$ используются в качестве алгебраического носителя. Построение алгоритмов этого типа на алгебрах, заданных над конечными полями характеристики два, представляет значительный интерес для повышения производительности и снижения схемотехнической сложности аппаратной реализации. **Цель:** разработать постквантовые алгоритмы цифровой подписи, в которых выполняются вычисления в конечных полях характеристики два. **Результаты:** предложены несколько четырехмерных конечных некоммутативных алгебр, заданных над полем $GF(2^z)$, в качестве алгебраических носителей схем цифровой подписи со скрытой группой. Разработаны рекомендации по выбору значения степени расширения z. В частных случаях значение z является степенью Мерсенна. По сравнению со схемами подписи, основанными на скрытой задаче дискретного логарифмирования, алгебраические алгоритмы подписи со скрытой группой, основанные на вычислительной сложности решения систем многих квадратных уравнений с многими неизвестными, рассматриваются как предпочтительные кандидаты на постквантовые криптосхемы. Предложены новые практичные алгоритмы подписи со скрытой группой. В двух алгоритмах подпись (e, $\mathbf{S}$) представляет собой целое число e и четырехмерный вектор $\mathbf{S}$. Верификация подписи выполняется по векторным уравнениям с тремя и четырьмя вхождениями элемента подписи $\mathbf{S}$. **Практическая значимость:** как и другие известные схемы подписи со скрытой группой, предложенные две схемы имеют достаточно малый размер подписи и открытого ключа. Благодаря сравнительно малой схемотехнической сложности аппаратной реализации и высокой производительности разработанные алгоритмы цифровой подписи представляют практический интерес и привлекательны как потенциальный прототип стандарта на постквантовые алгоритмы цифровой подписи.

Ключевые слова — постквантовые криптосхемы, компьютерная безопасность, электронная цифровая подпись, многомерная криптография, задача дискретного логарифмирования, конечные некоммутативные алгебры, ассоциативные алгебры, циклические группы, многомерная цикличность.

Для цитирования: Moldovyan A. A., Moldovyan N. A. Post-quantum algebraic signature algorithms with a hidden group. *Информационно-управляющие системы*, 2023, no. 1, pp. 29–40. doi:10.31799/1684-8853-2023-1-29-40, EDN: KSCBTZ
For citation: Moldovyan A. A., Moldovyan N. A. Post-quantum algebraic signature algorithms with a hidden group. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 29–40. doi:10.31799/1684-8853-2023-1-29-40, EDN: KSCBTZ

UDC 004.056
doi:10.31799/1684-8853-2023-1-41-50
EDN: MFUKDX

Articles

Information security audit for a manufacturing company

S. V. Shirokova[a], *PhD, Tech., Associate Professor, orcid.org/0000-0001-9384-1877*
O. V. Rostova[a], *PhD, Econ., Associate Professor, orcid.org/0000-0001-6581-3473, O.2908@mail.ru*
M. V. Bolsunovskaya[a], *PhD, Tech., Associate Professor, orcid.org/0000-0001-6650-6491*
L. A. Dmitrieva[a], *Junior Researcher, orcid.org/0000-0003-3831-7137*
T. O. Almataev[b], *PhD, Tech., Associate Professor, orcid.org/0000-0003-2373-9732*
[a]*Peter the Great St. Petersburg Polytechnic University, 29, Politekhnicheskaia St., 195251, Saint-Petersburg, Russian Federation*
[b]*Andijan Machine-Building Institute, 56, Bobur Shoh St., 170019, Andijan, Uzbekistan*

Introduction: *The number of information attacks on company information systems has now increased significantly. The unintended consequences of such attacks are both financial and reputational losses. To increase the effectiveness of information protection, a sound analysis of the level of information system security is necessary.* ***Purpose:*** *To justify the need and describe the information security audit procedure for a manufacturing company.* ***Results:*** *We have analyzed business operations of a certain company and collected the necessary information for an information system security audit. Having analyzed the approaches to threat identification and countermeasure techniques, as well as the specifics of the company in question, we have chosen a combined approach. The study of different risk analysis methods has allowed to substantiate the choice of FRAP methodology. As a result of the audit procedure the compliance of the information system to the information security standards has been assessed.* ***Practical relevance:*** *Recommendations for reducing risks associated with threats to information security have been developed. The implementation of the developed countermeasures to eliminate information security vulnerabilities will allow the company to avoid possible financial losses and avert the damage to the company's reputation.*

Keywords *— digital transformation, information systems security, project, information security standards, business process, audit.*

For citation: Shirokova S. V., Rostova O. V., Bolsunovskaya M. V., Dmitrieva L. A., Almataev T. O. Information security audit for a manufacturing company. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 41–50. doi:10.31799/1684-8853-2023-1-41-50, EDN: MFUKDX

Introduction

The relevance of information protection is due to the widespread use of information systems for the transmission, storage and processing of significant amounts of information, especially given the global trend towards an increase in the number of information attacks that lead to significant financial and reputational losses [1, 2]. For effective protection against attacks, organizations need a sound analysis of the security level of the information system, especially taking into account the ongoing digital transformation [3, 4]. This is exactly what security audits are used for.

The purpose of this study was to substantiate the need and describe the information security audit procedure for a manufacturing company, as well as to develop recommendations based on the analysis. In accordance with this goal, the following tasks were set:

1. Gather all necessary information about the company, the information system used, important business processes and the current information security system.

2. Investigate the types of information security audits and methods used to identify threats, vulnerabilities and countermeasures. Compare them and select the most appropriate ones for the company under investigation.

3. Conduct a security audit of the company's information system using the selected methods.

4. Develop recommendations to eliminate or minimize the identified information security risks of the company.

Drawing up an audit plan is not a trivial task, as it requires the auditor to be experienced in the methods of constructing such plans, and also depends on the objectives of the audit and the features of the audited object. Information security audits can be conducted according to a number of criteria, including requirements defined on the basis of one or more standards, as well as policies and requirements established by stakeholders.

Security audit of the company's information assets allows you to get an idea of the security status of the infrastructure and information security management processes. An information security audit is a test of the ability to successfully counter information security threats. Conducting an independent audit allows you to identify risks in a timely manner and objectively assess the compliance of the parameters characterizing the information security regime with the required level [5, 6].

According to the type of threats, it is possible to distinguish natural, associated with the impact of natural physical processes, and artificial, caused by the impact on the human information environment. Artificial threats can be unintentional (system failures, computer and communication equipment failures, employee errors) and intentional (caused by deliberate actions of people).

All sources of threats to information security are divided into three main groups [7]:

1) threats caused by the actions of subjects. These sources can be predicted and appropriate measures taken;

2) threats caused by technical means (man-made sources);

3) natural sources of threats. Such sources of threats are completely unpredictable, and therefore measures against them should always be applied.

It is the type of threat that determines the nature and features of anti-risk measures.

Rationale for auditing information security

A company engaged in the production of textile materials uses a CRM (Customer Relationship Management) system in its activities. The company uses a CRM-system, namely vTiger CRM, an open source customer relationship management system. The vTiger CRM-system has been specially modified in accordance with the requirements and objectives of the company. It has added modules, reports and functions that are not provided in the standard system solution.

In order for each employee to have access only to the information that he needs to perform his duties, roles are created in the information system. Available modules and available actions with records of these modules are installed for each role.

For the effective operation of the company, it is necessary that the data from the information system be accessible to employees who have access to it, so that this data is reliable, not distorted and confidential. To do this, it is necessary to ensure the security of the company's information system.

Since the CRM-system was recently implemented, it was necessary to conduct a study of possible threats to the security of the information system. The purpose of the information security audit was not only to reveal vulnerabilities and identify risks, but also to describe possible consequences and develop countermeasures to improve the current level of information system security.

Comparative analysis of information security audit types

There are different classifications of audit types. In the work, a study of different types of information security audit depending on the methods and means used was conducted and their distinctive features were identified, which are presented in Table 1 [8–10].

■ ***Table 1.*** Comparison of the types of audit

Audit type	Criteria		
	Verification tools and methods	Result of the test	The ideal with which the result of the test is compared
Hardware audit	Conducting real attacks on the information system by experts using special software and special methods	It is aimed at identifying and fixing the vulnerabilities of the system's software and hardware	Set of known vulnerabilities in the software and the expected test result
Expert audit	Collection and analysis of information on IS, analysis of organizational and administrative documents and information flows of the enterprise	Identifying global errors in the corporate network topology, using security tools, identifying vulnerabilities of information system	Requirements of the company's management for protection, as well as the auditor's own experience
Audit of compliance with standards	Collection and analysis of information on IS and subsequent comparison with the description of the standard	The degree of compliance of the tested IS with the selected standards, as well as recommendations for bringing the security of the IS in accordance with standards	Abstract description of the state of information security which given in the standards
Complex audit	Depends on the set of procedures that will be implemented during this audit		

Regardless of the type of audit used, in general, the information system security audit process consists of the following stages:

– initiation of the audit procedure;
– collecting the audit information;
– data analysis;
– making recommendations and preparing an audit report.

Recommendations are determined by the approach used, the characteristics of the information system being audited, the state of information security in the company, and the level of detail used during the audit.

The use of different types of auditing can be done individually or in combination, it depends on the needs of the company.

The information security standards approach to data analysis begins with selecting the standard that will form the basis, or a set of additional standards. The standards define a core set of information security requirements.

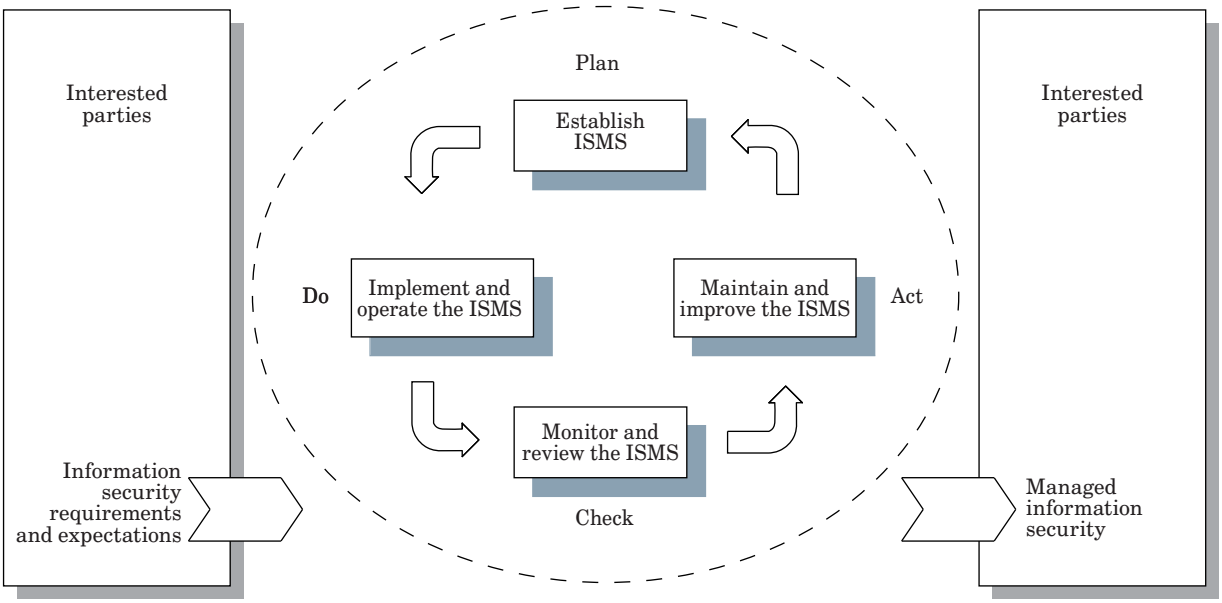
The ISO / IEC 27001 standard assumes the use of a process approach to create, implement, support the information security management system (ISMS). The implementation and use of a set of processes within the company that are interrelated. When creating the processes of the ISMS, the Deming cycle (PDCA) can be used, which is considered in the standard. The Fig. 1 shows the process of creating managed information security, where the process input is submitted to the requirements and expected results in the field of information security of the parties concerned, and on the output we receive managed information security. If desired, to enhance their reputation in the eyes of customers, companies can be certified for compliance with the requirements of ISO / IEC 27001.

For the company in question, a comprehensive information security audit was selected, which includes the following steps [12, 13]:

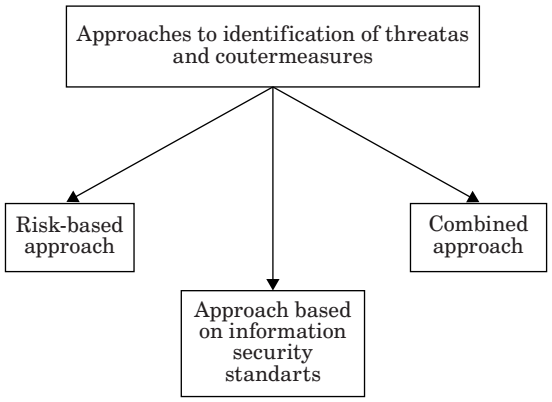
– analysis and construction of a model of interaction of components of information systems used in the provision of business processes of the organization;
– identification and determination of the value of information assets that are most important for ensuring the functioning of business processes. This stage allows you to take into account the unique structure of the system and the point of view of the company's management on critical assets and threats;
– study of the features of the existing information security system, analysis of the settings of standard security tools in the company, server operating systems and communication equipment;
– conducting penetration tests into the company's information environment;
– identification of threats to the most significant information assets, assessment of their level, probability of manifestation and development of recommendations taking into account the requirements and standards of information security formed at previous stages.

The description of the stages is due to the need to specify this procedure for auditing information security in the company in question.

Depending on the selected audit type, different approaches are used to identify threats and countermeasures [14–16]. Variants of existing approaches to identify threats and countermeasures, depend-



■ ***Fig. 1.*** Stages of building and using the ISMS [11]



■ ***Fig. 2.*** Types of approaches to identification of threats and countermeasures

ing on the selected audit approach, are presented in Fig. 2.

Let's describe each approach in more detail.

– The information security standards approach. In this approach, an information security standard is selected, compliance with which will be verified by the information system.

– Risk-based approach. This approach uses risk analysis methods to determine an individual set of information system security requirements [17, 18].

– Combined approach. A basic set of security requirements applicable to an information system is defined by a standard. Additional requirements, are formed on the basis of risk analysis.

For the company in question, a combined approach was chosen.

In addition, it should be noted that information security risk management methods for standard information system implementation projects and pilot projects, when high-tech solutions with a high degree of uncertainty are implemented, will differ significantly [19].

Information security risk analysis methods

Risk analysis is a study of information system security, in order to determine the key IT assets that are important to the company, as well as to identify threats against which they need to be protected. Key assets may include: information resources that support business processes; electronic media: system and application software; hardware; paper media. The importance or value of an asset is determined by the amount of damage that would occur if the confidentiality, availability, or integrity of the asset were to be compromised.

The development of adequate countermeasures to protect these assets occurs during risk management. At the same time, one must consider the fact that the value of the information security system should not exceed the value of the information being protected. The cost of the information security system includes the one-time costs of its development and implementation, as well as the operational costs of maintaining it. Detailed cost items, as well as an assessment of the effects of implementation are presented in detail in the publications [20, 21].

In the course of risk assessment and analysis, the following stages are distinguished:

1) identifying the key assets of the organization;

2) analyzing which key assets need to be protected;

3) creation of a list of expected threats and vulnerabilities that may lead to the emergence of these threats;

4) assessing the likelihood that security threats will materialize;

5) assessing the level of impact with employee involvement;

6) qualitative or quantitative risk assessment;

7) interpretation of the obtained results.

Qualitative or quantitative risk analysis can be done. Most companies usually choose a qualitative risk assessment because it is an easier way to assess risks. In this case, also qualitatively assess the levels of damage from the implementation of the attack, as well as the level of probability of threat of the attack itself.

In order to more reasonably choose a method of information security risk analysis for the enterprise in question, a comparative analysis of the most well-known and popular methods was made: CRAMM, RiskWatch, FRAP, OCTAVE.

This list presents methods that perform qualitative, quantitative, and comprehensive risk assessment. It will also describe one method that is used only in the framework of internal audit. Let's describe each of the methods a little.

1. *CRAMM* is based on an integrated risk analysis approach, i.e. Qualitative and quantitative assessments are used simultaneously [22]. Also, the method is considered universal, because Suitable for companies of different sizes and working in different areas. In addition, this technique allows you to justify the costs of building or upgrading the information security system and prevents unnecessary costs.

CRAMM divides the procedure of risk analysis into three stages.

In the first stage, the research boundaries are first defined, then within the boundaries, the system resources are identified and their value is determined. After this, the critical resources of the information system are determined on the basis of the data obtained. If the level of criticality of system resources is low enough, then this information system is considered to require a basic level of protection, and this level does not require a detailed assessment

of information security threats, therefore, the second stage of the procedure is skipped.

At the second stage, security threats are identified for the system resources under consideration, as well as an analysis of the probability of their occurrence and threat analysis. On the basis of available information, risks of threats are calculated. It should be noted that countermeasures existing in the company are not taken into account in order to avoid an incorrect evaluation of the effectiveness of countermeasures.

At the third stage, a list of adequate countermeasures is developed to reduce risks and develop recommendations for working with them. The auditor also justifies the proposed countermeasures. Further, the company's management analyzes the information received, assesses the labor costs and material costs of implementing or improving the protection system, benefits the business and decides which of the recommended will be implemented.

The merits of this method include:

– availability of a software tool that implements this method;

– well-tested.

The disadvantages of the CRAMM method include:

– sufficient time-consuming procedure;

– requires a high level of qualification of the auditor.

CRAMM is suitable for existing information systems, not for those that are under development.

2. The *RiskWatch* methodology is based on a quantitative risk assessment. Risk is estimated through a numerical value, in this case through the size of annual loss expectancy and an estimation of return on investment [23]. The procedure for risk analysis based on the RiskWatch method consists of four stages.

At the first stage, a description of the research object is given: the type of organization in question, the requirements for the security of the information system, and the structure of the information system.

The second stage provides a detailed description of system resources, losses, and incident classes. In this case, incident classes themselves are obtained after comparing the loss category and resource category.

At the third stage of this procedure, a quantitative risk assessment is carried out. The effect of using the recommended protection is determined using the return on investment indicator.

At the fourth stage, reports with analysis results are generated. When analyzing risks, the RiskWatch method uses special estimates, called LAFE (Local Annual Frequency Estimate) and SAFE (Standard Annual Frequency Estimate).

The merits of this method of risk analysis include:

– in the course of this procedure, in addition to the risk assessment, the effectiveness of the implemented information security measures is calculated;

– a report is generated, on the basis of which a decision can be made about the necessity and expediency of introducing the recommended means of protection;

– there is appropriate software that implements this method.

Disadvantages of RiskWatch is:

– it is difficult to adapt to Russian companies;

– it is quite problematic to obtain estimates of LAFE and SAFE for our conditions.

3. The *FRAP* methodology is based on a qualitative assessment of the risks of implementing information security threats [24]. At the same time, it is based on the principle that an unprotected information system is assessed, which makes it possible to evaluate the effectiveness of new implemented information security tools, i.e. At the initial stage of the analysis, we believe that there is no protection system.

The risk assessment procedure should consist of the following steps:

Identification of the resources to be protected. Information can be obtained by interviewing company employees and by analysing documentation accompanying the information system.

Identification of possible threats. A list of possible threats can be used, in which company representatives mark what they believe to be the most likely threats to the resources in question. In addition, threats can be identified based on statistics for the information system in question or for similar systems.

Risk assessment. The probability of each threat and the damage it can cause is determined. A final risk assessment is then carried out.

Development of countermeasures to eliminate risks or reduce them to a level acceptable by the company's management. The incremental cost of acquiring and implementing the proposed information security tools is identified.

Preparation of analytical reports.

Advantages of the FRAP methodology:

– a detailed description of how to obtain the information required for system and vulnerability analysis;

– the scales used to assess the likelihood of threat occurrence and the level of damage are simplified in this case, as they contain only three criteria.

The disadvantages are that this methodology requires an auditor with a very high level of expertise.

4. The *OCTAVE* method is a method of rapid assessment of critical threats, assets and vulnerabilities [25]. This method of analysis and risk management is used for internal audit by employees of

the company, i.e. It is planned to create a certain group of company employees (technical specialists and management units). This technique qualitatively assesses the risks of information security.

The OCTAVE method consists of eight steps:

At the first step, it is necessary to define qualitative indicators for risk assessment, for example, the level of costs for information security, criticality of information resources.

The second step is the identification of key resources, as well as the compilation of profiles for each of them. The profile implies a description of the characteristics and characteristics of the resource in order to identify security requirements for it.

At the third step, the storage locations for the resources in question are identified, the security of these places is analyzed, bottlenecks are identified.

At the fourth step, critical information security locations are identified to detect the obvious threats as quickly as possible.

At the fifth step, a tree of threat scenarios is compiled.

At the sixth step, a risk is defined for each resource to assess its criticality.

At the seventh step, possible damage from the threat to risk prioritization is calculated.

Countermeasures are being developed at the eighth step.

The advantages of OCTAVE include:

— the methodology is designed to organize different areas;

— ease of use;

— suitable for regular use.

The disadvantages include the following:

— calculated only for internal audit;

— does not involve any residual risk management mechanisms.

Based on the analysis of existing risk analysis methods, FRAP was chosen because it is suitable for external auditing (unlike OCTAVE), and because it is based on a qualitative risk assessment (unlike RiskWatch and CRAMM), which is easier to use, and because in this case a qualitative risk assessment is sufficient.

Audit results

At the first stage, a basic set of requirements was formed based on information security standards, which were divided into the following groups:

1) physical access control and premises monitoring;

2) hardware information system;

3) network support for the information system;

4) system software;

5) application software;

6) organizational support.

The report described the current state of the information security system, optimal structure, identified remarks and threats, measures to reduce threats.

The next stage was the formation of additional requirements, taking into account the specific functioning of the information system on the basis of risk analysis. This approach is quite time consuming and requires a high level of auditor skills. In the course of using this method, an individual set of requirements for the security of the information system is determined.

To determine the key resources of the company, the method of "brainstorming" the company's employees was used. As a result of the discussion, the assets were determined, we will consider some of them:

— servers on which all important information for the company is stored and processed (data about customers, orders, etc.);

— database, where all the information necessary for the functioning of the organization is stored;

— electronic media, on which copies of contracts and agreements with customers and suppliers are stored.

Assets with different levels of criticality were selected for greater visibility of the method.

Identification of possible threats

Possible threats to information security for previously defined resources were identified by brainstorming the company's employees. As a result of the discussion, possible threats to each of the resources were identified. Let's look at them in more detail.

Possible threats to servers:

– malicious server damage or destruction;

– server failure due to technical reasons;

– destruction of the server by a natural disaster, such as fire.

Possible threats to the database where all the information necessary for the organization's operation is stored:

– unintentional modification or deletion of information from the database by company employees;

– leakage of confidential information as a result of copying information from a database using an external electronic medium.

Possible threats to electronic media that store copies of contracts and agreements with customers and suppliers:

– destruction or damage of electronic media by an attacker, resulting in data loss;

– destruction of electronic media due to fire.

At the next stage, a qualitative assessment of the likelihood of threats and an assessment of damage in case of their manifestation was carried out (Table 2).

According to the methodology applied, each level of risk corresponds to a specific response: A — risk

■ ***Table 2.*** Qualitative assessment of the probability of the implementation of threats and assessment of damage

Threats	Probability	Damage	Risk
Malicious server damage or destruction	Low	High	B
Server failure due to technical reasons	High	High	A
Destruction of the server by natural cataclysm, for example, by fire	Average	High	B
Unintentional modification or deletion of information from the database by employees of the company	High	Average	B
The leakage of confidential information as a result of copying information from the database using an external electronic medium	Average	Average	B
Destroying or damaging an electronic medium by an attacker, which results in data loss	Low	Low	D
Destruction of the electronic carrier in consequence of a fire	Average	Low	C

action should be taken immediately on a mandatory basis; B — risk action should be taken; C — monitoring of the situation is required; D — no action is required at this time.

Recommendations

After the audit of the company's information security, a list of recommendations was compiled.

On the basis of the table obtained, one can see that only one of the risks under consideration is a critical level: server failure due to technical reasons.

To reduce this risk it is recommended:

— use of RAID technology (fault tolerance technology);

— maintaining a constant temperature, through the use of air conditioning (main, backup);

— reservation of important information to separate servers or external storage media;

— conduct a continuous check of the server room to detect breakdowns or malfunctions.

To reduce risks in case of inadvertent modification or deletion of information from the database, company employees should implement the following:

— reservation of important information to separate servers or external storage media;

— implementation of the functionality of storing the history of database changes.

The risk associated with the leakage of confidential information as a result of copying information from the database using an external electronic medium is high enough to reduce it:

— implementation of the DLP (Data Leak Prevention — Information Leakage Prevention System), which will analyze the data streams along the protected perimeter;

— introduction of organizational and administrative documents, which will describe: what information is confidential, the requirements for working with such information, as well as the appropriate sanctions for their failure to comply;

— implementation of active monitoring systems for employee workstations.

To reduce the risk associated with the destruction of electronic media that stores copies of contracts and agreements with customers and suppliers, you can use the same information security tools as to protect servers from fire, but this is not mandatory, because the damage from the threat is estimated as low and most likely the cost of implementing protection will be higher than the damage caused.

The risk of destruction or damage to electronic media by an attacker, resulting in data loss, is low. The damage is also assessed as low, so there is no need to take any measures to manage this risk at this time.

At the organizational level, it is recommended:

— regularly review the security policy when making changes to the company's system;

— include in the job responsibilities of all employees the task of ensuring information security;

— the IT service checks the personnel being hired;

— remind an employee of the company that they have signed an agreement to comply with the trade secret regime;

— regularly send information about incidents and viruses to all employees of the company;

— when creating a company's information security system, you need to consider the option that the company's employees have sufficient knowledge to perform unauthorized access to information when they have the opportunity;

— place the equipment in rooms that cannot be accessed by persons who are not connected with the maintenance of this room;

— in the event of a hard disk failure, call a representative of the supplier and inspect the damage in the presence of responsible representatives of the company;

— update antivirus databases every day;
— conduct regular third-party security audits of the company, as well as penetration tests performed by independent experts
— regularly use the vulnerability scanner to track changes in the security of the company's information system;
— ensure that the company's information system meets any published security standard;
— use only certified local and network security tools in the company;
— exclude the presence of service personnel in the company's office without the presence of controlling persons;
— develop a list of documents and recommendations on non-disclosure of confidential information by employees.

Conclusions

The study substantiated the feasibility of implementing an information security system at a manufacturing enterprise, and described in detail the procedure for auditing information security in a particular organization. In accordance with the goal, the following tasks were solved.

Collected and analyzed information about the company. Important business processes of the company include the processes of production of products (fabrics). Special attention is paid to the technological process of production. For effective operation of these business processes, up-to-date information from the information system is necessary, access to which should not be difficult for authorized persons. Also, the information must be securely protected from access by unauthorized persons. To ensure these requirements, it is necessary to conduct a security audit of the information system.

Possible types of information system security audits are analyzed. A review of methods for identifying threats and countermeasures is made: an approach based on risk analysis, an approach based on information security standards, and a combined approach. We selected a comprehensive audit with a combined approach to identifying threats and countermeasures. The ISO/IEC 27001 standard was chosen as the information security standard, which describes the requirements for information security. FRAP was chosen as the risk analysis method.

The company's information system security audit was conducted taking into account the selected procedure. Existing threats to information security were identified and analyzed. The most important of them are: server failure due to technical reasons, as well as copying information to external removable media for further disclosure.

Developed recommendations to eliminate existing threats to information security for the company. The most important implementation recommendations are the purchase and installation of an air conditioning and fire alarm system in the server room, as well as the implementation of a DLP-system designed to prevent leaks of confidential information outside the corporate network. In addition, the cost of all recommended information security measures for threats found was 183,000 rubles, which is included in the company's acceptable information security budget of 210,000 rubles. If these recommendations are implemented (organizational and software-technical), the security of the information system will increase, and the risks of threats found will decrease, which will allow the company to avoid financial and reputational losses in the future.

Financial support

The research is funded by the Ministry of Science and Higher Education of the Russian Federation under the strategic academic leadership program "Priority 2030" (Agreement 075-15-2021-1333 dated 30.09.2021).

References

1. Whitman M. E. In defense of the realm: understanding the threats to information security. *International Journal of Information Management*, 2004, vol. 24, no. 1, pp. 43–57. doi:10.1016/j.ijinfomgt.2003.12.003
2. Anisimov V. G., Anisimov E. G., Saurenko T. N., Zotova E. A. Models of forecasting destructive influence risks for information processes in management systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 5, pp. 18–23. doi:10.31799/1684-8853-2019-5-18-23
3. Babash A. V., Baranova E. K. *Aktual'nye voprosy zashchity informacii* [Actual issues of information protection]. Moscow, INFRA-M Publ., 2017. 111 p. (In Russian). doi:10.12737/monography_58dbc380aa3a4
4. Amirova E. F., Voronkova O. Yu., Zakirova N. R., Stepanenko O. G., Doguchaeva S. M., Murzagalina G. M. Internet of things as a tool for development of russias digital economy. *International Journal of Mechanical Engineering and Technology*, 2019, vol. 10, no. 2, pp. 1011–1019.
5. Anisimov V. G., Zegzhda P. D., Suprun A. F., Anisimov E. G. The problem of innovative development of information security systems in the transport sector. *Automatic Control and Computer Sciences*, 2018, vol. 52, no. 8, pp. 1105–1110. doi:10.3103/S0146411618080035

6. Shmeleva A. S., Suloeva S. B., Rostova O. V. Use of agile management tools in projects of information security systems implementation. *Problems of Information Security. Computer Systems*, 2021, no. 4, pp. 123–136. doi:10.48612/jisc/5zkx-22b9-8 kam
7. Vostretsova E. V. *Osnovy informacionnoj bezopasnosti* [Fundamentals of information security]. Yekaterinburg, Ural'skij universitet Publ., 2019. 204 p. (In Russian).
8. Astahov A. *Audit of Information Systems Security*. Available at: http://iso27000.ru/chitalnyi-zai/audit-informacionnoibezopasnosti/audit-bezopasnosti-informacionnyh-sistem (accessed 25 August 2022).
9. *Audit informacionnoj bezopasnosti* [Information security audit]. Available at: https://intuit.ru/studies/courses/600/456/lecture/10226 (accessed 15 September 2022).
10. Maksimova E. A. Audit of information security. *Information Protection. Inside*, 2006, no. 6(12), pp. 64–65 (In Russian).
11. State Standard R ISO/IEC 27001-2006. *Information technology. Security techniques. Information security management systems. Requirements*. Moscow, Standartov Publ., 2008. 31 p. (In Russian).
12. *Kompleksnyj audit* [Comprehensive audit]. Available at: http://www.pointlane.ru/security_a/complex (accessed 24 September 2022).
13. Sabillon R., Serra-Ruiz J., Cavaller V., Cano J. A comprehensive cybersecurity audit model to improve cybersecurity assurance: The CyberSecurity Audit Model (CSAM). *2017 Intern. Conf. on Information Systems and Computer Science (INCISCOS)*, 2017, pp. 253–259. doi:10.1109/INCISCOS.2017.20
14. Pandey S. K. A comparative study of risk assessment methodologies for information systems. *Bulletin of Electrical Engineering and Informatics*, 2012, vol. 1, no. 2, pp. 111–122.
15. Hashim N. A., Abidin Z. Z., Puvanasvaran A. P., Zakaria N. A., Ahmad R. Risk assessment method for insider threats in cyber security: A review. *International Journal of Advanced Computer Science and Applications*, 2018, vol. 9, no. 11. doi:10.14569/IJACSA.2018.091119
16. Shirokova S., Kislova E., Rostova O., Shmeleva A., Tolstrup L. Company efficiency improvement using agile methodologies for managing IT projects. *ACM Intern. Conf. Proc. Series (DTMIS 2020)*, 2020. doi:10.1145/3446434.3446465
17. Kuzminykh I., Ghita B., Sokolov V., Bakhshi T. Information security risk assessment. *Encyclopedia*, 2021, vol. 1, no. 3, pp. 602–617. doi:10.3390/encyclopedia1030050
18. Baranova E. K., Chernova M. V. Comparative analysis of programming tools for cybersecurity risk assessment. *Problems of Information Security. Computer Systems*, 2014, no. 4, pp. 160–168.
19. Zharova M., Shirokova S., Rostova O. Management of pilot IT projects in the preparation of energy resources. *E3S Web of Conf.*, 2019, vol. 110, 02033. doi:10.1051/e3sconf /201911002 033
20. Anisiforov A. B. *Metodiki ocenki effektivnosti informacionno-tekhnologicheskih proektov v biznese* [Methods of evaluating the effectiveness of information technology projects in business]. Saint-Petersburg, Politekhnicheskij universitet Publ., 2018. 127 p. (In Russian).
21. Grozdova A., Shirokova S., Rostova O., Shirokova A., Shmeleva A. Rationale for information and technological support for the enterprise investment management. *Lecture Notes in Networks and Systems*, 2022, vol. 387, pp. 181–190. doi:10.1007/978-3-030-93872-7_15
22. Yang T., Berger E. D., Kaplan S. F., Moss E. B. CRAMM: Virtual memory support for garbage-collected applications. *Proc. of the 7th USENIX Symp. on Operating Systems Design and Implementation — OSDI'06*, 2006, pp. 103–116.
23. Kouns J., Minoli D. *Information Technology Risk Management in Enterprise Environments: A Review of Industry Practices and a Practical Guide to Risk Management Teams*. 2010. doi:10.1002/9780470558133
24. Nurul A. H., Zaheera Z. A., Puvanasvaran A. P., Zakaria N. A., Ahmad R. Risk assessment method for insider threats in cyber security: A review. *Int. J. Adv. Comput. Sci*, 2018, vol. 9, no. 11, pp. 126–130. doi:10.14569/IJACSA.2018.091119
25. Jufri M. T., Hendayun M., Suharto T. Risk-assessment based academic information system security policy using octave Allegro and ISO 27002. *Proc. of the 2nd Intern. Conf. on Informatics and Computing, ICIC 2017*, 2018. doi:10.1109/IAC.2017.8280541

УДК 004.056
doi:10.31799/1684-8853-2023-1-41-50
EDN: MFUKDX

Аудит информационной безопасности производственной компании

С. В. Широкова[а], канд. техн. наук, доцент, orcid.org/0000-0001-9384-1877
О. В. Ростова[а], канд. экон. наук, доцент, orcid.org/0000-0001-6581-3473, O.2908@mail.ru
М. В. Болсуновская[а], канд. техн. наук, доцент, orcid.org/0000-0001-6650-6491
Л. А. Дмитриева[а], младший научный сотрудник, orcid.org/0000-0003-3831-7137
Т. О. Алматаев[б], канд. техн. наук, доцент, orcid.org/0000-0003-2373-9732
[а]Санкт-Петербургский политехнический университет Петра Великого, Политехническая ул., 29, Санкт-Петербург, 195251, РФ
[б]Андижанский машиностроительный институт, Бобур шох ул., 56, Андижан, 170019, Узбекистан

Введение: количество информационных атак на информационные системы компаний в настоящее время значительно увеличилось. Нежелательные последствия таких воздействий заключаются как в финансовых, так и в репутационных потерях. Для повышения эффективности защиты информации необходим обоснованный анализ уровня безопасности информационной системы. **Цель:** обосновать необходимость и описать процедуру аудита информационной безопасности для производственной компании. **Результаты:** проанализирована деятельность компании, собрана необходимая информация для проведения аудита безопасности информационной системы. На основе анализа подходов к выявлению угроз и контрмер, а также специфики рассматриваемой компании был выбран комбинированный подход. Исследование различных методов анализа рисков позволило обосновать выбор методологии FRAP. В результате проведения процедуры аудита даны оценки соответствия информационной системы стандартам информационной безопасности. **Практическая значимость:** разработаны рекомендации по снижению рисков, связанных с угрозами информационной безопасности. Внедрение разработанных контрмер по устранению уязвимостей информационной безопасности позволит компании предотвратить возможные финансовые потери и ущерб репутации компании.

Ключевые слова — цифровая трансформация, безопасность информационных систем, проект, стандарты информационной безопасности, бизнес-процесс, аудит.

Для цитирования: Shirokova S. V., Rostova O. V., Bolsunovskaya M. V., Dmitrieva L. A., Almataev T. O. Information security audit for a manufacturing company. *Информационно-управляющие системы*, 2023, № 1, с. 41–50. doi:10.31799/1684-8853-2023-1-41-50, EDN: MFUKDX
For citation: Shirokova S. V., Rostova O. V., Bolsunovskaya M. V., Dmitrieva L. A., Almataev T. O. Information security audit for a manufacturing company. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 41–50. doi:10.31799/1684-8853-2023-1-41-50, EDN: MFUKDX

УДК 621.391
doi:10.31799/1684-8853-2023-1-51-60
EDN: UBBHKD

Научные статьи
Articles

Оценка среднего возраста информации в системах со случайным доступом и множественным выходом

А. В. Борисовская[a]***,*** *ассистент, orcid.org/0000-0002-0561-4226*
А. М. Тюрликов[a]***,*** *доктор техн. наук, профессор, orcid.org/0000-0001-7132-094X, turlikov@k36.org*
[a]*Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ*

Введение: *важным направлением в исследовании интернета вещей является анализ систем случайного множественного доступа, которые могут обеспечить устойчивую работу при большом числе устройств, в частности исследование среднего возраста информации для таких систем.* ***Цель:*** *исследовать средний возраст информации в системе со случайным доступом и множественным выходом.* ***Результаты:*** *сценарий интернета вещей, в котором все устройства находятся на одинаковом расстоянии от базовой станции, описан моделью со случайным доступом и множественным выходом, предложенной С. Г. Фоссом в 2017 г. Отмечено, что эта модель в отличие от других систем случайного множественного доступа обеспечивает стабильную работу системы при потенциально неограниченном числе устройств. В настоящей работе впервые применительно к этой модели предложен способ определения среднего возраста информации. Данный способ позволяет по последовательности моментов поступления сообщений в систему и выхода сообщений из системы определить средний возраст информации. Исследована зависимость среднего возраста информации от интенсивности входного потока. Выявлено, что средний возраст информации в системе с множественным выходом конечен при любой интенсивности входного потока, отличной от нуля.* ***Практическая значимость:*** *предложенный способ определения среднего возраста информации для системы с множественным выходом позволяет сравнивать по этому показателю различные сценарии систем интернета вещей и определять целесообразность использования систем с множественным выходом с учетом специфики рассматриваемого сценария.* ***Обсуждение:*** *в данной работе рассматривалась упрощенная модель системы со случайным доступом и множественным выходом. Однако результаты, полученные для этой модели, можно обобщить для более сложных моделей.*

Ключевые слова *— интернет вещей, средний возраст информации, случайный множественный доступ, множественный выход.*

Для цитирования: Борисовская А. В., Тюрликов А. М. Оценка среднего возраста информации в системах со случайным доступом и множественным выходом. *Информационно-управляющие системы*, 2023, № 1, с. 51–60. doi:10.31799/1684-8853-2023-1-51-60, EDN: UBBHKD

For citation: Borisovskaya A. V., Turlikov A. M. Estimation of the average age of information in random access systems with multiple departure. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 51–60 (In Russian). doi:10.31799/1684-8853-2023-1-51-60, EDN: UBBHKD

Введение

За последние несколько лет интернет вещей (Internet of Things — IoT) успел занять важное место в нашей жизни. IoT [1] позволяет создавать гигантские сенсорные сети, которые могут передавать данные между собой или на общий центр без участия человека. Системы IoT помогают нам следить за окружающей средой, вовремя реагировать на чрезвычайные ситуации, удаленно контролировать состояние здоровья. В системах IoT постоянно возрастает число устройств, что привело к появлению двух взаимосвязанных проблем: 1) хорошо исследованные характеристики таких систем (задержка и т. п.) не позволяют в полной мере учесть особенности данных, которые передаются в таких системах; 2) с увеличением числа устройств система может перестать работать стабильно.

Опишем более детально первую проблему. В теории систем массового обслуживания (СМО) одним из основных показателей качества работы систем является средняя задержка. Она исследуется давно, в том числе и в системах интернета вещей. В системах мониторинга окружающей среды или лесных пожаров необходимо, чтобы сообщения об изменении показателей температуры, влажности или степени загрязнения воздуха приходили на общий центр своевременно. При этом из-за ограниченных ресурсов канала в таких системах сообщения будут приходить с задержкой. Уменьшить среднюю задержку в системе можно за счет уменьшения интенсивности появления новых сообщений, т. е. за счет увеличения периода анализа окружающей обстановки. Однако при таком подходе изменение важных показателей системы мониторинга может быть получено несвоевременно, так как сообщения будут отправляться реже. Таким образом, оптимизация средней задержки в системах IoT не приводит к оптимизации времени получения

свежих данных о состоянии системы. Именно этот факт и объясняет сформулированную ранее первую проблему. Для ее решения возникла необходимость введения нового показателя качества функционирования систем, отражающего актуальность полученной информации. Такой показатель был предложен в работе [2] и получил название возраст информации. В англоязычной литературе для него используется термин Age of Information (AoI). Он может играть важную роль не только в системах IoT [3], но и в вопросах надежности различных систем передачи данных [4].

Один из путей решения второй проблемы систем IoT — это использование систем случайного доступа с множественным выходом [5, 6], которые могут обеспечить устойчивую работу при любой интенсивности входного потока. Как было отмечено выше при обсуждении первой проблемы, в последнее время средний возраст информации интенсивно исследовался для различных моделей систем IoT. Однако для систем, которые могут обеспечить устойчивую работу при любой интенсивности входного потока, такие исследования не проводились. Поэтому исследование среднего возраста информации для системы с множественным выходом является актуальным.

Обзор исследований среднего возраста информации в системах IoT

Впервые понятие среднего возраста информации для систем IoT было сформулировано в работе [2]. В этой работе, а затем в работах [7–9] был описан способ вычисления среднего возраста информации для простейших СМО, таких как M|D|1, M|M|1 и D|M|1. В работе [10] исследуется средний возраст информации для более сложных СМО с прямым и обратным порядком обслуживания. В англоязычной литературе существуют сокращенные названия для прямого (First-Come-First-Served — FCFS) и обратного (Last-Come-First-Served — LCFS) порядка обслуживания. Возраст информации для систем с несколькими источниками рассматривался в работе [11].

В большинстве систем IoT для сценариев с большим числом устройств используется случайный множественный доступ (СМД) [12]. Например, в сетях, построенных по технологии LoRaWAN, используется алгоритм случайного множественного доступа ALOHA [12, 13]. В настоящее время появляются исследования среднего возраста информации в системах СМД [14–21], учитывающие специфические особенности таких систем.

Для большинства систем передачи данных, в том числе и для систем СМД, средняя задержка является монотонно возрастающей функцией от интенсивности входного потока. Существует небольшой класс систем, в которых средняя задержка немонотонна. Например, в системе СМД с обратной двоичной связью «успех-неуспех» [22] средняя задержка имеет высокие значения при низких интенсивностях входного потока, затем с ростом интенсивности она убывает и снова возрастает, когда система перестает быть стабильной. В отличие от средней задержки, средний возраст информации не является монотонно возрастающей функцией от интенсивности входного потока. Во многих системах, например в СМО с порядком обслуживания FCFS и в системах СМД с алгоритмом ALOHA, средний возраст информации убывает с ростом интенсивности входного потока на интервале малых значений и неограниченно возрастает, когда система перестает быть стабильной. В некоторых системах, например в системе M|M|1 с порядком обслуживания LCFS, средний возраст информации с ростом интенсивности входного потока монотонно убывает. Он продолжает уменьшаться даже в тот момент, когда система перестает быть стабильной. При этом, как и в случае со средней задержкой, при заданной интенсивности входного потока желательно, чтобы средний возраст информации был как можно меньше [14].

Понятие среднего возраста информации в СМО и системах СМД

Рассмотрим понятие среднего возраста информации согласно [2]. Возраст информации $\Delta(t)$ — функция от времени, значение которой увеличивается линейно и уменьшается в момент завершения обслуживания очередной заявки. Будем считать, что в начале работы системы $\Delta(0) = 0$. Обозначим через T_{in}^{i} время появления в системе заявки от устройства i, а через T_{out}^{i} — время завершения обслуживания заявки от устройства i. Тогда в момент времени T_{out}^{i} возраст информации $\Delta\left(T_{out}^{i}\right)$ уменьшится на величину, равную $T_{in}^{i} - T_{in}^{i-1}$.

Обозначим через T время наблюдения за работой системы. Тогда средний возраст информации $\bar{\Delta}_T$ на интервале наблюдения $[0, T]$ определяется как отношение площади под графиком функции $\Delta(t)$ к интервалу времени работы системы T [2]:

$$\bar{\Delta}_T = \frac{1}{T}\int_0^T \Delta(t)\mathrm{d}t.$$

Средний возраст информации $\bar{\Delta}_T$ на полубесконечном интервале наблюдения $[0, +\infty)$ определяется следующим образом:

$$\bar{\Delta} = \lim_{T\to\infty} \bar{\Delta}_T. \quad (1)$$

Рассмотрим СМО с порядком обслуживания FCFS и случайным временем обслуживания. Пример работы такой системы представлен на рис. 1. Предположим, что за работой системы наблюдают в течение периода времени T. Будем считать, что в начальный момент времени в системе нет ни одной заявки. Пусть за период наблюдения в системе были обслужены три заявки. На рис. 1 показано, как будет меняться возраст информации в такой системе на интервале от 0 до T. Функция $\Delta(t)$ будет иметь пилообразный вид.

В данном примере средний возраст информации на интервале от 0 до T — отношение площади заштрихованной области к длине интервала наблюдения T. Покажем, как можно вычислить средний возраст информации, используя данные о моментах поступления заявок в систему и моментах окончания обслуживания. Обозначим через s_1 площадь трапеции $A_0A_1B_1T_{in}^1$, через s_2 — площадь трапеции $T_{in}^1A_2B_2T_{in}^2$, через s_3 — площадь трапеции $T_{in}^2A_3B_3T_{in}^3$, а через s_4 — площадь треугольника $T_{in}^3B_3T_{out}^3$. Тогда средний возраст информации на интервале от 0 до T в данном примере можно вычислить следующим образом:

$$\bar{\Delta}_T = \frac{1}{T}\left(s_1 + s_2 + s_3 + s_4\right). \quad (2)$$

Найдем площади каждой из этих фигур отдельно. Площадь трапеции $A_0A_1B_1T_{in}^1$ можно вычислить как разность площадей двух треугольников $A_0A_1T_{out}^3$ и $T_{in}^1B_1T_{out}^1$. Данные треугольники являются прямоугольными и равнобедренными. Следовательно, площадь трапеции $A_0A_1B_1T_{in}^1$ будет равна

$$s_1 = \frac{1}{2}\left(T_{out}^1\right)^2 - \frac{1}{2}\left(T_{out}^1 - T_{in}^1\right)^2. \quad (3)$$

Площади трапеций $T_{in}^1A_2B_2T_{in}^2$ и $T_{in}^2A_3B_3T_{in}^3$ вычисляются аналогично:

$$s_i = \frac{1}{2}\left(T_{out}^i - T_{in}^{i-1}\right)^2 - \frac{1}{2}\left(T_{out}^i - T_{in}^i\right)^2, \quad (4)$$

где $i \in \{2, 3\}$.

Площадь треугольника $T_{in}^2B_3T_{out}^3$ можно вычислить по формуле

$$s_4 = \frac{1}{2}\left(T_{out}^3 - T_{in}^3\right)^2. \quad (5)$$

Подставив (3), (4) и (5) в выражение (2), получим

$$\bar{\Delta}_T = \frac{1}{2T}\left(\left(T_{out}^1\right)^2 - \left(T_{out}^1 - T_{in}^1\right)^2 + \right.$$
$$\left. + \left(T_{out}^2 - T_{in}^1\right)^2 - \left(T_{out}^2 - T_{in}^2\right)^2 + \left(T_{out}^3 - T_{in}^2\right)^2\right). \quad (6)$$

По аналогии с рассмотренным выше примером можно получить выражение для среднего возраста информации в общем виде. Обозначим через n число заявок в системе, которые были обслужены за время T от начала работы системы. Тогда средний возраст информации в этой системе можно вычислить по следующей формуле:

$$\bar{\Delta}_T = \frac{1}{T}\left(s_1 + \sum_{i=2}^{n} s_i + \tilde{s}\right). \quad (7)$$

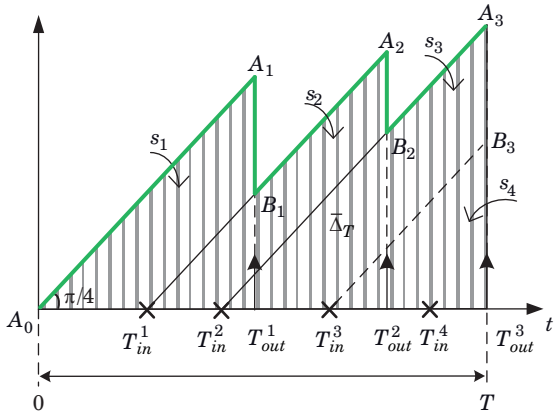
Здесь $\tilde{s}$ — площадь треугольника, вычисляемая как

$$\tilde{s} = \frac{1}{2}\left(T_{out}^n - T_{in}^n\right)^2. \quad (8)$$

Подставив (3), (4) и (8) в выражение (7), получим

$$\bar{\Delta}_T = \frac{1}{2T}\left(\left(T_{out}^1\right)^2 - \left(T_{out}^1 - T_{in}^1\right)^2 + \right.$$
$$\left. + \sum_{i=2}^{n}\left(\left(T_{out}^i - T_{in}^{i-1}\right)^2 - \left(T_{out}^i - T_{in}^i\right)^2\right) + \left(T_{out}^n - T_{in}^n\right)^2\right). \quad (9)$$

Используя выражение (9), можно по результатам имитационного эксперимента вычислить значение $\bar{\Delta}_T$. В работе [2] показано, как на осно-



■ ***Рис. 1.*** Типовой вид зависимости возраста информации от времени

■ ***Fig. 1.*** Typical view of the dependence of the age of information on time

ве (9) для системы M|M|1 можно получить явное выражение зависимости среднего возраста информации от интенсивности входного потока λ и интенсивности обслуживания μ:

$$\bar{\Delta} = \frac{1}{\mu}\left(1 + \frac{1}{\rho} + \frac{\rho^2}{1-\rho}\right),$$

где $\rho = \lambda/\mu$. В работах [7–9] рассматриваются способы вычисления среднего возраста информации и для других СМО.

Рассмотрим систему M|D|1, которая является одной из простейших СМО с порядком обслуживания FCFS и постоянным временем обслуживания. Предположим, что время работы системы разделено на интервалы одинаковой длины — окна. Длительность одного окна равна времени обслуживания одной заявки. Для упрощения анализа системы будем считать, что это время равняется единице. Будем предполагать, что обслуживание заявки может начаться только в начале окна. Назовем эту систему синхронной системой M|D|1. Наибольший интерес представляет синхронная система M|D|1, так как в дальнейшем мы будем сравнивать с ней более сложные синхронные системы.

Средний возраст информации на интервале от 0 до T для этой системы можно оценить с помощью имитационного моделирования, используя выражение (9). Моделирование синхронной системы M|D|1 можно организовать следующим образом.

1. Задать параметры системы: время работы системы T, измеряемое в окнах; интенсивность появления заявок в системе λ.

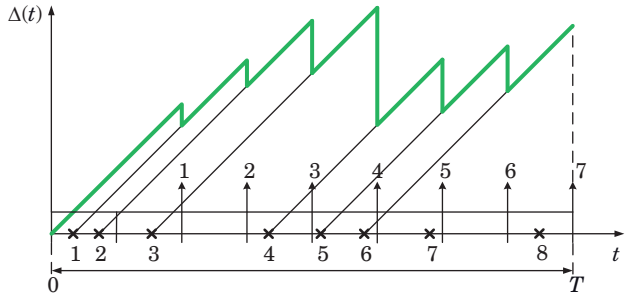
2. Сгенерировать для каждой заявки время ее появления в системе в соответствии с пуассоновским потоком интенсивности λ. Вычислить количество заявок, появившихся в системе за T окон.

3. Найти для каждой заявки время ее обслуживания в соответствии с порядком обслуживания FCFS. Вычислить количество заявок, обслуженных в системе за T окон.

4. Оценить средний возраст информации на интервале от 0 до T, используя выражение (9).

Пример изменения возраста информации в синхронной системе M|D|1 на интервале от 0 до T приведен на рис. 2. В данной системе возраст информации может уменьшиться только в конце окна, так как система является синхронной.

В отличие от системы M|D|1, в системе с алгоритмом ALOHA могут возникать конфликты. Конфликты нарушают порядок обслуживания и делают его случайным процессом. При вычислении среднего возраста информации в таких системах принято учитывать только те сообщения, которые покинули систему не позже появившихся после них сообщений [14]. Другими словами, если сообщение i покинуло систему позже сообщения j и $j > i$, то информация в нем считается неактуальной и средний возраст информации продолжает расти. Пример изменения возраста информации в системе с алгоритмом ALOHA на интервале от 0 до T представлен на рис. 3.



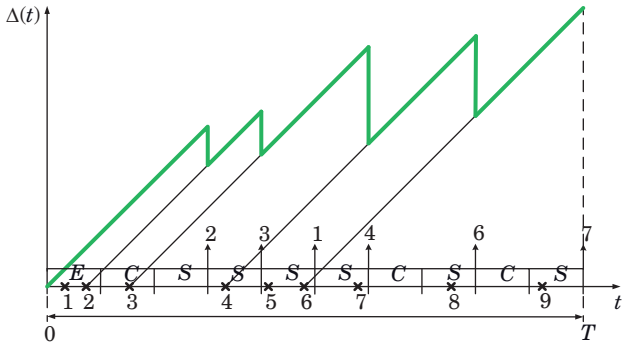
■ ***Рис. 2.*** Пример изменения возраста информации в синхронной системе M|D|1

■ ***Fig. 2.*** An example of changing the age of information in the synchronous system M|D|1

Для алгоритма ALOHA существенно сложнее, чем для СМО из работ [2, 7–9], получить явную зависимость среднего возраста информации от интенсивности входного потока. В работе [14] получены только оценки среднего возраста информации для алгоритма ALOHA.

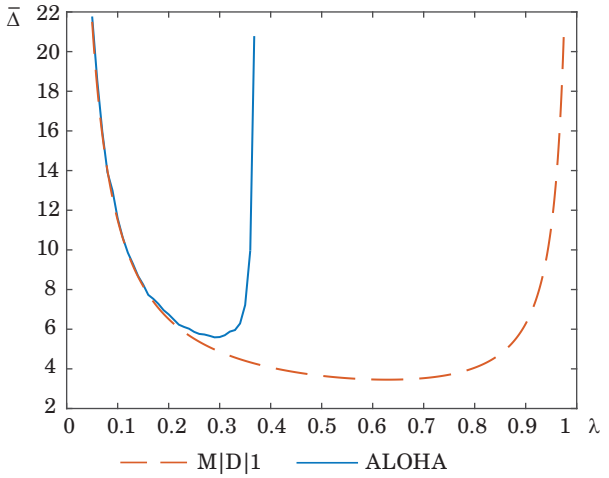
При большом числе окон имитационное моделирование позволяет получить оценку среднего возраста информации (1) с любой заданной точностью. На рис. 4 показаны зависимости среднего возраста информации от интенсивности входного потока для синхронной системы M|D|1 и системы с алгоритмом ALOHA. Данные зависимости получены с помощью имитационного моделирования при T = 100 000.

Вопросы стабильности СМО и систем СМД с неограниченным числом устройств исследуются давно [23, 24]. Под средней задержкой $\bar{D}(\lambda)$ в таких системах понимается среднее время нахождения сообщения в системе при интенсивно-



■ ***Рис. 3.*** Пример изменения возраста информации в системе с алгоритмом ALOHA

■ ***Fig. 3.*** An example of changing the age of information in the ALOHA algorithm



■ ***Рис. 4.*** Средний возраст информации в системе M|D|1 и в системе с алгоритмом ALOHA

■ ***Fig. 4.*** Average age of information in the M|D|1 system and in the ALOHA algorithm

сти входного потока λ. Для каждой из этих систем известна критическая интенсивность входного потока λ_c, которая определяется следующим образом: $\lambda_c \triangleq \sup\{\lambda : \bar{D}(\lambda) < \infty\}$. Другими словами, λ_c — это интенсивность, до которой система работает устойчиво. Для системы M|D|1 $\lambda_c = 1$, а для системы с алгоритмом ALOHA $\lambda_c = e^{-1}$. Средний возраст информации в обеих системах конечен на интервале $(0, \lambda_c)$ и неограниченно возрастает на границах этого интервала. Для наглядности на рис. 4 зависимость $\bar{\Delta}(\lambda)$ для системы M|D|1 приведена при λ в диапазоне от 0,05 до 0,975 < 1, а для системы с алгоритмом ALOHA – в диапазоне от 0,05 до 0,3678 < e^{-1}.

Модель системы

Модель системы СМД с множественным выходом была впервые предложена в работе [5]. Доказано, что она работает стабильно при любой интенсивности входного потока [5]. В этой модели в качестве зоны действия базовой станции рассматривается поверхность сферы, что позволяет избежать краевых эффектов. Однако моделирование и анализ такой модели является сложной задачей. В работе [6] предложена упрощенная модель системы с множественным выходом. В качестве территории, на которой находятся устройства, в этой модели рассматривается окружность. При таком подходе также отсутствуют краевые эффекты, но становится проще моделировать и анализировать модель. Основные свойства системы при этом сохраняются. Поэтому в текущей работе будем рассматривать модель системы с множественным выходом из работы [6].

Опишем модель системы из работы [6] в виде набора допущений.

1. В системе имеется базовая станция. Зоной действия базовой станции считается окружность. Примем длину окружности за единицу.

2. Процесс появления устройств в системе описывается пространственным точечным пуассоновским процессом интенсивности λ. В момент появления устройство содержит единственное сообщение. После успешной передачи устройство покидает систему. В данной модели устройство и сообщение являются тождественными понятиями.

3. Время в системе разделено на окна одинакового размера. Считается, что устройства знают моменты разделения окон. Предполагается, что передача сообщения занимает одно окно и может осуществляться только в начале окна.

4. В окне может произойти одна из трех ситуаций: «успех» (передает одно устройство), «пусто» (ни одно устройство не передает) или «конфликт» (передают два или более устройств).

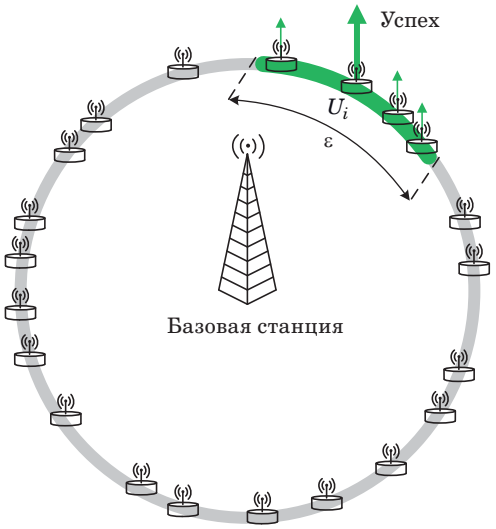
5. Предполагается, что в системе имеется обратная связь. В конце каждого окна базовая станция передает с помощью обратной связи всем устройствам информацию о событии в окне.

6. Все устройства, которые находятся в системе, являются активными. В начале каждого окна активные устройства с вероятностью $p_t = 1/N_t$ принимают решение о передаче сообщения. N_t — число устройств в системе в начале окна t, которое известно всем устройствам.

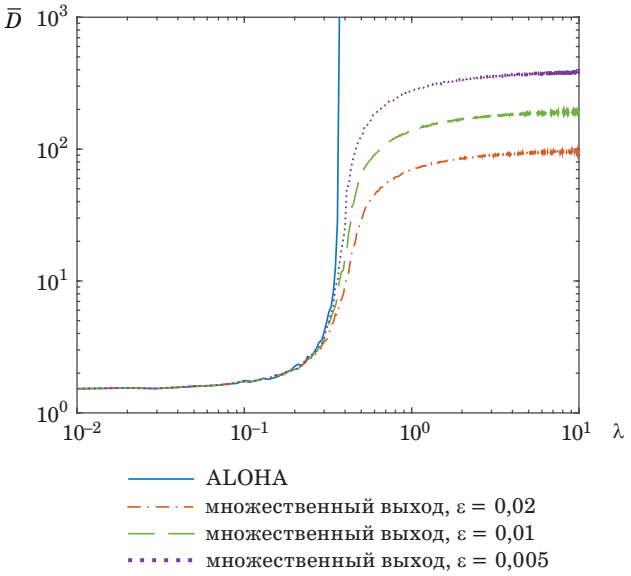
7. Количество устройств, покидающих систему, зависит от количества соседей устройства, у которого был «успех». Систему покидают устройства, находящиеся на расстоянии r от устройства, у которого был «успех». Через ε обозначим длину дуги окружности, равную $2r$.

В допущении 2 используется пространственный точечный пуассоновский процесс интенсивности λ. Это означает, что устройства появляются в системе в случайные моменты времени. Количество устройств, появившихся в одном окне, распределено по закону Пуассона с параметром λ. А устройства равномерно распределены по окружности. Модель данной системы представлена на рис. 5.

Зависимости средней задержки $\bar{D}$ от интенсивности входного потока λ для системы с множественным выходом и системы с алгоритмом ALOHA (рис. 6) получены с помощью имитационного моделирования при $T = 100\,000$. Для системы с множественным выходом приведено три зависимости при разных значениях ε (ε = 0,005, ε = 0,01, ε = 0,02). Полученные результаты иллюстрируют, что в системе с алгоритмом ALOHA задержка конечна при λ < e^{-1}, а в системе с мно-



■ ***Рис. 5.*** Модель системы с множественным выходом
■ ***Fig. 5.*** Model of a system with multiple departure



■ ***Рис. 6.*** Средняя задержка в системе с множественным выходом (при различных значениях ε) и в системе с алгоритмом ALOHA
■ ***Fig. 6.*** Average delay in the system with multiple departure (for different values of ε) and in the ALOHA algorithm

жественным выходом задержка конечна при любой интенсивности входного потока.

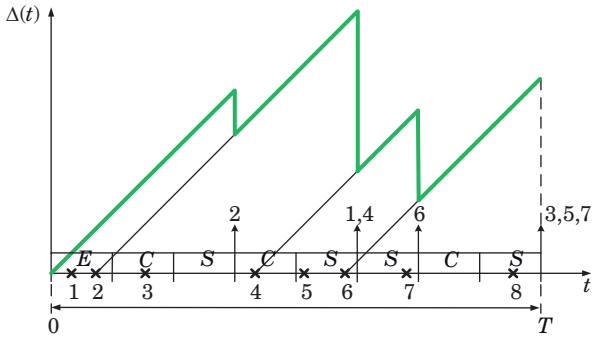
Средний возраст информации для системы с множественным выходом

Как и в системе с алгоритмом ALOHA, при вычислении среднего возраста информации в системе с множественным выходом учитываются только те сообщения, которые покинули систему не позже появившихся после них сообщений. В отличие от системы с алгоритмом ALOHA, в системе с множественным выходом в случае успешной передачи сообщения систему могут покинуть несколько устройств одновременно. Тогда при вычислении среднего возраста информации учитывается только то сообщение, которое появилось в системе последним. На рис. 7 представлен пример изменения возраста информации для системы с множественным выходом на интервале от 0 до T.

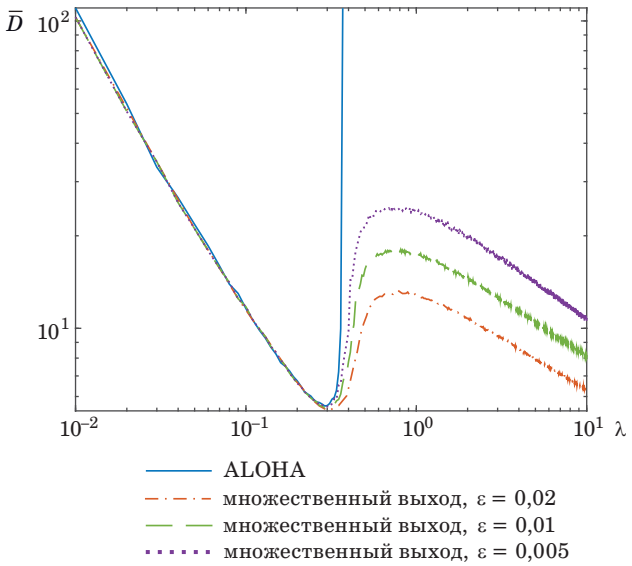
Зависимости среднего возраста информации от интенсивности входного потока для системы с множественным выходом при различных значениях параметра ε и системы с алгоритмом ALOHA (рис. 8) получены с помощью имитационного моделирования при $T = 100\,000$. Результаты показали, что в этой системе средний возраст информации, как и средняя задержка, конечен для любой интенсивности входного потока. Однако в отличие от средней задержки средний возраст информации не является монотонно возрастающей функцией от интенсивности входного потока.

Зависимости средней задержки и среднего возраста информации для системы с множественным выходом при ε = 0,01 и для системы с алгоритмом ALOHA показаны на рис. 9. Полученные результаты иллюстрируют необычные эффекты поведения данных характеристик. Проанализируем их на качественном уровне.

В системе с алгоритмом ALOHA средняя задержка с ростом интенсивности возрастает из-за увеличения количества конфликтов. Средний возраст информации в системе с алгоритмом ALOHA при низких интенсивностях входного потока имеет высокие значения из-за редкого появления сообщений в системе. С ростом интенсивности время между сообщени-

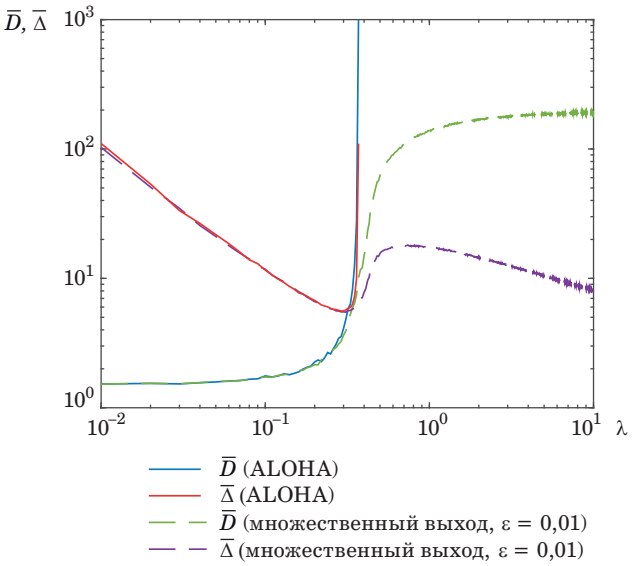


■ ***Рис. 7.*** Пример изменения возраста информации в системе с множественным выходом
■ ***Fig. 7.*** An example of changing the age of information in the system with multiple departure



■ ***Рис. 8.*** Средний возраст информации в системе с множественным выходом (при различных значениях ε) и в системе с алгоритмом ALOHA

■ ***Fig. 8.*** Average age of information in the system with multiple departure (for different values of ε) and in the ALOHA algorithm



■ ***Рис. 9.*** Средняя задержка и средний возраст информации в системе с множественным выходом и в системе с алгоритмом ALOHA

■ ***Fig. 9.*** Average delay and average age of information in the system with multiple departure and in the ALOHA algorithm

ями в системе сокращается, так как сообщения в системе появляются чаще. Это приводит к уменьшению среднего возраста информации. При интенсивности входного потока, близкой к λ_c, нарушается прямой порядок выхода сообщений из системы. Сообщения, которые вышли из системы не в свою очередь, учитываются при вычислении средней задержки и не учитываются при вычислении среднего возраста информации. Поэтому в данном случае средний возраст информации становится меньше средней задержки. При λ_c средняя задержка и средний возраст информации неограниченно возрастают, так как количество конфликтов становится слишком большим и сообщения прекращают выходить из системы.

В системе с множественным выходом при низких интенсивностях входного потока средняя задержка и средний возраст информации ведут себя так же, как и в системе с алгоритмом ALOHA. С ростом интенсивности средняя задержка сначала резко возрастает из-за увеличения количества конфликтов в системе, но потом практически стабилизируется (незначительно растет) из-за увеличения количества покидающих систему сообщений. Средний возраст информации с ростом интенсивности тоже резко возрастает, так как из-за большого числа конфликтов увеличивается промежуток между успешными передачами сообщений. Однако потом средний возраст информации убывает, так как увеличивается количество покидающих систему сообщений и сообщения чаще выходят из системы.

Таким образом, в обеих системах наблюдается интересный эффект: при низких интенсивностях входного потока средняя задержка меньше среднего возраста информации, а при высоких — больше.

Заключение

В данной работе была рассмотрена модель системы со случайным доступом и множественным выходом из работы [6], которая стабильна при любой интенсивности входного потока и потенциально неограниченном числе устройств в системе. Впервые предложен способ оценки среднего возраста информации для рассматриваемой системы и исследована зависимость среднего возраста информации от интенсивности входного потока. Показано, что средний возраст информации в системе с множественным выходом конечен при любой интенсивности входного потока, отличной от нуля.

В настоящее время такая характеристика, как средний возраст информации, начинает широко применяться для количественной оценки актуальности информации в системах передачи данных. С учетом этого результаты работы могут быть использованы при исследовании различных сценариев применения IoT, в которых важны вопросы стабильной работы системы при большом числе устройств.

В работе для исследования зависимости среднего возраста информации от интенсивности входного потока использовалось имитационное моделирование. Цель дальнейших исследований — детальное изучение случайных процессов, описывающих систему с множественным выходом, и получение значений для оценок среднего возраста информации без использования имитационного моделирования.

Финансовая поддержка

Исследование выполнено при финансовой поддержке Российского научного фонда, грант № 22-19-00305 «Пространственно-временные стохастические модели беспроводных сетей с большим числом абонентов».

Литература

1. **Ding J., Nemati M., Ranaweera C., and Choi J**. IoT connectivity technologies and applications: A survey. *IEEE Access*, 2020, vol. 8, pp. 67646–67673. doi:10.1109/ACCESS.2020.2985932
2. **Kaul S., Yates R., and Gruteser M.** Real-time status: How often should one update? *2012 Proc. IEEE INFOCOM*, IEEE, 2012, pp. 2731–2735. doi:10.1109/INFCOM.2012.6195689
3. **Abd-Elmagid M. A., Pappas N., and Dhillon H. S.** On the role of age of information in the Internet of Things. *IEEE Communications Magazine*, 2019, vol. 57(12), pp. 72–77. doi:10.1109/MCOM.001.1900041
4. **Bogatyrev V. A., Bogatyrev A. V., Bogatyrev S. V.** The probability of timeliness of a fully connected exchange in a redundant real-time communication system. *2020 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, 2020, pp. 1–4. doi:10.1109/WECONF48837.2020.9131517
5. **Foss S., Turlikov A., and Grankin M.** Spatial random multiple access with multiple departure. *IEEE Intern. Symp. on Information Theory (ISIT)*, IEEE, 2017, pp. 2728–2731. doi:10.1109/ISIT.2017.8007025
6. **Borisovskaya A., Glebov A., and Turlikov A.** Estimation of average delay in systems with unsourced random access and multiple departure. *2021 XVII Intern. Symp. "Problems of Redundancy in Information and Control Systems" (REDUNDANCY)*, IEEE, 2021, pp. 28–33. doi:10.1109/REDUNDANCY52534.2021.9606453
7. **Kosta A., Pappas N., and Angelakis V.** Age of information: A new concept, metric, and tool. *Foundations and Trends® in Networking*, 2017, vol. 12(3), pp. 162–259. doi:10.1561/1300000060
8. **Sun Y., Kadota I., Talak R., and Modiano E.** Age of information: A new metric for information freshness. *Synthesis Lectures on Communication Networks*, 2019, vol. 12(2), pp. 1–224.
9. **Yates R. D., Sun Y., Brown D. R., Kaul S. K., Modiano E., and Ulukus S.** Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communications*, 2021, vol. 39(5), pp. 1183–1210. doi:10.1109/JSAC.2021.3065072
10. **Inoue Y., Masuyama H., Takine T., and Tanaka T.** A general formula for the stationary distribution of the age of information and its application to single-server queues. *IEEE Transactions on Information Theory*, 2019, vol. 65(12), pp. 8305–8324. doi:10.1109/TIT.2019.2938171
11. **Yates R. D., and Kaul S. K.** The age of information: Real-time status updating by multiple sources. *IEEE Transactions on Information Theory*, 2019, vol. 65(3), pp. 1807–1827. doi:10.1109/TIT.2018.2871079
12. **Burkov A. A.** Signal power and energy-per-bit optimization problems in mMTC systems. *Информационно-управляющие системы*, 2021, № 5, с. 51–58. doi:10.31799/1684-8853-2021-5-51-58
13. **Kim D., Georgiev G., and Markovskaya N.** A model of random multiple access in unlicensed spectrum systems. *2022 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, 2022. doi:10.1109/WECONF55058.2022.9803810
14. **Chen X., Gatsis K., Hassani H., and Bidokhti S. S.** Age of information in random access channels. *IEEE Transactions on Information Theory. arXiv: 1912.01473v6*, 2022.
15. **Pan H., Chan T. T., Li J., and Leung V. C.** Age of information with collision-resolution random access. *IEEE Transactions on Vehicular Technology*, 2022, vol. 71(10), pp. 11295–11300. doi:10.1109/TVT.2022.3189399
16. **Feng J., Pan H., and Chan T. T.** Low-power random access for timely status update: Packet-based or connection-based? *arXiv preprint arXiv:2210.03962*, 2022.
17. **Munari A., and Frolov A.** Average age of information of irregular repetition slotted ALOHA. *GLOBECOM 2020-2020 IEEE Global Communications Conf.*, IEEE, 2020, pp. 1–6. doi:10.1109/GLOBECOM42002.2020.9322355
18. **Chen H., Gu Y., and Liew S. C.** Age-of-information dependent random access for massive IoT networks. *IEEE INFOCOM 2020 — IEEE Conf. on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, 2020, pp. 930–935. doi:10.1109/INFOCOMWKSHPS50562.2020.9162973
19. **Yates R. D., and Kaul S. K.** Age of information in uncoordinated unslotted updating. *IEEE Intern. Symp. on Information Theory (ISIT)*, IEEE, 2020, pp. 1759–1764. doi:10.1109/ISIT44484.2020.9174098
20. **Munari A.** Modern random access: An age of information perspective on irregular repetition slotted

ALOHA. *IEEE Transactions on Communications*, 2021, vol. 69(6), pp. 3572–3585. doi:10.1109/TCOMM.2021.3060429
21. **De Jesus G. G. M., Rebelatto J. L., and Souza R. D.** Age-of-information dependent random access in multiple-relay slotted ALOHA. *IEEE Access*, 2022, vol. 10, pp. 112076–112085. doi:10.1109/ACCESS.2022.3216616
22. **Чебунин М. Г., Фосс С. Г.** О стабильности систем случайного множественного доступа с минимальной обратной связью. *Сибирские электронные математические известия*, 2019, т. 16, с. 1805–1821. doi:10.33048/semi.2019.16.128
23. **Цыбаков Б. С., Михайлов В. А.** Свободный синхронный доступ пакетов в широковещательный канал с обратной связью. *Проблемы передачи информации*, 1978, т. 14, вып. 4, с. 32–59.
24. **Capetanakis J.** Tree algorithms for packet broadcast channels. *IEEE Transactions on Information Theory*, 1979, vol. 25, iss. 5, pp. 505–515. doi:10.1109/TIT.1979.1056093

UDC 621.391
doi:10.31799/1684-8853-2023-1-51-60
EDN: UBBHKD

Estimation of the average age of information in random access systems with multiple departure

A. V. Borisovskaya[a], Assistant Professor, orcid.org/0000-0002-0561-4226
A. M. Turlikov[a], Dr. Sc., Tech., Professor, orcid.org/0000-0001-7132-094X, turlikov@k36.org
[a]Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation

Introduction: An important direction in the study of the Internet of Things is the analysis of random multiple access systems that can provide stable operation with a large number of devices. An urgent task within this direction is the study of the average age of information for such systems. **Purpose:** To study the average age of information in random access systems with multiple departure. **Results:** Using a model with random access and multiple departure proposed by S. G. Foss in 2017, we describe the scenario of the Internet of Things, in which all devices are at the same distance from the base. It should be noted that the above mentioned model, unlike other random multiple access systems, ensures stable operation of the system with a potentially unlimited number of devices. In this paper, for the first time in relation to this model, we propose a method for computing the average age of information by the sequence of moments when messages enter the system and the sequence of moments when messages leave the system. We study the dependence of the average age of information on the input stream intensity. The research results show that the average age of information in a system with multiple departure is finite for any nonzero input stream intensity. **Practical relevance:** The proposed method for computing the average age of information for a system with multiple departure allows to compare scenarios of Internet of Things systems by this indicator. It also allows to determine the feasibility of using systems with multiple departure, taking into account the specifics of the scenario under consideration. **Discussion:** In this paper, we consider a simplified model of random access system with multiple departure. However, the results obtained for this model can be generalized for more complex models.

Keywords — Internet of Things, average age of information, random multiple access, multiple departure.

For citation: Borisovskaya A. V., Turlikov A. M. Estimation of the average age of information in random access systems with multiple departure. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 51–60 (In Russian). doi:10.31799/1684-8853-2023-1-51-60, EDN: UBBHKD

Financial support
The study was financially supported by the Russian Science Foundation, grant No. 22-19-00305 “Spatio-temporal stochastic models of wireless networks with a large number of subscribers”.

References

1. Ding J., Nemati M., Ranaweera C., and Choi J. IoT connectivity technologies and applications: A survey. *IEEE Access*, 2020, vol. 8, pp. 67646–67673. doi:10.1109/ACCESS.2020.2985932
2. Kaul S., Yates R., and Gruteser M. Real-time status: How often should one update? *2012 Proc. IEEE INFOCOM*, IEEE, 2012, pp. 2731–2735. doi:10.1109/INFCOM.2012.6195689
3. Abd-Elmagid M. A., Pappas N., and Dhillon H. S. On the role of age of information in the Internet of Things. *IEEE Communications Magazine*, 2019, vol. 57(12), pp. 72–77. doi:10.1109/MCOM.001.1900041
4. Bogatyrev V. A., Bogatyrev A. V., Bogatyrev S. V. The probability of timeliness of a fully connected exchange in a redundant real-time communication system. *2020 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, 2020, pp. 1–4. doi:10.1109/WECONF48837.2020.9131517
5. Foss S., Turlikov A., and Grankin M. Spatial random multiple access with multiple departure. *IEEE Intern. Symp. on Information Theory (ISIT)*, IEEE, 2017, pp. 2728–2731. doi:10.1109/ISIT.2017.8007025
6. Borisovskaya A., Glebov A., and Turlikov A. Estimation of average delay in systems with unsourced random access and multiple departure. *2021 XVII Intern. Symp. “Problems of Redundancy in Information and Control Systems” (REDUNDANCY)*, IEEE, 2021, pp. 28–33. doi:10.1109/REDUNDANCY52534.2021.9606453
7. Kosta A., Pappas N., and Angelakis V. Age of information: A new concept, metric, and tool. *Foundations and Trends® in Networking*, 2017, vol. 12(3), pp. 162–259. doi:10.1561/1300000060
8. Sun Y., Kadota I., Talak R., and Modiano E. Age of information: A new metric for information freshness. *Synthesis Lectures on Communication Networks*, 2019, vol. 12(2), pp. 1–224.
9. Yates R. D., Sun Y., Brown D. R., Kaul S. K., Modiano E., and Ulukus S. Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communications*, 2021, vol. 39(5), pp. 1183–1210. doi:10.1109/JSAC.2021.3065072

10. Inoue Y., Masuyama H., Takine T., and Tanaka T. A general formula for the stationary distribution of the age of information and its application to single-server queues. *IEEE Transactions on Information Theory*, 2019, vol. 65(12), pp. 8305–8324. doi:10.1109/TIT.2019.2938171
11. Yates R. D., and Kaul S. K. The age of information: Real-time status updating by multiple sources. *IEEE Transactions on Information Theory*, 2019, vol. 65(3), pp. 1807–1827. doi:10.1109/TIT.2018.2871079
12. Burkov A. A. Signal power and energy-per-bit optimization problems in mMTC systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 51–58. doi:10.31799/1684-8853-2021-5-51-58
13. Kim D., Georgiev G., and Markovskaya N. A model of random multiple access in unlicensed spectrum systems. *2022 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, 2022. doi:10.1109/WECONF55058.2022.9803810
14. Chen X., Gatsis K., Hassani H., and Bidokhti S. S. Age of information in random access channels. *IEEE Transactions on Information Theory. arXiv: 1912.01473v6*, 2022.
15. Pan H., Chan T. T., Li J., and Leung V. C. Age of information with collision-resolution random access. *IEEE Transactions on Vehicular Technology*, 2022, vol. 71(10), pp. 11295–11300. doi:10.1109/TVT.2022.3189399
16. Feng J., Pan H., and Chan T. T. Low-power random access for timely status update: Packet-based or connection-based? *arXiv preprint arXiv:2210.03962*, 2022.
17. Munari A., and Frolov A. Average age of information of irregular repetition slotted ALOHA. *GLOBECOM 2020-2020 IEEE Global Communications Conf.*, IEEE, 2020, pp. 1–6. doi:10.1109/GLOBECOM42002.2020.9322355
18. Chen H., Gu Y., and Liew S. C. Age-of-information dependent random access for massive IoT networks. *IEEE INFOCOM 2020 – IEEE Conf. on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, 2020, pp. 930–935. doi:10.1109/INFOCOMWKSHPS50562.2020.9162973
19. Yates R. D., and Kaul S. K. Age of information in uncoordinated unslotted updating. *IEEE Intern. Symp. on Information Theory (ISIT)*, IEEE, 2020, pp. 1759–1764. doi:10.1109/ISIT44484.2020.9174098
20. Munari A. Modern random access: An age of information perspective on irregular repetition slotted ALOHA. *IEEE Transactions on Communications*, 2021, vol. 69(6), pp. 3572–3585. doi:10.1109/TCOMM.2021.3060429
21. De Jesus G. G. M., Rebelatto J. L., and Souza R. D. Age-of-information dependent random access in multiple-relay slotted ALOHA. *IEEE Access*, 2022, vol. 10, pp. 112076–112085. doi:10.1109/ACCESS.2022.3216616
22. Chebunin M. G., and Foss S. G. On stability of multiple access systems with minimal feedback. *Sibirskie Elektronnye Matematicheskie Izvestiya* [Siberian Electronic Mathematical Reports], 2019, vol. 16, pp. 1805–1821 (In Russian). doi:10.33048/semi.2019.16.128
23. Tsybakov B. S., and Mikhailov V. A. Free synchronous packet access in a broadcast channel with feedback. *Problems of Information Transmission*, 1978, vol. 14, iss. 4, pp. 259–280 (In Russian).
24. Capetanakis J. Tree algorithms for packet broadcast channels. *IEEE Transactions on Information Theory*, 1979, vol. 25, iss. 5, pp. 505–515. doi:10.1109/TIT.1979.1056093

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.

UDC 621. 391
doi:10.31799/1684-8853-2023-1-61-70
EDN: PCFDIS

Articles

Signal detection amid noise using order statistics: detector sensitivity analysis and parameter choice

D. S. Osipov[a,b], *PhD, Tech., Senior Reseacher, orcid.org/0000-0003-0400-7181, d_osipov@iitp.ru*
[a]*Kharkevich Institute for Information Transmission Problems of the RAS, 19, bld.1, Bolshoy Karetny per., 127051, Moscow, Russian Federation*
[b]*National Research University «Higher School of Economics», 20, Myasnitskaya St., 101000, Moscow, Russian Federation*

Introduction: *Developing physical level techniques for machine type communications is a challenging task. In particular developing low complexity channel estimation with acceptable precision or maintaining accurate power control is cumbersome. One possible way to solve this problem is to use the reception techniques based on order statistics that do not require any form of channel estimation or power control. This paper deals with a communication system that uses frequency shift keying in a dynamically allocated instantaneous frequency band and an order-statistics-based receiver previously proposed by the author.* ***Purpose:*** *To analyze the sensitivity of the system under consideration to both noise and interference power variation and to explore the receiver parameter choice.* ***Results:*** *Simulation-based capacity analysis demonstrates that the receiver is resistant to signal-to-noise variations and thus can provide the desired performance even if the signal of the user under consideration is subject to drastic power variation. It is demonstrated that the number of possible symbol values being assigned maximum reliability optimized for the worst case number of users yields capacity close to optimal for a lower number of users. Finally, simulation confirms that the performance of the communication system under consideration is not dependent on the choice of reliability values. Thus values that minimize hardware complexity can be chosen.* ***Practical relevance:*** *The results obtained prove that the detector under consideration is practically usable and can be applied in a variety of real-life scenarios. The hardware complexity can be minimized while preserving the performance.*

Keywords *— machine type communications, frequency hopping, dynamically allocated hopset, noncoherent reception, α-detector, channel capacity.*

For citation: Osipov D. S. Signal detection amid noise using order statistics: detector sensitivity analysis and parameter choice. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 61–70. doi:10.31799/1684-8853-2023-1-61-70, EDN: PCFDIS

Introduction

Machine-type communications (MTC) are expected to play prevailing role in the future communication systems development [1, 2]. The signals transmitted in these systems are subject to severe mixed interference that is due to both multi-user interference and background noise. Thus channel estimation with desired precision requires high computational complexity [3, 4]. Power control is also very challenging [5–7] since closed loop methods cannot be used due to limitations on packet size and energy consumption as well as the burst-like traffic. Thus traditional techniques that are used in conventional digital communications systems turn out to be ill-suited for the problem in question. There are basically two approaches to address the problem. The first one makes use of the blind or semi-blind detection techniques with error correction coding and interference cancellation in order to enable multi-user reception [8–10]. Despite the many attractive features this approach enjoys interference cancellation leads to error propagation risk and requires the information on the number of active users. Successive interference cancellation also leads to increased delay that can be undesirable especially in critical machine type communications. Another approach is to use single user reception techniques that can withstand severe interference. In recent decades several order statistics-based reception techniques that meet this requirement were proposed in [11–15].

This paper deals with the α-detector proposed in [15] for the communication system that will be described in the following section. The α-detector is an order statistics-based receiver that uses only measurements obtained from the channel (powers of the signals) and avoids using any kind of side information by using measurements ordering only and assigning one reliability value to the α possible symbol values that correspond to signals with greatest powers and another to the remaining $q - \alpha$ ones, then using the reliability values in question for soft-input decoding of the error correction code in use. Thus both the value of the parameter α and the reliability values are to be chosen. Herein below the way those parameters are chosen affects the performance of the system is studied and the ways to choose those parameters in optimal way are revealed. The paper also deals with the problem of the system sensitivity to both multi-user interference and background noise power variations. Although some results on the matter were obtained as a by-product in [16–17] this problem has never been studied comprehensively.

Transmission and reception

Let us consider a single user reception in an uplink transmission scenario. Apart from the user under consideration K active users are assumed to transmit to the base station (following [15] those will be referred to as "interfering"). Each of the K active users is assumed to transmit codewords of an $C(N, k, d)$ code, each symbol of the codeword being mapped into a weight 1 vector of length q. Thus each codeword is mapped into a binary matrix $\mathbf{M}$ of the size $q \times N$. Each column of $\mathbf{M}$ contains a weight 1 vector. The matrix $\mathbf{M}$ is then complemented with an all-zero matrix $\mathbf{Z}$ of the size $(q - Q) \times N$:

$$\tilde{\mathbf{M}} = \begin{bmatrix} \mathbf{M} \\ \mathbf{Z} \end{bmatrix}.$$

The matrix $\tilde{\mathbf{M}}$ is then permuted column-wise (permutations are assumed to be independent, equiprobable and pseudo-random) and transmitted via the channel using OFDM.

The reception starts with the reception of N OFDM symbols corresponding to the codeword and inverse permutations of the corresponding columns in frequency domain. The receiver then extracts the first q rows of the resulting matrix and squares its elements to obtain the matrix $\mathbf{\Omega}$.

The α-detector

The α-detector has been proposed in [15]. It is convenient to consider the performance of the detector in question as a two stage process. Within the first stage reliability estimates for each symbol are to be assigned. The α-detector assigns reliability value λ_1 to the α possible values of the symbols that corresponds to the α subcarriers with greatest energies λ_0 to the remaining $q - \alpha$ ones. The second stage boils down to aggregating reliability estimates of the symbols to compute reliability estimates for each codeword and choosing a codeword with maximum reliability. Thus the α-detector is essentially a soft-input order statistics-based decoder/demodulator. To introduce the detector in question in a more formal way let us derive $\mathbf{\Omega}^{\downarrow}$ — the matrix obtained by sorting the matrix $\mathbf{\Omega}$ column-wise in descending order. Each element of the matrix of the reliability estimates $\mathbf{D}_{(\alpha,\lambda_0,\lambda_1)}$ is given by

$$\mathbf{D}_{(\alpha,\lambda_0,\lambda_1)}(t, z) = \begin{cases} \lambda_0 & \mathbf{\Omega}(t, z) < \mathbf{\Omega}^{\downarrow}(t, \alpha) \\ \lambda_1 & \mathbf{\Omega}(t, z) \geq \mathbf{\Omega}^{\downarrow}(t, \alpha) \end{cases},$$

where t is the column number; z is the row number; $\mathbf{\Omega}^{\downarrow}(t, \alpha)$ is the α greatest element of the t-th column, and $\lambda_1 > \lambda_0 \geq 0$, where λ_1 is the reliability assigned to the α possible values of the t-th symbol corresponding to α greatest elements of the column and λ_0 is the reliability of the remaining $q - \alpha$ values.

The decoder then computes the reliability estimate for the g-th codeword

$$S_g = \sum_{j=1}^{q} \sum_{k=1}^{N} \left(\mathbf{D}_{(\alpha,\lambda_0,\lambda_1)}(j, k) \cdot \mathbf{X}_g(j, k) \right), \qquad (1)$$

for each codeword and chooses

$$g^* = \arg \max_{g \in [1,\ldots,M]} S_g, \qquad (2)$$

where g^* is the set of numbers of codewords that correspond to maximal reliability value. If $|g^*| = 1$, i.e. there is only one such codeword, the detector declares this codeword to be the decoded codeword. Otherwise decoding failure is declared.

An equivalent (α, p)-channel and its capacity

The reception process described above corresponds to the vector channel depicted in Fig. 1. In the following section an equivalent (α, p)-channel introduced in [16] will be considered. In our consideration of the channel in question we follow notation introduced in [17].

Let us define

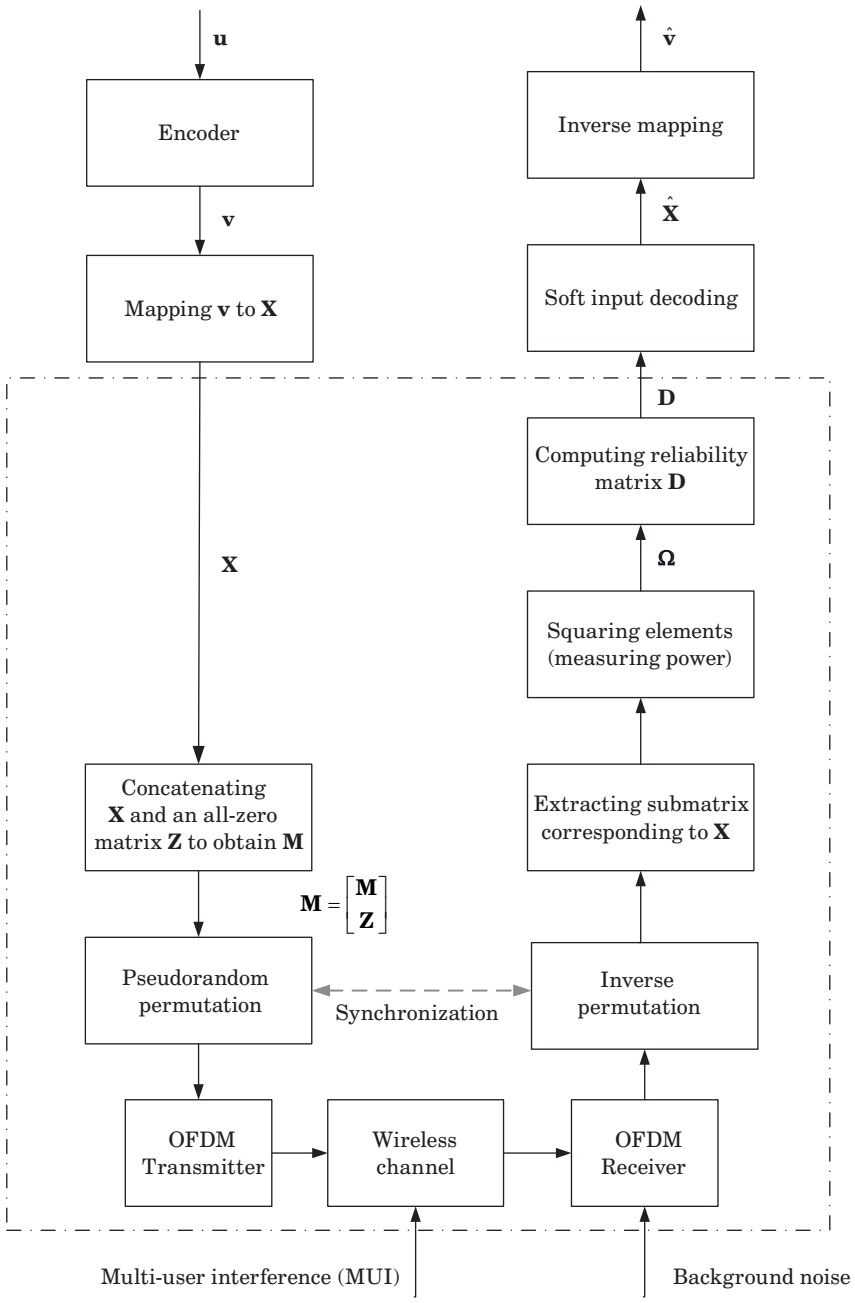
$$\mathbb{B}_q^x = \left\{ \mathbf{b} = \left(b_1, \ldots, b_q\right)^{\mathrm{T}} : \forall \ i \in \{1 : q\} b_i \in \{0, 1\}, \right.$$
$$\left. w_H(\mathbf{b}) = x \right\},$$

where $\mathbb{B}_q^x$ is the set of all binary column vectors of length q with Hamming weight x and the sets

$$\mathbb{S}^1(\mathbf{z}, \alpha) = \left\{ \mathbf{s} : \mathbf{s} \in \mathbb{B}_q^{\alpha}, \ \mathbf{s} \wedge \mathbf{z} = \mathbf{z} \right\};$$
$$\mathbb{S}^0(\mathbf{z}, \alpha) = \left\{ \mathbf{s} : \mathbf{s} \in \mathbb{B}_q^{\alpha}, \ \mathbf{s} \wedge \mathbf{z} \neq \mathbf{z} \right\}.$$

The channel is defined in the following way

$$\forall \ \alpha \geq 2; \ q > \alpha; \ \mathbf{x} \in \mathbb{B}_q^1, \mathbf{y} \in \mathbb{B}_q^{\alpha} \ \frac{1}{2} < p < 1;$$
$$\sum_{\mathbf{y} \in \mathbb{S}^1(\mathbf{x},\alpha)} p(\mathbf{y} \mid \mathbf{x}) = p;$$
$$\forall \ \mathbf{x} \in \mathbb{B}_q^1, \mathbf{y}_a \in \mathbb{S}^1(\mathbf{x}, \alpha), \mathbf{y}_b \in \mathbb{S}^1(\mathbf{x}, \alpha),$$
$$a \neq b : p(\mathbf{y}_a \mid \mathbf{x}) = p(\mathbf{y}_b \mid \mathbf{x}) = p_1;$$
$$\forall \ \mathbf{x} \in \mathbb{B}_q^1, \mathbf{y}_n \in \mathbb{S}^0(\mathbf{x}, \alpha), \mathbf{y}_l \in \mathbb{S}^0(\mathbf{x}, \alpha),$$
$$n \neq l : p(\mathbf{y}_n \mid \mathbf{x}) = p(\mathbf{y}_l \mid \mathbf{x}) = p_0. \qquad (3)$$

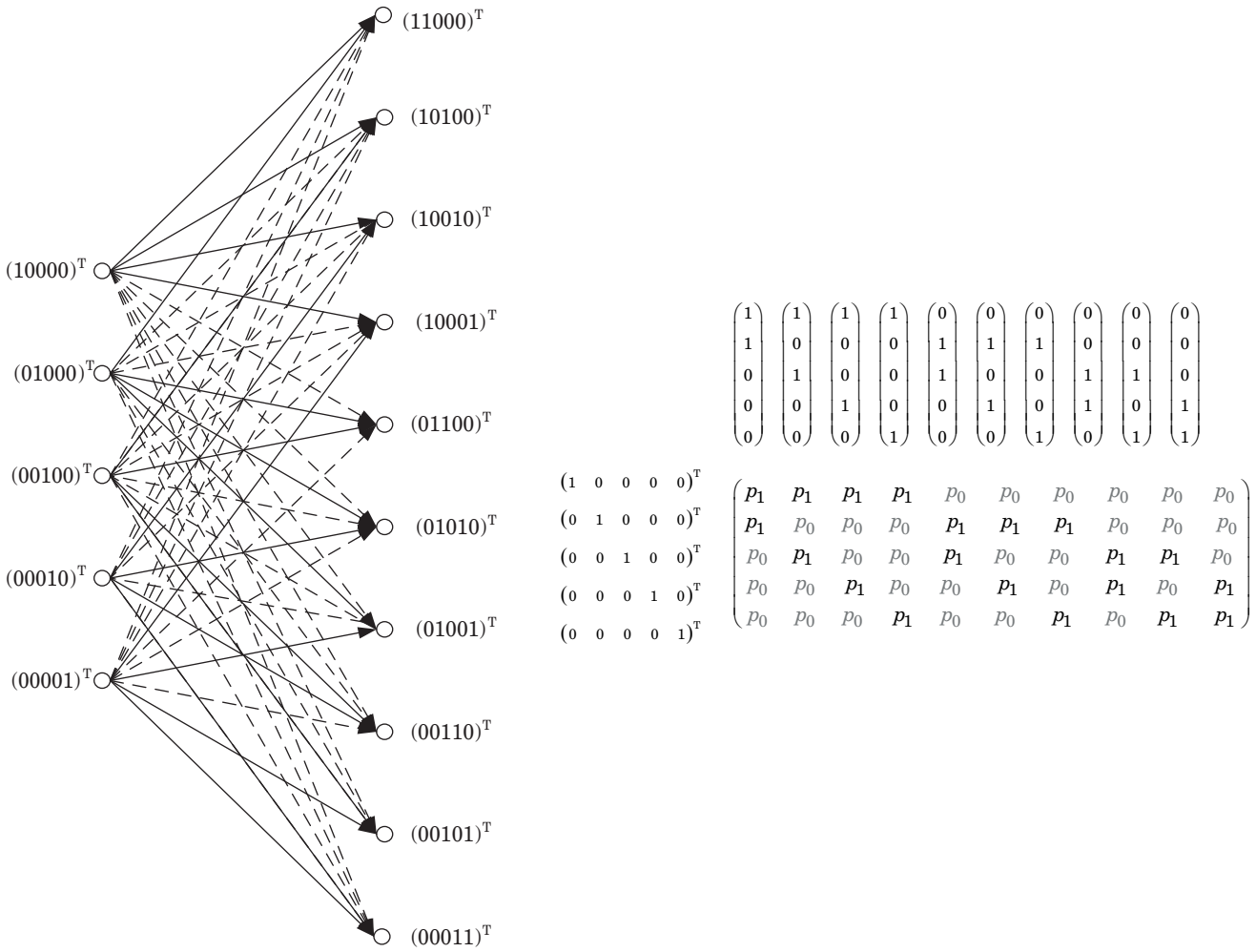


■ ***Fig. 1.*** The communication system and the vector channel under consideration

Equation (3) can be interpreted in the following way: the detector forms a list of α subcarriers with greatest output energy. The probability p is then interpreted as the probability of the fact that the subcarrier that corresponds to the symbol sent will be in the list. An example of the equivalent (α, p)-channel diagram and transition matrix (for the case $q = 5$ and $\alpha = 2$) is given in Fig. 2. In order to keep the size of the output alphabet moderate we have chosen relatively small values of q and α respectively. As can be seen from Fig. 2 since $\alpha = 2$ any vector of weight 2 can appear at the output of the channel since each time the detector chooses α-subchannels with maximum energy and assigns reliability λ_1 to the respective symbol values (these symbols are labelled with 1. Later it will be shown that the output vectors can be represented as binary ones and this representation doesn't affect the performance of the decoder).

Analytical expression for the channel capacity of the (α, p)-channel was obtained in [16] and is given by

$$C_{(\alpha,p)} = \log_2(q) - p\log_2(\alpha) - (1-p)\times \\ \times \log_2(q-\alpha) - H(p), \quad (4)$$



■ ***Fig. 2.*** The equivalent (α, p)-channel diagram and transition matrix (for $q = 5$ and $\alpha = 2$)

where $H(p)$ is the binary entropy given by

$$H(p) = -p\log_2(p) - (1-p)\log_2(1-p).$$

To describe the performance of the equivalent channel we will use the normalized version of the equation (4)

$$C^{\alpha}_{norm}(\alpha, p) = \frac{C_{(\alpha,p)}}{\log_2(q)} =$$

$$= 1 - \frac{p\log_2(\alpha) + (1-p)\log_2(q-\alpha) + H(p)}{\log_2(q)}. \quad (5)$$

Communication system performance evaluation: simulation-based channel capacity-aided approach

In what follows we will use the normalized channel capacity given by (5) in order to reveal some of the properties of the communication system employing α-detector. The probability p depends on the background noise, multi-user interference and various parameters such as α, q and Q. In what follows simulation will be used in order to obtain $\hat{p}$ an estimate of the probability p that will be used to compute $C^{\alpha}_{norm}(\alpha, \hat{p})$. It is thus essential to describe the simulation setup in use.

Let us assume that the cardinality of the set of all subcarriers available to the users is set to $Q = 4096$. The background noise is modelled as additive white Gaussian noise (AWGN) described by signal-to-noise ratio (SNR)

$$SNR = 10\log_{10}\left(\frac{E_s}{\log_2(q)E_N}\right),$$

where E_s is the energy of the signal transmitted by the user under consideration; E_N is the noise energy (in the entire band). The multi-user interference is modeled in the following way: we assume that each of the K interfering users transmit signals similar to that of the user under consideration. The phase

of each signal is modeled as a random variable with circular uniform distribution. The power of all signals transmitted by each user is the same P_t while the power at the receiver end for each (say i-th) interfering user $P_{r,i}$ is derived in the following way: the power ratio has log-normal distribution [18], i.e. [19]

$$L(d_i) = 10\log\left(\frac{P_t}{P_{r,i}}\right) =$$

$$= \bar{L}(d_i) + X_\sigma = L_{fs}(d_0) + 10\gamma_0 \log\left(\frac{d_i}{d_0}\right) + X_\sigma,$$

where d_i is the distance between the i-th interfering user; X_σ is a Gaussian random variable with mean 0 and variance σ; L_{fs} is the free-space path loss given by Friis law [20]; d_0 is the reference distance and γ_0 is the path-loss exponent, whereas the respective power ratio for the user under consideration also has log-normal distribution and is given by:

$$L(d^*) = 10\log\left(\frac{P_t}{P_r^*}\right) =$$

$$= \bar{L}(d^*) + X_\sigma = L_{fs}(d_0) + 10\gamma_0 \log\left(\frac{d^*}{d_0}\right) + X_\sigma,$$

where d^* is the distance between the user under consideration and the receiver; P_r^* is the power of the signal transmitted by the user under consideration at the receiver end and d_0 is reference distance. The signal-to-interference ratio thus depends on the values

$$\mu_i = \log_{10}\frac{d^*}{d^i},$$

where $i \in \{1, \ldots, K\}$.

In what follows we assume, that the values μ_i are random variables equiprobably distributed on a one dimensional gird. In particular we consider 2 scenarios. Within the scope of the Scenario 1 the values of μ_i are assumed to be equiprobably distributed on [0:0.01:2]. This scenario thus boils down to the assumption that $d_i \le d^*$ and therefore $\bar{L}(d_i) \le \bar{L}(d^*)$. Within the scope of the Scenario 2 the values of μ_i are assumed to be equiprobably distributed on [–2:0.01:2].

Let us first consider the problem of finding the optimal value of the parameter α. Fig. 3, *a* and *b* depicts the dependency of normalized capacity on the value of α for $q = 256$, different values of K, SNR and Scenario 1 and Scenario 2 respectively.

First and foremost it is worth pointing out that for all scenarios and parameters under consideration there is a value of α that corresponds to the maximum value of $C_{norm}^{\alpha}(\alpha, \hat{p})$. In what follows this value will be referred to as optimal. The optimal value lies within the range $(q/16, q/2)$. The maximum value of $C_{norm}^{\alpha}(\alpha, \hat{p})$ is less than 0.5, i.e. only relatively low rate codes can be used. For small α ($\alpha \le q/16$) the $C_{norm}^{\alpha}(\alpha, \hat{p})$ is close to zero. Thus conventional frequency-shift keying (FSK) demodulation (i.e. the case that corresponds to $\alpha = 1$) is not suitable. For large α normalized capacity doesn't depend on the number of interfering users or SNR. The optimal value of the parameter α for smaller number of interfering users is less than that for the greater number of interfering users. The maximum value of $C_{norm}^{\alpha}(\alpha, \hat{p})$ decreases as K increases (for the parameters under consideration 200 more interfering users results in approximately 15% decrease in maximum normalized capacity that can be obtained) but remains the same when SNR varies in a broad interval (10 dB in our case). For comparison we present dependences of normalized capacity on the value of α for $q = 16$, different values of K, SNR and Scenario 1 and Scenario 2 (Fig. 4, *a* and *b* respectively).

The observations we made for $q = 256$ are valid for $q = 16$. Again one can notice that normalized capacity depends on the number of interfering users but not on the SNR value. Although it is true in a very broad range of SNR values it is interesting to find out the region where this doesn't hold. To do so we present results for fixed K ($K = 500$) and q ($q = 256$) and plot values of the normalized capacity for different SNR values (Fig. 5, *a* and *b* respectively).

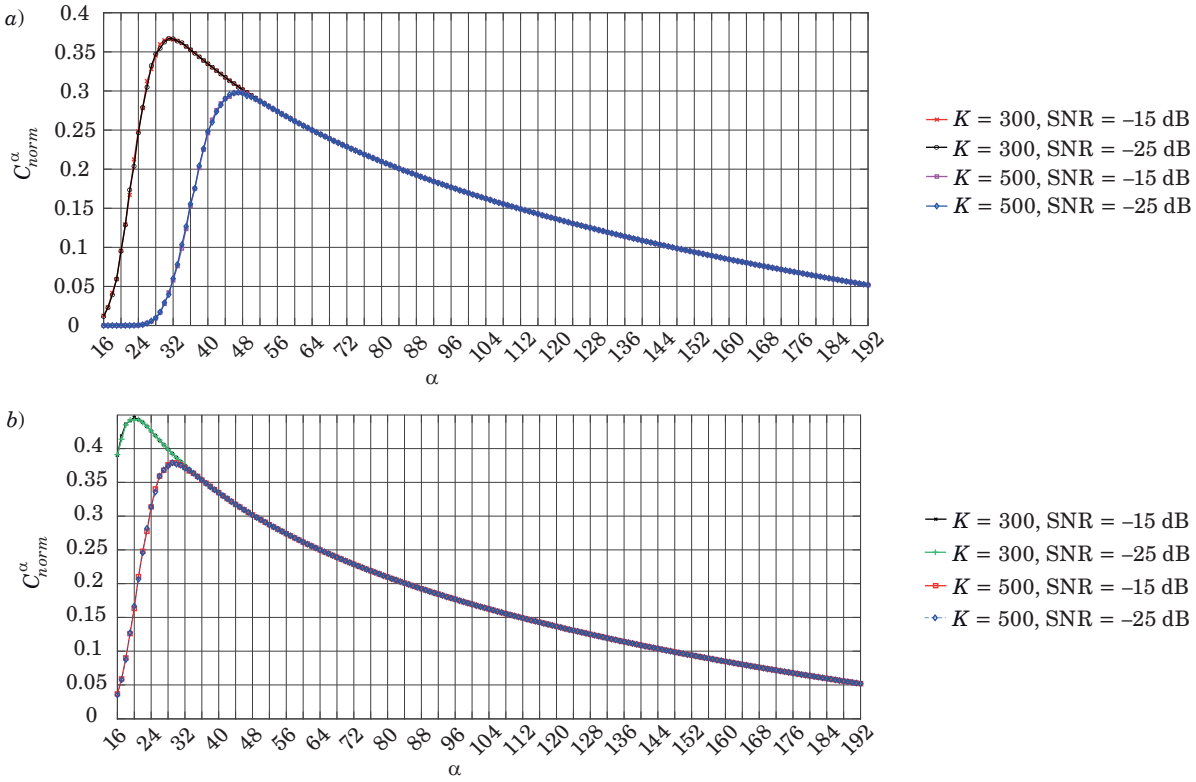
Let us note that the curves for SNR from –30 to –34 dB almost coincide for both scenarios. However for SNR < –34 normalized capacity decreases slowly as SNR decrease. For comparison results for $q = 16$ are presented in Fig. 6, *a* and *b* respectively.

For small q the trend is similar although normalized capacity decreases slowly as SNR decreases for SNR < –32. Nevertheless one can argue that normalized capacity of the equivalent channel (for fixed values of α and K) is almost the same for a very broad range of SNR values.

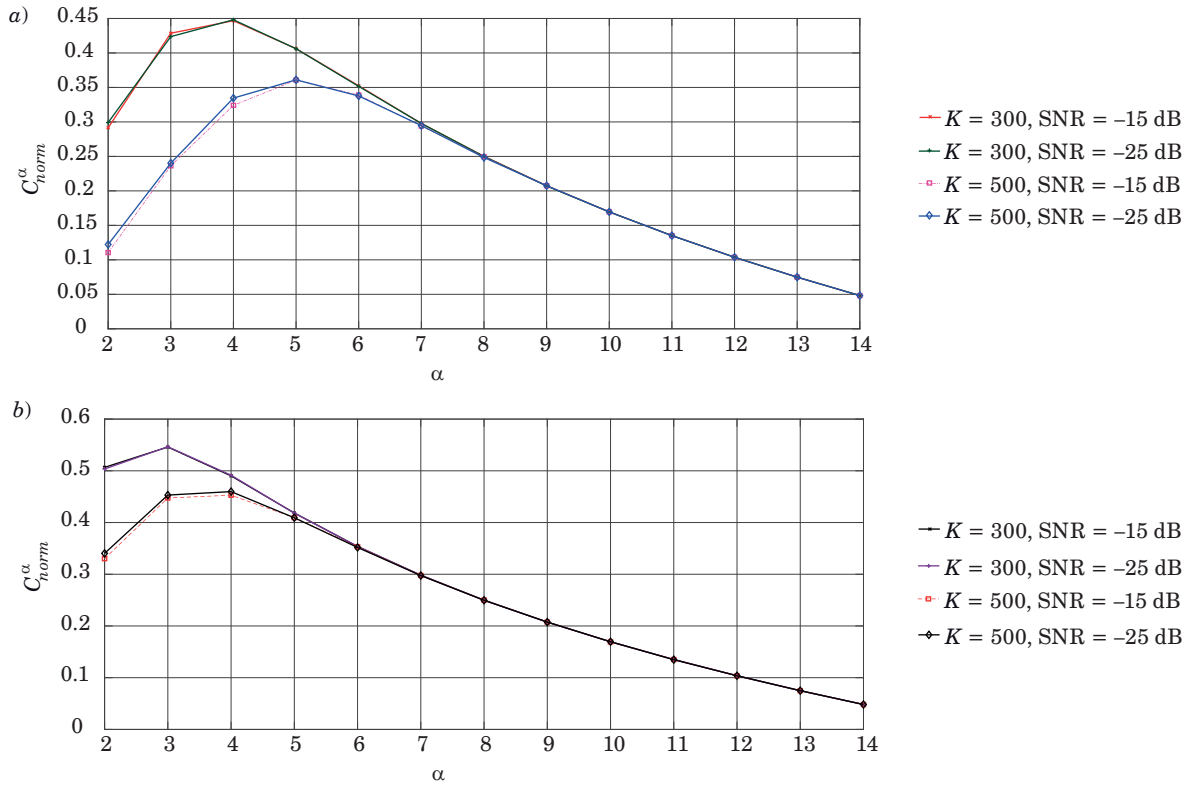
Let us now summarize our findings made in this section:

– the capacity of the equivalent channel remains the same for a very wide range of SNR values. Thus even if the power of the signal at the receiver side exhibits drastic variation (e.g. due to small scale fading or the transmitter mobility) the performance of the detector will not degrade unless SNR is not close to the threshold value;

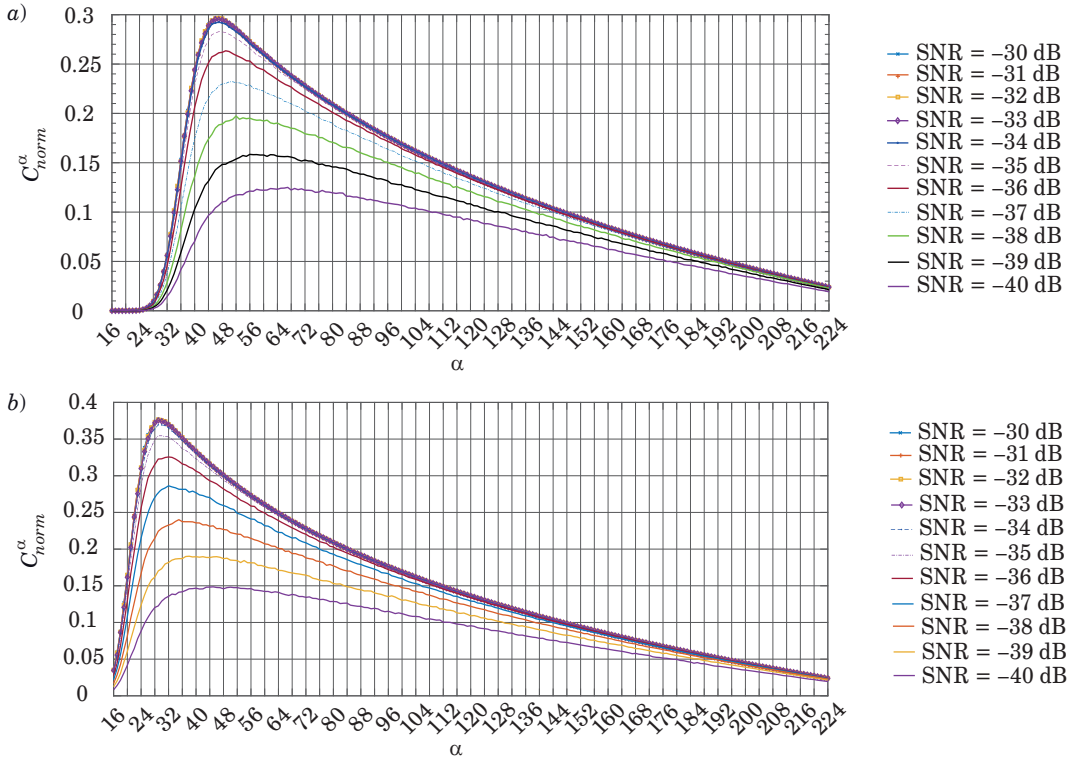
– the capacity of the equivalent channel depends on the power of the multi-user interference and thus on the number of interfering users. However the capacity degrades slowly as the number of interfering users increases. Thus even if the number of



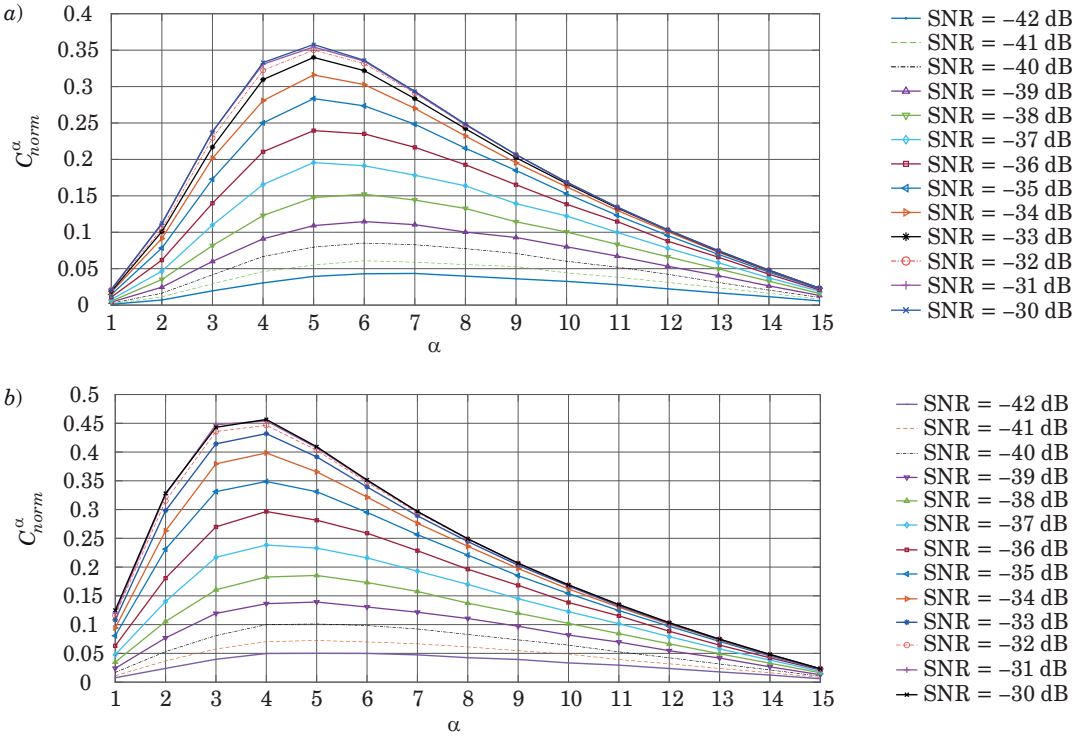
■ ***Fig. 3.*** Normalized capacity of the equivalent (α, p)-channel vs. α for $q = 256$, various values of K (number of interfering users), SNR and Scenario 1 (*a*) and Scenario 2 (*b*)



■ ***Fig. 4.*** Normalized capacity of the equivalent (α, p)-channel vs. α for $q = 16$, various values of K (number of interfering users), SNR and Scenario 1 (*a*) and Scenario 2 (*b*)



■ ***Fig. 5.*** Normalized capacity of the equivalent (α, p)-channel vs. α for $q = 256$, $K = 500$, various SNR values and Scenario 1 (*a*) and Scenario 2 (*b*)



■ ***Fig. 6.*** Normalized capacity of the equivalent (α, p)-channel vs. α for $q = 16$, various SNR values and Scenario 1 (*a*) and Scenario 2 (*b*)

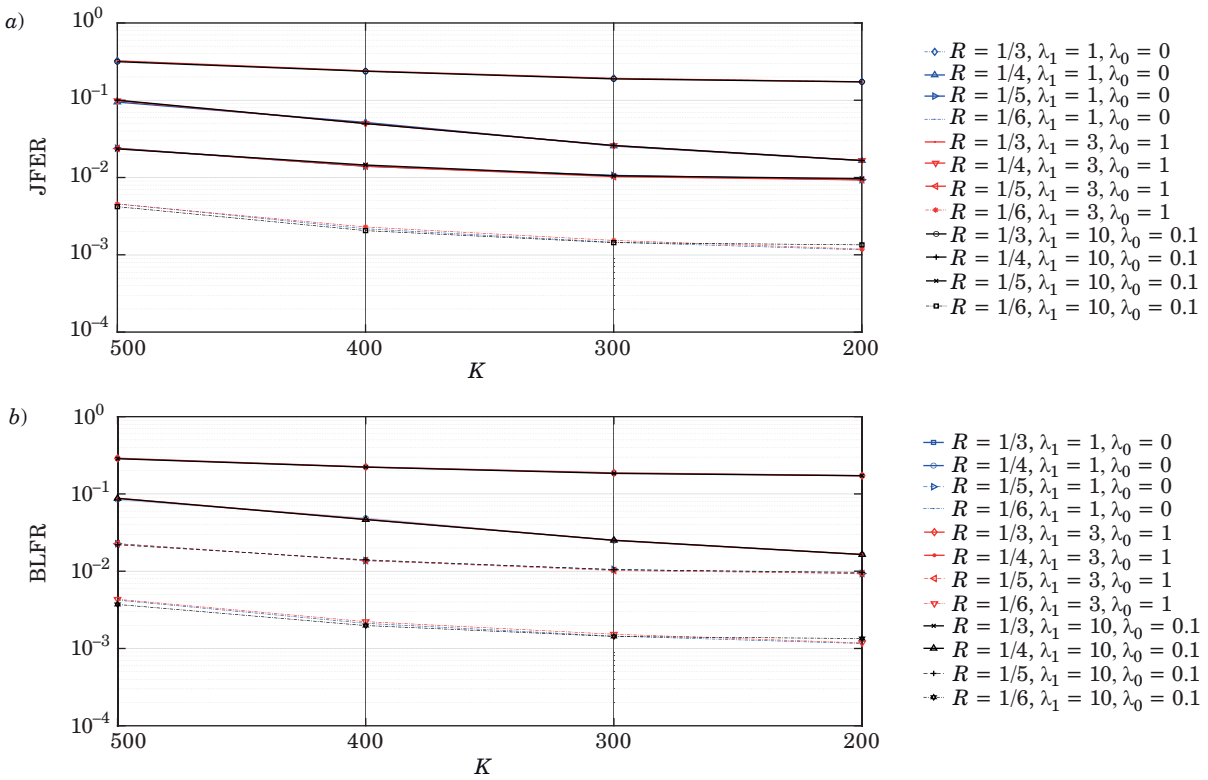
interfering users change the performance will not change in any meaningful way;

— for any K there exists value of the α parameter that results in channel with highest capacity. For different K this value is different however the optimal value changes slowly as K increases (decreases). Thus onse a certain value of α is fixed the performance of the detector will be very close to optimal even if the number of interfering users will vary.

Communication system performance evaluation: simulation-based performance-aided approach

In our consideration of the equivalent (α, p)-channel reliability values were not used since channel capacity doesn't depend on labelling. However since reliabilities of the codewords depend on the reliabilities values it is essential to find out how the choice of λ_1 and λ_0 affects the detector performance. In what follows we shall use the same simulation scenario we described above. The error correction codes in use are maximum distance separable codes obtained by appropriate puncturing Reed — Solomon code $C_{16}(15, 2, 14)$ to the desired rate. Since detection given by the rule (2) can result in both decoding failure and erroneous decoding in what follows we shall consider joint probability that either decoding failure or error will take place — Joint Failure and Error Rate (JFER). For different values of the number of interfering users we use the same value of α, the one that maximizes the capacity for the maximum number of interfering users under consideration, i.e. $K = 500$ and SNR = –25 dB. JFER vs. the number of interfering users is plotted for various values of λ_0 and λ_1 and Scenario 1. As can be seen from Fig. 7, *a*, the performance of the communication system that uses α-detector remains the same for different values of λ_0 and λ_1 for any rate R under consideration.

For comparison we present curves for Block Failure Rate (BLFR) vs. the number of interfering users for the same values λ_0 and λ_1 and code rate R (Fig. 7, *b*). BLFR is the same for different values of λ_0 and λ_1 for any rate R under consideration. Thus performance doesn't depend on the values of λ_0 and λ_1 (as long as $0 \le \lambda_0 < \lambda_1$) and therefore $\lambda_0 = 0$ and $\lambda_1 = 1$ can be used. Although the matrix $\mathbf{D}_{(\alpha,\lambda_0,\lambda_1)}$ need not be stored since codeword reliability values (1) can be computed on the fly the values of the codeword reliability values in question (both intermediate and final values) should be stored. Thus the choice $\lambda_0 = 0$ and $\lambda_1 = 1$ minimizes the hardware complexity since there is no need to store any values for λ_0 and λ_1 and the value for each codeword



■ ***Fig. 7.*** JFER (*a*) and BLFR (*b*) vs. the number of interfering users K for various values of λ_0 and λ_1, code rates R and Scenario 1

requires $\lceil \log_2(N) \rceil$ bits for storage only. The choice $\lambda_0 = 0$ and $\lambda_1 = 1$ minimizes the hardware complexity of the (final) decision making step given by the decoding rule (2) as well since it depends on the bit width of the input. It's worth noting that the curves for JFER and BLFR given in Fig. 7, *a* and *b* respectively are very similar. That confirms that the probability of decoding failure is much higher than that of erroneous decoding for the α-detector. Moreover the fact that both JFER and BLFR curves exhibit minor change as K varies confirms our previous conclusion that the performance of the detector doesn't change significantly as the number of interfering users vary.

Conclusion

Hereinabove a communication system that uses frequency shift keying in a dynamically allocated hopset with α-detector is considered. This paper address the sensitivity of the communication system under consideration to both noise and interference power variation and detector parameters choice. Even though results obtained in [16] suggested that the communication system that makes use of the α-detector is resilient to background noise intensity variations this problem has never before been studied comprehensively. Hereinabove we have demonstrated that the communication system under consideration exhibits such feature in a vast range of SNR values for various values parameters and scenarios. Moreover the fact that the capacity of the equivalent channel exhibits threshold behaviour i.e. for any specific set of parameters and scenario there is a critical value of SNR such that for any SNR value less than the critical value the capacity of the equivalent channel starts to degrade as SNR reduces has been revealed for the first time. The analytical bound obtained in [17] demonstrated that the performance of the detector is not affected by the choice of λ_0 and λ_1. This paper shows that it is true for the communication system under consideration for all rates, different system parameters and scenarios. Thus the values of λ_0 and λ_1 can be chosen in the way that minimizes the hardware complexity. It has been demonstrated that the value of α that yields maximum capacity of the equivalent channel varies slowly with SNR. Thus parameters choice optimized for certain conditions yields performance close to optimal even if the parameters of the communication system change. To the best of the author's knowledge none of the results discussed above has been obtained before.

References

1. Jiang W., Han B., Habibi M. A., and Schotten H. D. The road towards 6G: A comprehensive survey. *IEEE Open Journal of the Communications Society*, 2021, vol. 2, pp. 334–366. doi:10.1109/OJCOMS.2021.3057679
2. Chen X., Ng D. W. K., Yu W., Larsson E. G., Al-Dhahir N., and Schober R. Massive access for 5G and beyond. *IEEE Journal on Selected Areas in Communications*, 2021, vol. 39, no. 3, pp. 615–637. doi:10.1109/JSAC.2020.3019724
3. Ahn Y., Kim W., and Shim B. Active user detection and channel estimation for massive machine-type communication: Deep learning approach. *IEEE Internet of Things Journal*, 2022, vol. 9, no. 14, pp. 11904–11917. doi:10.1109/JIOT.2021.3132329
4. Ahn J., Shim B., and Lee K. B. EP-based joint active user detection and channel estimation for massive machine-type communications. *IEEE Transactions on Communications*, 2019, vol. 67, no. 7, pp. 5178–5189. doi:10.1109/TCOMM.2019.2907853
5. Zhang S., and Chen M. Cluster formation and power control for M2M-enabled two-tier cellular network. *2018 IEEE 4th Intern. Conf. on Computer and Communications* (*ICCC*), IEEE, 2018, pp. 673–677. doi:10.1109/CompComm.2018.8781006
6. Han D., Minn H., Tefek U. and Lim T. J. Network dimensioning, QoE maximization, and power control for multi-tier machine-type communications. *IEEE Transactions on Communications*, 2019, vol. 67, no. 1, pp. 859–872. doi:10.1109/TCOMM.2018.2875735
7. Zhang Y., Xia W., Zhao H., Xu W., Wong K.-K., and Yang L. Cell-free IoT networks with SWIPT: Performance analysis and power control. *IEEE Internet of Things Journal*, 2022, vol. 9, no. 15, pp. 13780–13793. doi:10.1109/JIOT.2022.3143531
8. Yeduri S. R., Thummaluri U., Jeeru S., Kumar A., Dubey A., and Cenkeramaddi L. R. SIC-RSRA for massive machine-to-machine communications in 5G cellular IoT. *2022 IEEE Wireless Communications and Networking Conf.* (*WCNC*), IEEE, 2022, pp. 25–30. doi:10.1109/WCNC51071.2022.9771752
9. Qiao L., Zhang J., Gao Z., Ng D. W. K., Renzo M. D., and Alouini M.-S. Massive access in media modulation based massive machine-type communications. *IEEE Transactions on Wireless Communications*, 2022, vol. 21, no. 1, pp. 339–356. doi:10.1109/TWC.2021.3095484
10. Qiao L., Zhang J., Gao Z., Chen S., and Hanzo L. Compressive sensing based massive access for IoT relying on media modulation aided machine type communications. *IEEE Transactions on Vehicular Technology*, 2020, vol. 69, no. 9, pp. 10391–10396. doi:10.1109/TVT.2020.3006318
11. Viswanathan R., and Gupta S. Nonparametric receiver for FH-MFSK mobile radio. *IEEE Transactions on Communications*, 1985, vol. 33, no. 2, pp. 178–184. doi:10.1109/TCOM.1985.1096265

12. Gulliver T. A., and Felstead C. Nonparametric diversity combining for fast frequency hopping. *Conf. Proc. IEEE MILCOM*, IEEE, 1995, pp. 60–64. doi:10.1109/MILCOM.1995.483272

13. Annampedu V., Roganov V. V., and Viswanathan R. Two rank order tests for M-ary detection. *IEEE Transactions on Information Theory*, 2000, vol. 46, no. 2, pp. 585–594. doi:10.1109/18.825823

14. Kreshchuk A., and Potapov V. Better goodness-of-fit statistics for coded FSK decoding. *Electronic Notes in Discrete Mathematics*, 2017, vol. 57, pp. 139–145. doi.org/10.1016/j.endm.2017.02.024

15. Osipov D. Reduced-complexity robust detector in a DHA FH OFDMA system under mixed interference. *7th Intern. Workshop "Multiple Access Communications"*, *MCOM 2014*, August 27–28, 2014, Halmstad, Sweden. 2014, vol. 8715. doi.org/10.1007/978-3-319-10262-7_3

16. Osipov D. *On the channel capacity of an order statistics-based single-user reception in a multiple access system*. In: *Multiple Access Communications. MACOM 2015*. Lecture Notes in Computer Science, Springer, Cham, 2015. Vol. 9305. doi.org/10.1007/978-3-319-23440-3_8

17. Osipov D. S. An upper bound on error probability in communication systems with single-user reception based on order statistics. *Automation and Remote Control*, 2020, vol. 81, pp. 107–117. doi.org/10.1134/S0005117920010099

18. Pätzold M. *Mobile Radio Channels*. 2nd ed. Chichester, John Wiley Sons, 2011. 583 p.

19. Rappaport T. S. *Wireless Communications: Principles Practices*. 2nd ed. Upper Saddle River, N.J., Prentice Hall, 2002. 707 p.

20. Friis H. T. A note on a simple transmission formula. *Proc. of the IRE*, 1946, vol. 34, iss. 5, pp. 254–256. doi:10.1109/JRPROC.1946.234568

УДК 621.391
doi:10.31799/1684-8853-2023-1-61-70
EDN: PCFDIS

Обнаружение сигнала на фоне помех на основе порядковых статистик: анализ чувствительности и выбор параметров детектора

Д. С. Осипов[а,б], канд. техн. наук, старший научный сотрудник, orcid.org/0000-0003-0400-7181, d_osipov@iitp.ru
[а]Институт проблем передачи информации им. А. А. Харкевича РАН, Б. Каретный пер., 19, стр. 1, Москва, 127051, РФ
[б]Национальный исследовательский университет «Высшая школа экономики», Мясницкая ул., 20, Москва, 101000, РФ

Введение: разработка методов физического уровня для систем межмашинной связи является нетривиальной задачей. В частности проблематична разработка методов оценивания характеристик канала с низкой сложностью и методов контроля мощности с необходимой точностью. Один из возможных способов решения этой проблемы — использование методов приема, основанных на порядковых статистиках и не требующих каких-либо техник оценивания или контроля мощности. **Цель:** исследовать чувствительность системы связи, использующую частотно-позиционное кодирование в динамически выделяемом диапазоне, и приемник на основе порядковых статистик, предложенный автором, к изменениям мощности фонового шума, многопользовательских помех и к выбору параметров приемников. **Результаты:** анализ вычисленной с использованием моделирования значений пропускной способности эквивалентного канала подтверждает, что рассматриваемый приемник устойчив к флуктуациям отношения сигнал/шум и, следовательно, может использоваться даже в том случае, если энергия сигнала на приемном конце подвержена значительным изменениям. Показано, что величина числа значений, которым приписывается наибольшее значение оценки достоверности, оптимизированная для наибольшего числа активных пользователей, гарантирует значения пропускной способности, близкие к оптимальному, даже если число активных пользователей меняется. Также с помощью моделирования показано, что вероятностные характеристики не зависят от выбора значений оценок достоверности. Это говорит о том, что можно использовать значения, минимизирующие аппаратную сложность детектора. **Практическая значимость:** полученные результаты подтвердили то, что детектор рассматриваемого типа допускает практическую реализацию и использование в широком диапазоне сценариев. Аппаратная сложность может быть минимизирована без ущерба для качества связи.

Ключевые слова — межмашинная связь, псевдослучайно переключаемые частоты, динамически выделяемые поддиапазоны, некогерентный прием, α-детектор, пропускная способность.

Для цитирования: Osipov D. S. Signal detection amid noise using order statistics: detector sensitivity analysis and parameter choice. *Информационно-управляющие системы*, 2023, № 1, с. 61–70. doi:10.31799/1684-8853-2023-1-61-70, EDN: PCFDIS

For citation: Osipov D. S. Signal detection amid noise using order statistics: detector sensitivity analysis and parameter choice. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 61–70. doi:10.31799/1684-8853-2023-1-61-70, EDN: PCFDIS

АЛМАТАЕВ Тожибой Орзикулович

Профессор кафедры автомобилестроения Андижанского машиностроительного института, Узбекистан.
В 1980 году с отличием окончил Ташкентский автомобильно-дорожный институт по специальности «Эксплуатация и ремонт дорожных машин».
В 1987 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 150 научных публикаций и 13 свидетельств о государственной регистрации программ для ЭВМ и патентов на изобретения и полезные модели.
Область научных интересов — материаловедение, математическое моделирование, управление проектами.
Эл. адрес: talor58@mail.ru

БОЛСУНОВСКАЯ Марина Владимировна

Доцент Санкт-Петербургского политехнического университета Петра Великого.
В 1989 году окончила Ленинградский политехнический институт им. М. И. Калинина, в 1997 году — Санкт-Петербургский технический университет по специальности «Менеджмент».
В 2003 году защитила диссертацию на соискание ученой степени кандидата технических наук.
Является автором 178 научных публикаций и 19 свидетельств о государственной регистрации программ для ЭВМ и патентов на изобретения и полезные модели.
Область научных интересов — обработка изображений, трехмерное моделирование, проектирование информационных систем и баз данных, управление проектами.
Эл. адрес: bolsun_mv@spbstu.ru

БОРОВКОВ Алексей Иванович

Профессор, проректор по цифровой трансформации, руководитель Научного центра «Передовые цифровые технологии» Санкт-Петербургского политехнического университета Петра Великого.
В 1978 году окончил Ленинградский политехнический институт им. М. И. Калинина по специальности «Динамика и прочность машин».
В 1985 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 150 научных публикаций.
Область научных интересов — вычислительная механика и компьютерный инжиниринг, мульти- и трансдисциплинарные компьютерные технологии для решения промышленных задач и др.
Эл. адрес:
borovkov@compmechlab.com

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

БОРИСОВСКАЯ Анна Владимировна

Ассистент кафедры инфокоммуникационных технологий и систем связи Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2012 году окончила Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Комплексная защита объектов информатизации».
Является автором 11 научных публикаций.
Область научных интересов — системы связи, случайный множественный доступ.
Эл. адрес: borisovskaya@k36.org

БУРАКОВ Вадим Витальевич

Начальник отдела суперкомпьютерных технологий и разработки ПО Инжинирингового центра CompMechLab Санкт-Петербургского политехнического университета Петра Великого.
В 1996 году окончил Государственную академию аэрокосмического приборостроения по специальности «Вычислительные системы, комплексы и сети».
В 2011 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 60 научных публикаций.
Область научных интересов — программная инженерия, качество программ.
Эл. адрес:
burakov@compmechlab.com

ДМИТРИЕВА Лидия Алексеевна

Младший научный сотрудник лаборатории промышленных систем потоковой обработки данных Центра НТИ Санкт-Петербургского политехнического университета Петра Великого.
В 2010 году окончила Санкт-Петербургский государственный политехнический университет по специальности «Маркетинг».
Область научных интересов — управление проектами, бизнес-аналитика, информационно-технологическая поддержка бизнеса.
Эл. адрес: lagli@mail.ru

МОЛДОВЯН Николай Андреевич

Профессор, заведующий научно-исследовательским отделом проблем информационной безопасности Санкт-Петербургского института информатики и автоматизации РАН, заслуженный изобретатель РФ.
В 1975 году окончил Кишиневский политехнический институт по специальности «Полупроводниковые приборы».
В 2001 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 250 научных публикаций и 60 патентов на изобретения.
Область научных интересов — информационная безопасность, криптография, электронная цифровая подпись, блочные шифры.
Эл. адрес: nmold@mail.ru

РОСТОВА Ольга Владимировна

Доцент Высшей школы бизнес-инжиниринга Санкт-Петербургского политехнического университета Петра Великого.
В 2001 году окончила Санкт-Петербургский государственный технический университет по специальности «Экономика».
В 2008 году защитила диссертацию на соискание ученой степени кандидата экономических наук.
Является автором более 100 научных публикаций.
Область научных интересов — управление проектами внедрения информационных систем, оценка эффективности внедрения информационных систем, информационная безопасность.
Эл. адрес: O.2908@mail.ru

МОЛДОВЯН Александр Андреевич

Профессор, главный научный сотрудник лаборатории кибербезопасности и постквантовых криптосистем Санкт-Петербургского института информатики и автоматизации РАН.
В 1974 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Автоматизированные системы управления».
В 2005 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 200 научных публикаций и 60 патентов на изобретения.
Область научных интересов — компьютерная безопасность, защита информации, криптография, протоколы электронной цифровой подписи.
Эл. адрес: maa1305@yandex.ru

ОСИПОВ Дмитрий Сергеевич

Старший научный сотрудник лаборатории информационных технологий передачи, анализа и защиты данных Института проблем передачи информации им. А. А. Харкевича РАН, Москва.
В 2003 году окончил Московский государственный технический университет по специальности «Системы автоматического управления».
В 2008 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 35 научных публикаций.
Область научных интересов — теория передачи информации, разработка и исследование моделей систем множественного доступа, технологии защиты данных, передаваемых по беспроводным каналам связи, и др.
Эл. адрес: d_osipov@iitp.ru

СЕБЕРРИ Дженифер

Профессор, директор Центра компьютерных исследований безопасности Австралийского государственного университета Вуллонгонг (Wollongong), основатель школы криптографии Австралии, Вуллонгонг, Австралия.
В 1966 году получила степень бакалавра в университете Нового Южного Уэльса, в 1969 году — магистра естественных наук в университете Ла Троб, Австралия.
В 1971 году защитила диссертацию на соискание ученой степени доктора наук (PhD).
Является автором более 450 научных публикаций и шести монографий.
Область научных интересов — дискретная математика, комбинаторика, матрицы Адамара, безопасные криптоалгоритмы, передача информации.
Эл. адрес: jennie@uow.edu.au

СЕРГЕЕВ Михаил Борисович

Профессор, заведующий кафедрой вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения, почетный работник высшего профессионального образования РФ.
В 1980 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Электронные вычислительные машины».
В 2001 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 100 научных публикаций и 14 патентов на изобретения.
Область научных интересов — теория разрядных вычислений, методы проектирования спецпроцессоров для систем контроля и управления и др.
Эл. адрес: mbse@mail.ru

ШИРОКОВА Светлана Владимировна

Доцент Высшей школы бизнес-инжиниринга Санкт-Петербургского политехнического университета Петра Великого.
В 1993 году окончила Санкт-Петербургский государственный технический университет по специальности «Электронные вычислительные машины, комплексы, системы и сети».
В 1998 году защитила диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 150 научных публикаций.
Область научных интересов — управление проектами разработки и внедрения информационных систем, системный анализ в проектировании и управлении.
Эл. адрес: swchirokov@mail.ru

ТЮРЛИКОВ Андрей Михайлович

Профессор, заведующий кафедрой инфокоммуникационных технологий и систем связи Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1980 году окончил Ленинградский институт авиационного приборостроения по специальности «Информационные системы управления».
В 2011 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 150 научных публикаций.
Область научных интересов — многоабонентные системы связи, системы дистанционного обучения, протоколы передачи данных в реальном масштабе времени, алгоритмы сжатия видеоинформации.
Эл. адрес: turlikov@k36.org

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