ISSN 2686-679X

ВЕСТНИК РГГУ

Серия
«Информатика.
Информационная безопасность.
Математика»

Научный журнал

RSUH/RGGU BULLETIN

"Information Science.
Information Security. Mathematics"
Series

Academic Journal

Основан в 2018 г.
Founded in 2018

3
2022

VESTNIK RGGU. Seriya “Informatica. Informacionnaya bezopasnost. Matematica”

RSUH/RGGU BULLETIN. “Information Science. Information Security. Mathematics” Series
Academic Journal

There are 4 issues of the printed version of the journal a year.

Founder and Publisher
Russian State University for the Humanities (RSUH)

RSUH/RGGU BULLETIN. “Information Science. Information Security. Mathematics” series is included: in the Russian Science Citation Index; in the List of leading scientific magazines journals and other editions for publishing PhD research findings peer-reviewed publications fall within the following research area:

20.00.00 Informatics

81.93.29 Information security, data protection

27.00.00 Mathematics

Objectives and areas of research

RSUH/RGGU BULLETIN. “Information Science. Information Security. Mathematics” series publishes the results of research by scientists from RSUH and other universities and other Russian and foreign academic institutions. The areas covered by contributions include theoretical and applied computer science, up-to-date IT, means and technologies of information protection and information security as well as the issues of theoretical and applied mathematics including analytical and imitation models of different processes and objects. Special emphasis is put on articles and reviews covering research in indicated directions in the areas of social and humanitarian problems and also issues of personnel training for these directions.

RSUH/RGGU BULLETIN. “Information Science. Information Security. Mathematics” series is registered by Federal Service for Supervision of Communications Information Technology and Mass Media. 25.05.2018, reg. No. FS77-72977

Editorial staff office: 6, Miusskaya sq., Moscow, Russia, 125047

tel: +7 (916) 250-90-85

e-mail: grnat@rambler.ru

ВЕСТНИК РГГУ. Серия «Информатика. Информационная безопасность. Математика»
Научный журнал

Выходит 4 номера печатной версии журнала в год.

Учредитель и издатель – Российский государственный гуманитарный университет (РГГУ)

ВЕСТНИК РГГУ, серия «Информатика. Информационная безопасность. Математика», включен: в систему Российского индекса научного цитирования (РИНЦ); в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук по следующим научным специальностям и соответствующим им отраслям науки:

20.00.00 Информатика

81.93.29 Информационная безопасность, защита информации

27.00.00 Математика

Цели и область

В журнале «Вестник РГГУ», серия «Информатика. Информационная безопасность. Математика», публикуются результаты научных исследований ученых и специалистов РГГУ, а также других университетов и научных учреждений России и зарубежных стран. Направления публикаций включают теоретическую и прикладную информатику, современные информационные технологии, методы, средства и технологии защиты информации и обеспечения информационной безопасности, а также проблемы теоретической и прикладной математики, включая разработку аналитических и имитационных моделей процессов и объектов различной природы. Особое внимание уделяется статьям и обзорам, посвященным исследованиям по указанным направлениям в области социальных и гуманитарных проблем, а также вопросам подготовки кадров по соответствующим специальностям для данных направлений.

ВЕСТНИК РГГУ, серия «Информатика. Информационная безопасность. Математика», зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 25.05.2018 г., регистрационный номер ПИ № ФС77-72977.

Адрес редакции: 125047, Россия, Москва, Миусская пл., 6

Тел: +7 (916) 250-90-85

электронный адрес: grnat@rambler.ru

Учредитель и издатель
Российский государственный гуманитарный университет (РГГУ)

Ответственный за выпуск:

Н.В. Гришина, кандидат технических наук, доцент, Российский государственный гуманитарный университет (РГГУ)

CONTENTS

СОДЕРЖАНИЕ

УДК 004.45
DOI: 10.28995/2686-679X-2022-3-8-22

Обзор методологий перехода от требований к решениям в системной инженерии

Владимир А. Бочаров
Московский государственный технический университет им. Н.Э. Баумана, Москва, Россия, bocharovva@student.bmstu.ru

Тамара М. Волосатова
Московский государственный технический университет им. Н.Э. Баумана, Москва, Россия, tamaravol@gmail.com

Юрий А. Солоделов
Московский авиационный институт, Москва, Россия, yasolodelov@gmail.com

Михаил В. Филиппов
Московский государственный технический университет им. Н.Э. Баумана, Москва, Россия, filippov.mike@mail.ru

Аннотация. Создание правил переходов между этапами проектирования: цели потребителя системы, системные требования, технические решения и потребительские свойства систем остаются одним из самых важных аспектов разработки сложных систем. В данной статье рассматриваются решения руководства INCOSE по данной теме и обосновывается необходимость введения регулярности в связи между этапами проектирования для проверки и обоснованности каждого действия. Для упрощения переходов между этапами сравниваются между собой три методологии: унифицированный язык моделирования UML, онтологический язык OntoUML и методология моделирования IDEF0; объясняется, почему предлагается применить нотацию IDEF0 совместно с подходом Goal-Oriented Requirements Engineering (разработка требований, ориентированная на достижение цели), который позволяет компенсировать некоторые недостатки указанной нотации.

Ключевые слова: системный инжиниринг, INCOSE, функции, требования, цели, связи

Для цитирования: Бочаров В.А., Волосатова Т.М., Солоделов Ю.А., Филиппов М.В. Обзор методологий перехода от требований к решениям в системной инженерии // Вестник РГГУ. Серия «Информатика. Информационная безопасность. Математика». 2022. № 3. С. 8–22. DOI: 10.28995/2686-679X-2022-3-8-22

An overview of methodologies for moving from requirements to solutions in the system engineering

Vladimir A. Bocharov
Bauman Moscow State Technical University, Moscow, Russia, bocharovva@student.bmstu.ru

Tamara M. Volosatova
Bauman Moscow State Technical University, Moscow, Russia, tamaravol@gmail.com

Yurii A. Solodelov
Moscow Aviation Institute, Moscow, Russia, yasolodelov@gmail.com

Mikhail V. Filippov
Bauman Moscow State Technical University, Moscow, Russia, filippov.mike@mail.ru

Abstract. The creation of rules for transitions between design stages: the goals of the consumer of the system, system requirements, technical solutions and consumer properties of systems, remains one of the most important aspects of the development of complex systems. The article discusses the INCOSE decisions on the topic and justifies the need to introduce regularity in connection between the design stages in order to validate and validate each action. To simplify the transitions between the stages, three methodologies are compared with each other: the unified modeling language UML, the ontological language OntoUML and the modeling methodology IDEF0, the article also explains why it is suggested to apply the IDEF0 notation together with the Goal-Oriented Requirements Engineering approach, which allows compensating for some of the shortcomings of that notation

Keywords: system engineering, INCOSE, functions, requirements, goals, communications

For citation: Bocharov, V.A., Volosatova, T.M., Solodelov, Yu.A. and Filippov, M.V. (2022), "An overview of methodologies for moving from requirements to solutions in the system engineering", *RSUH/RGGU Bulletin. "Information Science. Information Security. Mathematics" Series*, no. 3, pp. 8–22, DOI: 10.28995/2686-679X-2022-3-8-22

Введение

В разработке сложной наукоемкой системы есть важные составляющие для ее внедрения: начиная от времени разработки до запуска системы в эксплуатацию и заканчивая обнаружением скрытых потребностей и последующим выполнением требований заказчиков. Перечисленные выше составляющие необходимо учитывать с самого первого этапа разработки, когда из технического задания заказчика необходимо сформировать требования, цели и структуру системы, которые непосредственно влияют на последующие этапы проектирования. Для выполнения данного этапа используется комплексная методология «системный инжиниринг» [Manenti, Ebrahimiarjestan, Yang, Yu 2019].

Системный инжиниринг (Systems Engineering, SE), управляющий всеми техническими и управленческими этапами, необходимыми для преобразования набора потребностей, ожиданий и ограничений заказчика в решение и поддержку этого решения на протяжении всей его жизни [ISO/IEC/IEEE 24765 2010], является важным элементом при разработке наукоемкой продукции. SE необходим, так как позволяет решить три проблемы инженерии: проблему сложности (недооценки сложности проекта), проблему непонимания (целей проекта, отношений внутри элементов системы, решения проблем, возникших в течение жизненного цикла проекта) и проблему коммуникации (между инженерами в команде, между организациями, внутри проекта), что было доказано во многих исследованиях (например, [Suranto 2015]). В связи с этим совершенствование SE способствует более эффективному процессу разработки сложной системы.

Инженерное проектирование в значительной степени зависит от модели разработки системы и процессов, методов и инструментов, которые поддерживают его. Предприятия ожидают, что инженерное проектирование системы будет происходить максимально легко и предсказуемо, а там, где это необходимо, будут внесены соответствующие изменения. ISO/IEC/IEEE 15288 [ISO/IEC/IEEE 15288 2015] определяет общие и стандартные процессы системного проектирования, которые подходят для

различных областей приложения. Международным советом по системной инженерии подготовлено «Руководство по системной инженерии INCOSE» (далее именуемое «Руководство INCOSE» [Walden, Roedler, Forsberg, Hamelin, Shortell 2015]), в котором подробно описывается применение стандарта и предоставляются рекомендации по адаптации и масштабированию данных процессов.

В настоящее время руководство INCOSE (текущее издание – версия 4, 2015 г.) во многих организациях становится основным справочником по созданию документов внутреннего процесса системного проектирования. Но его недостаток состоит в чрезвычайной сложности, являющейся препятствием для потенциальных пользователей (особенно новичков в системной инженерии) и не позволяющей быстро получить исчерпывающее представление о системной инженерии и ее приложениях. В частности, руководство INCOSE описывает процессы жизненного цикла линейно и последовательно [Walden, Roedler, Forsberg, Hamelin, Shortell 2015, p. 3–4], используя схемы Input-Process-Output (IPO) [Yang, Cormican, Yu 2017], но без общей картины того, как они связаны.

Но самая главная проблема SE и подобных ему современных методологий в том, что они работают в пределах четырех этапов проектирования систем: цели потребителя системы, системные требования, технические решения и потребительские свойства систем, но не решают проблему переходов между данными этапами, что приводит к неоптимальному результату, длительному проектированию системы и частым ошибкам, как, например, при разработке американского самолета F-35. Примером проблемы, которая возникает при отсутствии правильного перехода между системными требованиями и техническим решением, может служить проблема 25-мм пушки, установленной на F-35, так как ее точность не соответствует требованиям из-за плохого крепления, которое часто растрескивается и перекашивается. Возможной причиной возникновения данной проблемы (фактически невыполнения ТЗ) может служить несогласованность технического решения и требований. Поэтому целью данной работы является обзор возможных методологий создания сложных систем (в частности, авиационного бортового оборудования), которые позволили бы связать все этапы определенными правилами и привели бы к следующим результатам:

– сделает процесс разработки регулярным: подчинит его заранее подготовленному и утвержденному процессу, использующему проверенные методы;

– позволит проверить, что каждое действие в ходе процесса разработки обоснованно и не выходит за рамки установленных порядков, а его результат – не противоречит поставленным целям;
– позволит проверять выбор вариантов реализации из множества возможных, используя аналитические методы.

Для решения данной задачи предлагается сравнить три наиболее популярных и широко используемых методологии описания функциональных моделей между собой: UML, OntoUML и IDEF0 с целью определения наиболее простого, понятного и удобного для доработки подхода.

Унифицированный язык моделирования UML

Унифицированный язык моделирования (UML) – это графический язык для визуализации, определения, построения и документирования программно-ресурсоемких систем [Booch, Rumbaugh, Jacobson 1999]. UML предоставляет стандартный способ написания чертежей системы, охватывающих концептуальные аспекты, классы, написанные на определенном языке программирования, схемы баз данных и повторно используемые программные компоненты. UML является стандартным обозначением, которое используется всеми, кто участвует в производстве, развертывании и обслуживании программного обеспечения. UML включает девять диаграмм для описания системы.

Однако у UML есть недостаток, связанный с концепцией анализа объектов/классов. Согласно источнику [Pergl, Sales, Rybola 2013], UML только констатирует факты наличия тех или иных свойств у классов/атрибутов, но у него не предусмотрена возможность анализа на пересечение: использование одного и того же атрибута/свойства в других местах, что приводит к тому, что в разных классах один атрибут может иметь разное значение (а должен иметь одинаковое). Также при использовании UML программные системы описываются построением набора диаграмм. Обычно эти диаграммы создаются независимо и содержат перекрывающуюся информацию. Например, зависимости, содержащиеся в диаграммах классов, и сообщения в диаграммах последовательности сообщений связаны [Muskens, Bril, Chaudron 2005]; таким образом, эта взаимосвязь между обеими диаграммами вызывает требование согласованности. Без соответствующих средств для оценки сходства между диаграммами несоответствия могут возникнуть или даже

в худшем случае останутся незамеченными, когда они возникнут. Несогласованность между диаграммами UML увеличивает вероятность ошибок и потенциально неправильных значений сходства между проектами разработки программного обеспечения. Другой момент, вызывающий беспокойство, заключается в том, что UML не устанавливает отношения упорядочения между диаграммами, инженеры-программисты склонны рассматривать определенный тип диаграммы в большей степени по сравнению с диаграммой другого типа [Lange, Chaudron 2005]. В качестве примера инженеры считали диаграммы классов преобладающими по сравнению с другими диаграммами [Salami, Ahmed 2013]. Кроме того, данный язык моделирования предназначен для предоставления информации пользователю, а как ее использовать/анализировать – в языке не предусмотрено, что приводит к необходимости интеграции UML с другими языками.

Онтологический язык OntoUML

OntoUML – это хорошо обоснованный онтологический язык для концептуального моделирования. OntoUML построен как расширение UML на основе Unified Foundational Ontology (UFO) [Дерюгина 2015].

Его основная цель – описание семантических сетей и подтверждение или опровержение данных, которые были описаны в сети, при этом появление информации в сети никак не описывается. Поэтому у него есть ограничение: в данном языке можно создавать различные эталоны систем, но нет инструмента для сравнения эталона с полученной системой. Таким образом, чтобы понять, насколько спроектированная система отличается от эталона и по каким причинам, необходимо интегрировать OntoUML с другими методами, которые позволят сделать подобный анализ. Также в OntoUML способ дифференциации сущностей, которые существуют в языке, неполноценный, так как основывается на объектно-атрибутном подходе. Данная неполноценность возникает из-за того, что невозможно обобщенно описать множество различных объектов, так как при переходе к конкретному множеству атрибут становится объектом, следовательно, теряется связь между разными уровнями абстракций. Таким образом, необходимо использовать методологию, которая использует подход, основанный на связях между объектами.

Методология моделирования IDEF0

IDEF0 – это методология моделирования структурных представлений функций, участвующих в процессах или сложных системах [Brundage, Lechevalier, Morris 2018].

Модель IDEF0 состоит из серии иерархических и связанных диаграмм, которые позволяют представить исследуемый процесс используя графические диаграммы, текст и глоссарий. Данные типы представлений связаны друг с другом. Диаграммы IDEF0 основаны на пяти элементах: функция, входные данные, элементы управления, механизмы и выходы. Связываются данные элементы через функцию: функция получает входные данные и элементы управления, использует механизмы, а также предоставляет выходы. Аналогичным образом схемы IPO, определенные в справочнике INCOSE, также содержат пять ключевых элементов: процесс, входы, средства управления, выходы, факторы реализации.

Хотя диаграмма IPO напоминает диаграмму IDEF0, она не показывает декомпозицию разрабатываемой системы [Walden, Roedler, Forsberg, Hamelin, Shortell 2015, p. 192]. Декомпозиция – это когда каждая диаграмма представляет собой процесс или действие и может быть разбита на более мелкие и связанные действия. Данный подход также полезен для того, чтобы скрыть ненужную сложность из поля зрения, пока не потребуется более глубокое понимание. Таким образом, IDEF0 позволяет создавать многоуровневые модели для представления сложных систем с помощью простых, что облегчает этап анализа для человека, а именно этого требует системная инженерия. На диаграммах IDEF0 функция, которая является процессом или деятельностью, представлена рамкой с глагольной фразой внутри. Блоки функций можно дополнительно разложить с более высоких уровней на более низкие уровни, показывая детальную иерархию действий процесса. Остальные четыре компонента, а именно входы, элементы управления, механизмы и выходы, представлены стрелками. Все прямоугольники и стрелки связаны между собой. Функциональные блоки разбиты на набор подфункций. Эта декомпозиция обозначается номером функции на диаграммах, где 0 – функция верхнего уровня. В отличие от UML и OntoUML данная методология проще для использования, а также позволяет использовать сравнение различных функций, так как описание происходит в пределах одного типа диаграммы; именно поэтому предлагается использовать IDEF0 для последующей доработки, чтобы выполнить поставленные задачи.

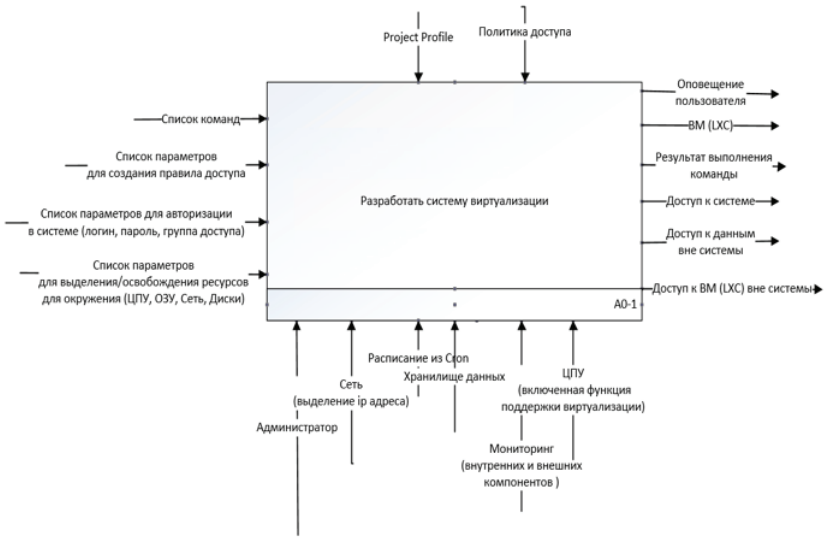

Рис. 1. Первый уровень IDEF0 «Системы Виртуализации»

Ограничения методологии IDEF0

Однако у данного метода, несмотря на его четкое представление системы в виде иерархической структуры, а также наличие прямых и обратных связей, есть ограничение: отсутствие связей между целями и функциями после проведения декомпозиции верхнего уровня, что не позволяет использовать этот метод на протяжении всего цикла проектирования. Далее данное ограничение будет продемонстрировано на примере использования метода IDEF0 при разработке «Системы Виртуализации». Первый уровень схемы показан на рис. 1.

Как видно из рис. 1, на первом уровне существует одна функция, которая имеет несколько входов, выходов, механизмов и управлений. Однако при проектировании системы инженеру трудно работать с функцией, которая имеет больше двух переменных. Для уменьшения количества переменных необходима декомпозиция системы, что позволяет разбить функцию на подфункции, которые обладают уже меньшим количеством параметров. Процесс декомпозиции происходит циклически до тех пор, пока у всех подфункций не будет минимального количества входов, то есть переменных. В идеале такое количество равно 1.

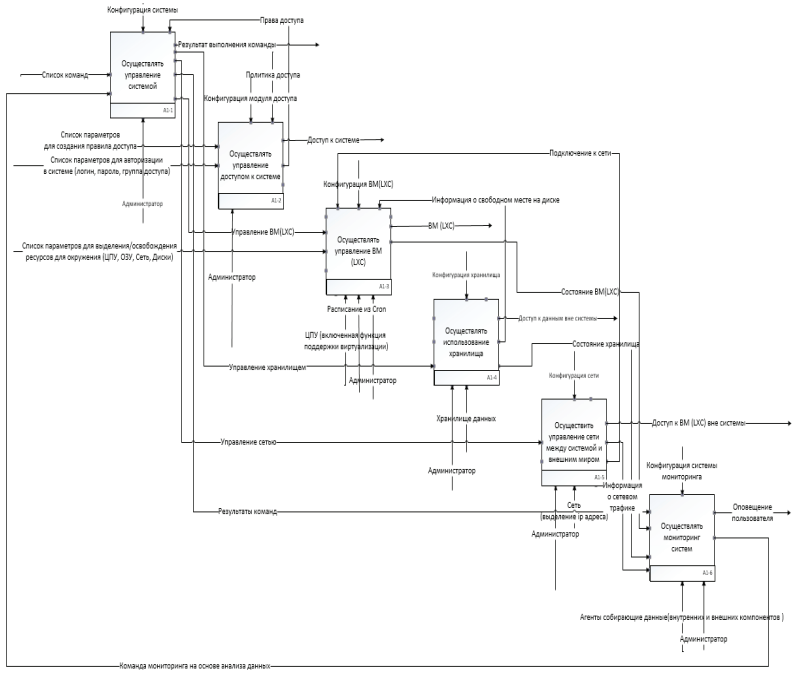

Рис. 2. Второй уровень IDEF0 «Системы Виртуализации»

На рис. 2 представлен второй уровень схемы «Системы Виртуализации».

В реальных проектах разработка системы основывается на целях, которые ставятся заказчиком перед исполнителем. Ограничение IDEF0 заключается в том, что цели связаны только с первым уровнем схемы, так как в названии функции подразумевается цель, которая стоит перед исполнителем. При декомпозиции и переходе на следующий уровень связь с целями разрывается, что приводит к тому, что описанная по методу IDEF0 система правильно функционирует, но может не соответствовать целям, которые были предъявлены заказчиком. Поэтому необходимо связать цели с функцией в IDEF0, которая описана на первом уровне, и также производить их декомпозицию для создания связей на следующих уровнях детализации функций системы. Для осуществления такой задачи необходимо, чтобы цели были описаны формально. Для этого подходят следующие методы: дерево целей, цели на основе показателей, Goal-Oriented Requirements Engineering [Lapouchnian 2005]

и др. Но инженер работает также с инженерными требованиями (Requirements Engineering, RE), поэтому представление целей в виде требований облегчает ему задачу, а для такого представления целей существует метод Goal-Oriented Requirements Engineering (Разработка требований, ориентированных на достижение цели – GORE).

Метод разработки требований на основе достижения цели

В начале 2000-х годов резко возросла популярность целенаправленных подходов к проектированию требований. Основной причиной этого является неадекватность традиционных подходов системного анализа (объектно-ориентированный подход, структурный анализ и т. д.) ([Ross 1977; Rumbaugh 2003]) при работе со все более сложными системами. На уровне требований эти подходы рассматривают требования как состоящие только из процессов и данных и не учитывают обоснование систем, что затрудняет понимание требований в отношении некоторых проблем высокого уровня в проблемной области. Большинство методов фокусируются только на моделировании и спецификации системы. Поэтому им не хватает поддержки для рассуждений о сложной системе, состоящей из будущей системы и ее окружения. Однако известно, что неверные предположения об окружающей среде системы ответственны за многие ошибки в спецификациях требований [Lamsweerde, Letier 2004]. Нефункциональные требования также, как правило, остаются за пределами спецификаций требований. Кроме того, традиционные методы моделирования и анализа не позволяют представлять и сравнивать альтернативные конфигурации системы, в которых более или менее функционально автоматизированы или исследуются различные распределения ответственности и т. д. GORE пытается решить вышеуказанные важные проблемы. Важно отметить, что процесс разработки целевых требований заканчивается там, где начинается большинство традиционных методов спецификации [Lamsweerde, Letier 2004]. В целом GORE фокусируется на деятельности, которая предшествует формулированию требований к системе. В подходах GORE обычно присутствуют следующие основные виды деятельности: выявление целей, уточнение целей и различные виды анализа целей, а также распределение ответственности за цели между агентами.

Большинство ранних исследований в области RE были сосредоточены на том, что должна делать система и как она должна это

делать. Это сводилось к выработке и уточнению довольно низкоуровневых требований к данным, операциям и т. д. Несмотря на то что необходимость в обосновании разрабатываемой системы была очевидна из ранних определений требований (например, «определение требований должно указывать, зачем нужна система» [Ross, Schoman 1977]), в литературе по RE мало внимания уделялось пониманию того, зачем нужна система и действительно ли спецификация требований отражает потребности заинтересованных сторон. Недостаточно внимания уделялось пониманию организационного контекста новой системы. В целом тенденция в исследованиях моделирования состояла в том, чтобы абстрагировать низкоуровневые программные конструкции до уровня требований, а не опускать абстракции требований вниз до уровня проектирования.

Заключение

Характерной чертой современных подходов к разработке сложных наукоемких систем является то, что каждый подход позволяет использовать только часть данных, необходимых для разработки сложных систем, и не учитывает данные, которые были получены другими методами. Иными словам, методы не связаны между собой отношением, которое бы позволило анализировать влияние одних данных на другие. В связи с этим была поставлена задача поиска решения данной проблемы и рассмотрены различные варианты.

В данной статье были исследованы нотации UML, OntoUML и IDEF0 с целью выбора одной из них для последующей доработки, которая позволила бы решить задачу перехода от требований к решениям в системной инженерии. В ходе анализа был выбран IDEF0 и продемонстрированы его недостатки:

1) IDEF0 не позволяет разделять требования и реализацию, а также не обеспечивает их связность;
2) В IDEF0 не предусмотрены преобразования требований высокого уровня в требования низкого уровня;
3) механизмы, с помощью которых получаются выходные данные, только упоминаются, но связи предикатов с входами и выходами не раскрываются;
4) IDEF0 не отображает поведение системы в разных режимах.

Для устранения данных недостатков необходимо менять нотацию, что является сложной задачей; альтернативой может являться использование различных методик на разных уровнях и создание связей между ними для обеспечения работы с общей информацией на всех уровнях разработки сложной наукоемкой системы.

В дальнейшем планируется объединение IDEF0 с другими методиками для устранения его недостатков, а после перейти к реализации полученной методики через объектно-ориентированный язык, который позволит описывать систему и представлять описание графически.

Литература

Дерюгина 2015 – *Дерюгина О.А.* Семантика и семантически эквивалентные трансформации UML-диаграмм классов // Труды МФТИ. 2015. Т. 7. № 2. С. 146–155.

Booch, Rumbaugh, Jacobson 1999 – *Booch G., Rumbaugh J., Jacobson I.* The Unified Modeling Language Use Guide // J. Database Manag. 1999. Vol. 10. P. 51–52.

Brundage, Lechevalier, Morris 2018 – *Brundage M.P., Lechevalier D., Morris K.C.* Toward Standards-Based Generation of Reusable Life Cycle Inventory Data Models for Manufacturing Processes // J. Manuf. Sci. Eng. 2018. Vol. 141. Issue 2.

ISO/IEC/IEEE 15288 2015 – ISO/IEC/IEEE 15288 ISO/IEC/IEEE International Standard – Systems and software engineering – System life cycle processes. Geneva; New York: International Organization for Standardization; IEEE, 2015.

ISO/IEC/IEEE 24765 2010 – ISO/IEC/IEEE 24765 ISO/IEC/IEEE International Standard – Systems and software engineering – Vocabulary. Geneva; New York: International Organization for Standardization; IEEE, 2010.

Lamsweerde, Letier 2004 – *Lamsweerde A., Letier E.* From Object Orientation to Goal Orientation: A Paradigm Shift for Requirements Engineering // Radical Innovations of Software and Systems Engineering in the Future: Proc. 9th Intern. Workshop, RISSEF 2002, Venice, Italy, October 2002 / M. Wirsing, A. Knapp, S. Balsamo (eds.). Berlin; Heidelberg: Springer, 2004. P. 156–166.

Lange, Chaudron 2005 – *Lange C.F., Chaudron M.R.* Experimentally investigating effects of defects in UML models. Eindhoven: Technische Universiteit Eindhoven, 2005 (Computer science reports, Vol. 0507).

Lapouchnian 2005 – *Lapouchnian A.* Goal-Oriented Requirements Engineering: An Overview of the Current Research. Toronto: University of Toronto, 2005.

Manenti, Ebrahimiarjestan, Yang, Yu 2019 – *Manenti G., Ebrahimiarjestan M., Yang, Lan, Yu, Ming.* Functional Modelling and IDEF0 to Enhance and Support Process Tailoring in Systems Engineering // International Symposium on Systems Engineering (ISSE) 2019. New York: IEEE, 2019. P. 1–8.

Muskens, Bril, Chaudron 2005 – *Muskens J., Bril R.J., Chaudron M.R.V.* Generalizing consistency checking between software views // WICSA' 05: Proceedings of the 5th Working IEEE/IFIP Conference on Software Architecture, Pittsburgh, PA, November 6–9. New York: IEEE, 2005. P. 169–180.

Pergl, Sales, Rybola 2013 – *Robert P, Sales T.P., Rybola Z.* Towards OntoUML for Software Engineering: From Domain Ontology to Implementation Model // Third

International Conference, MEDI 2013, Amantea, Italy, September 25–27, 2013. Berlin; Heidelberg: Springer, 2013.

Ross 1977 – *Ross D.* Structured Analysis (SA): A Language for Communicating Ideas // IEEE Transactions on Software Engineering. 1977. Vol. SE-3, issue 1.

Ross, Schoman 1977 – *Ross D., Schoman K.* Structured Analysis for Requirements Definition // IEEE Transactions on Software Engineering. 1977. Vol. SE-3, issue 1.

Rumbaugh 2003 – *Rumbaugh J.* Object-oriented analysis and design (OOAD) // Encyclopedia of Computer Science. London: John Wiley & Sons Ltd., 2003. P. 1275–1279.

Salami, Ahmed 2013 – *Salami H.O., Ahmed M.* Class Diagram Retrieval Using Genetic Algorithm // Machine Learning and Applications (ICMLA), 2013 12th International Conference. New York: IEEE, 2013. P. 96–101.

Suranto 2015 – *Suranto B.* Systems Enhineering: why is it important? // The 4th ICIBA 2015, International Conference on Information Technology and Business Applications, Palembang, Indonesia. Palembang: Bina Darma University, 2015. P. 20–21.

Walden, Roedler, Forsberg, Hamelin, Shortell 2015 – *Walden D.D., Roedler G.J., Forsberg K.J., Hamelin R.D., Shortell T.M.* Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities, 4th ed. Hoboken, NJ: John Wiley & Sons, Inc., 2015. 304 p.

Yang, Cormican, Yu 2017 – *Yang L., Cormican K., Yu.M.* Towards a methodology for systems engineering ontology development // 2017 IEEE International Systems Engineering Symposium (ISSE). New York: IEEE, 2017. P. 1–7.

References

Booch, G., Rumbaugh, J., Jacobson, I. (1999), "The Unified Modeling Language Use Guide", *J. Database Manag.*, vol. 10, pp. 51–52.

Brundage, M.P., Lechevalier, D. and Morris, K.C. (2018), "Toward Standards-Based Generation of Reusable Life Cycle Inventory Data Models for Manufacturing Processes", *J. Manuf. Sci. Eng.*, vol. 141, issue 2.

Deryugina, O.A. (2015), "Semantics and semantically equivalent transformations of the UML class diagrams", *Proceedings of Moscow Institute of Physics and Technology*, vol. 7, no. 2, pp. 146–155.

ISO/IEC/IEEE 15288 (2015), ISO/IEC/IEEE 15288 ISO/IEC/IEEE International Standard – Systems and software engineering – System life cycle processes, International Organization for Standardization; IEEE, Geneva; New York, Switzerland; USA.

ISO/IEC/IEEE 24765 (2010), ISO/IEC/IEEE 24765 ISO/IEC/IEEE International Standard – Systems and software engineering – Vocabulary, International Organization for Standardization; IEEE, Geneva; New York, Switzerland; USA.

Lamsweerde, A. and Letier, E. (2004), "From Object Orientation to Goal Orientation: A Paradigm Shift for Requirements Engineering", in Wirsing, M., Knapp, A. and Balsamo, S. (eds.), *Radical Innovations of Software and Systems Engineering in the Future: Proc. 9th Intern. Workshop, RISSEF 2002, Venice, Italy, October 2002*, Springer, Berlin, Heidelberg, Germany, pp. 156–166.

Lange, C.F. and Chaudron, M.R. (2005), Experimentally investigating effects of defects in UML models, Technische Universiteit Eindhoven, Eindhoven, The Netherlands.

Lapouchnian, A. (2005), Goal-Oriented Requirements Engineering: An Overview of the Current Research, University of Toronto, Toronto, Canada.

Manenti, G., Ebrahimiarjestan, M., Yang, Lan and Yu, Ming (2019), "Functional Modelling and IDEF0 to Enhance and Support Process Tailoring in Systems Engineering", *International Symposium on Systems Engineering (ISSE) 2019*, IEEE, New York, USA, pp. 1–8.

Muskens, J., Bril, R.J. and Chaudron, M.R.V. (2005), "Generalizing consistency checking between software views", *WICSA' 05: Proceedings of the 5th Working IEEE/IFIP Conference on Software Architecture*, Pittsburgh, PA, November 6–9, IEEE, New York, USA, pp. 169–180.

Robert, P, Sales, T.P. and Rybola, Z. (2013), "Towards OntoUML for Software Engineering: From Domain Ontology to Implementation Model", *Third International Conference, MEDI 2013, Amantea, Italy, September 25–27, 2013*, Springer, Berlin, Heidelberg, German.

Ross, D. (1977), "Structured Analysis (SA): A Language for Communicating Ideas", *IEEE Transactions on Software Engineering*, vol. SE-3, issue 1.

Ross, D. and Schoman, K. (1977), "Structured Analysis for Requirements Definition", *IEEE Transactions on Software Engineering*, vol. SE-3, issue 1.

Rumbaugh, J. (2003), "Object-oriented analysis and design (OOAD)", *Encyclopedia of Computer Science*, John Wiley and Sons Ltd., London, UK, pp. 1275–1279.

Salami, H.O. and Ahmed, M. (2013), "Class Diagram Retrieval Using Genetic Algorithm", *Machine Learning and Applications (ICMLA), 2013 12th International Conference*, IEEE, New York, USA, pp. 96–101.

Suranto, B. (2015), "Systems Enhineering: why is it important?", *The 4th ICIBA 2015, International Conference on Information Technology and Business Applications, Palembang, Indonesia*, Bina Darma University, Palembang, Indonesia, pp. 20–21.

Walden, D.D., Roedler, G.J., Forsberg, K.J., Hamelin, R.D., and Shortell, T.M. (2015), *Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities, 4th ed.*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 304 p.

Yang, L., Cormican, K. and Yu, M. (2017), "Towards a methodology for systems engineering ontology development", *2017 IEEE International Systems Engineering Symposium (ISSE)*, IEEE, New York, pp. 1–7.

Информация об авторах

Владимир А. Бочаров, аспирант, Московский государственный технический университет имени Н.Э. Баумана, Москва, Россия; 105005, Россия, Москва, 2-я Бауманская ул., д. 5; bocharovva@student.bmstu.ru

Тамара М. Волосатова, кандидат технических наук, доцент, Московский государственный технический университет имени Н.Э. Баумана, Москва, Россия; 105005, Россия, Москва, 2-я Бауманская ул., д. 5; tamaravol@gmail.com

Юрий А. Солоделов, Московский авиационный институт, Москва, Россия; 125080, Россия, Москва, Волоколамское шоссе, д. 4; yasolodelov@gmail.com

Михаил В. Филиппов, кандидат технических наук, доцент, Московский государственный технический университет имени Н.Э. Баумана, Москва, Россия; 105005, Россия, Москва, 2-я Бауманская ул., д. 5; filippov.mike@mail.ru

Information about the authors

Vladimir A. Bocharov, postgraduate student, Bauman Moscow State Technical University, Moscow, Russia; bld. 5, 2nd Baumanskaya Str., Moscow, Russia, 105005; bocharovva@student.bmstu.ru

Tamara M. Volosatova., Cand. of Sci. (Computer Science), associate professor, Bauman Moscow State Technical University, Moscow, Russia; bld. 5, 2nd Baumanskaya Str., Moscow, Russia, 105005; tamaravol@gmail.com

Yurii A. Solodelov, Moscow Aviation Institute, Moscow, Russia; bld. 4, Volokolamskoe highway, Moscow, Russia, 125080; yasolodelov@gmail.com

Mikhail V. Filippov, Cand. of Sci. (Computer Engineering), associate professor, Bauman Moscow State Technical University, Moscow, Russia; bld. 5, 2nd Baumanskaya Str., Moscow, Russia, 105005; filippov.mike@mail.ru

УДК 004.3
DOI: 10.28995/2686-679X-2022-3-23-42

Физические ограничения прогресса вычислительной техники

Андрей М. Подорожный
Российский государственный гуманитарный университет, Москва, Россия, poam@mail.ru

Аннотация. В работе проанализирована существующая элементная база компьютерной техники. Показано, что физические основы элементной базы за последние 70 лет не изменились: абсолютное большинство вычислительной техники создается на основе кремниевых полупроводниковых транзисторов. Произошло совершенствование технологий, уменьшение транзисторов до наноразмеров, что привело к росту вычислительной мощности на много порядков. Дальнейшее совершенствование кремниевой полупроводниковой технологии вступает в противоречие с фундаментальными физическими законами: необходимостью отвода тепла, невозможностью преодоления скорости света, неустранимых утечек тока за счет туннельного эффекта, вызванного квантовомеханическим соотношением неопределенностей. Несмотря на маркетинговые заявления, пока никому не удалось создать транзисторы меньше 25 нм. Закон Мура перестал работать, развитие элементной базы замедлилось. В связи с этим остро стоит проблема создания принципиально новой элементной базы компьютеров. Наиболее перспективными направлениями являются создание транзисторов на основе графена, а также квантовых компьютеров. Однако эти направления пока не могут служить полноценной заменой существующим технологиям.

Ключевые слова: закон Мура, кремниевые транзисторы, нанотехнологии, отвод тепла, скорость света, соотношение неопределенностей, фундаментальные ограничения развития

Для цитирования: *Подорожный А.М.* Физические ограничения прогресса вычислительной техники // Вестник РГГУ. Серия «Информатика. Информационная безопасность. Математика». 2022. № 3. С. 23–42. DOI: 10.28995/2686-679X-2022-3-23-42

Physical limitations of computer technology progress

Andrey M. Podorozhnyi
Russian State University for the Humanities, Moscow, Russia,
poam@mail.ru

Abstract. The paper analyzes the existing element base of computer technology. It is shown that the physical foundations of the element base have not changed over the past 70 years: the absolute majority of computing equipment is created on the basis of silicon semiconductor transistors. There was an improvement in technology, a reduction of transistors to nanoscale, which led to an increase in computing power by many orders of magnitude. Further improvement of the silicon semiconductor technology comes into conflict with fundamental physical laws: the need for heat removal, the impossibility of overcoming the speed of light, ineliminable current leakage due to the tunneling effect caused by the quantum mechanical uncertainty ratio. Despite marketing statements, so far no one has managed to create transistors smaller than 25 nm. Moore's law stopped working, the development of the element base slowed down. In this regard, the issue of creating a fundamentally new element base of computers is acute. The most promising areas are the creation of graphene-based transistors, as well as quantum computers. However, those areas cannot yet serve as a full-fledged replacement for existing technologies.

Keywords: Moore's law, silicon transistors, nanotechnology, heat removal, speed of light, uncertainty ratio, fundamental limitations of development

For citation: Podorozhnyi, A.M. (2022), "Physical limitations of computer technology progress", *RSUH/RGGU Bulletin. "Information Science. Information Security. Mathematics" Series,* no. 3, pp. 23–42, DOI: 10.28995/2686-679X-2022-3-23-42

За последние 70–80 лет электронно-вычислительная техника достигла фантастических успехов. Так, первая ламповая ЭВМ ЭНИАК, созданная в 1946 г., выполняла 5000 операций сложения и 360 операций умножения в секунду. А сегодня самый мощный суперкомпьютер (на ноябрь 2021 г.) имеет производительность 442 тыс. TFlops/s, то есть $4{,}42 \times 10^{17}$ операций над числами с плавающей точкой в секунду[1]. Ни в одной другой области техники не наблюдалось таких темпов развития. И это развитие преобразовало все сферы деятельности человеческого общества.

[1] Рейтинг и описания 500 самых мощных общественно известных вычислительных систем мира // TOP500.org. URL: https://www.top500.org/lists/top500/2021/11/ (дата обращения 5 июля 2022).

До недавнего времени развитие компьютерной техники определялось известным законом Мура. В 1965 г. Гордон Мур, один из руководителей Intel, сделал эмпирическое наблюдение об экспоненциальном росте мощности вычислительных устройств: каждый год число транзисторов в микросхемах удваивается. В 1975 г. он внес поправку: удвоение числа транзисторов на кристалле происходит каждые два года. Позже Дэвид Хаус, коллега Мура из Intel, тоже внес коррективу: производительность процессоров должн удваиваться каждые 18 месяцев за счет сочетания роста количества транзисторов и увеличения тактовой частоты. Это значит, что за 9 лет производительность процессоров вырастает в $2^6 = 64$ раза, а с 1975 по 2005 г. она должна была вырасти в $2^{20} = 1\,048\,576$ раз; примерно так и произошло.

К сожалению, в последние десять лет после уникального, впечатляющего предыдущего развития производительность процессоров растет сравнительно медленно. Это известный факт, он отчетливо прослеживается в фундаментальном описании характеристик всех выпущенных за последние 10–15 лет процессоров AMD [Дудкин 2021a] и Intel [Дудкин 2021b]. Так, процессоры AMD в 2009 г. выпускались по технологии 45 нм, и их тактовая частота находилась в пределах от 1,6 до 3,6 Ггц. А выпущенная в 2019 г. последняя серия производилась по технологии 7 нм, но тактовая частота каждого процессора выросла не так уж сильно: от 2,9 до 4 Ггц.

Разумеется, выросло число ядер: тогда было максимум 6, сейчас до 64, а также вырос размер кэш-памяти. По-прежнему идет увеличение числа транзисторов в процессоре как целом. Но примерно до 2005 г. это увеличение шло в одном процессорном ядре, что вело к пропорциональному росту тактовой частоты. А сегодня в отдельном ядре число транзисторов слабо увеличивается, несмотря на снижение их размеров в несколько раз.

Но зато растет число процессорных ядер, в которых осуществляются независимые параллельные вычисления. Однако полной параллельности достичь невозможно, в любом алгоритме присутствуют последовательные стадии. И, согласно закону Амдала[2], существуют границы эффективности выполнения задачи за счет распараллеливания.

В 2007 г. сам Мур заявил, что его закон не вечен, когда-нибудь он перестанет действовать из-за атомарных ограничений и влияния скорости света. С тех пор проблема замедления прогресса

[2] Закон Амдала // Википедия. URL: https://ru.wikipedia.org/wiki/Закон_Амдала (дата обращения 5 июля 2022).

в наноэлектронике становится все острее, постоянно предлагаются различные решения. Данная работа также посвящена анализу этой проблемы.

Этапы развития вычислительной техники

Принято считать, что история развития ЭВМ начинается с 30-х годов XX в., когда появилась реально работающая вычислительная техника, соединившая математическую логику с двоичной системой счисления. Элементной базой этих устройств были электромагнитные реле. Затем появились ЭВМ первого поколения на вакуумных лампах, где тормозящая процесс механика была устранена. Их сменили ЭВМ второго поколения, выполненные на транзисторах, на смену второму пришло третье поколение ЭВМ с элементной базой на микрочипах. Каждая смена поколений ЭВМ радикально улучшала характеристики компьютеров[3]. Но и внутри поколений непрерывно шло совершенствование технических параметров.

Четвертое поколение оказалось самым длительным: его определяют с конца 70-х годов прошлого века по настоящее время [Казакова 2011]. В ЭВМ 4-го поколения все вычисления по программам сосредоточены в одной микросхеме: центральном процессоре[4]. То есть здесь не было смены элементной базы, произошли изменения в архитектуре.

Но и при переходе от 2-го к 3-му поколению ЭВМ фактически тоже не произошло смены элементной базы. Разберемся подробнее.

В электромагнитном реле за счет протекающего в катушке электрического тока находящийся внутри сердечник движется и замыкает контакты. В электровакуумном триоде между катодом и анодом протекает электрический ток. Третий электрод, сетка за счет подачи на нее напряжения может ослабить этот ток и может его прекратить. Последнее аналогично размыканию контактов реле. В электровакуумных лампах с помощью небольшого сигнала на сетке можно управлять значительными величинами тока. Оче-

[3] Термины «ЭВМ» и «компьютер» можно считать синонимами, отличия только по времени употребления. Постепенно «компьютер» вытеснил «ЭВМ», в силу чего характеристики и возможности современных компьютеров и ЭВМ прошлого века несравнимы. Однако в юридической и иной официальной литературе по-прежнему пишется «ЭВМ», и все написанное там полностью относится к компьютерам.

[4] Позднее появился еще видеопроцессор.

видно, что в реле и вакуумных лампах одна и та же задача решается с помощью совершенно разных физических процессов.

В 1948 г. был создан первый транзистор, а первый компьютер на транзисторах был создан в 1953 г. Транзистор состоит из полупроводников, в которые добавлены примеси, создающие избыток свободных электронов (полупроводники n-типа) или недостаток электронов (полупроводники p-типа). Здесь аналогично электронной лампе слабым сигналом, подаваемым на базу (затвор), можно управлять большим током, идущим от эмиттера (исток) к коллектору (сток), вплоть до полного выключения тока. Существует несколько схем сочетания pn-полупроводников, но суть всегда одна: один сигнал управляет другим. Микросхемы компьютера, в том числе в процессоре, построены по схеме pn-переходов.

То есть при замене ламп на транзисторы тоже была успешно решена задача управления на другой, более удачной физической основе. С меньшими размерами, большей скоростью вычислений, меньшим потреблением энергии и другими свойствами (рис. 1).

Рис. 1. Элементная база поколений ЭВМ

В следующем, третьем поколении ЭВМ (с начала 60-х годов) элементная база стала выполняться на интегральных схемах. Здесь на одном кристалле стали формировать сначала десятки, а потом тысячи, миллионы, миллиарды транзисторов. И не только транзисторов, но также и других деталей: диодов, конденсаторов, сопротивлений, необходимых в тех или иных схемах (рис. 1).

На современном этапе технология производства микросхем чрезвычайно сложна и многостадийна. Сложно кратко описать эти технологические процессы. А детально они описаны во многих источниках, например в [Родионов 2019]. Осуществляется работа с особо чистыми веществами в стерильной среде, используют наношаблоны для литографии, эпитаксиальные процессы выращивания полупроводниковых пленок, создаются сверхмалые изолирующие и проводящие структуры и прочее. Не менее сложна задача автоматизированного проектирования структуры процессора: схемотехнического, электрического, физического и др.

Детали современных микропроцессоров достигли наноразмеров. Технологии их создания настолько сложны, что становятся практически невоспроизводимыми. Некоторые операции способно делать только одно предприятие в мире, и оно снабжает своей продукцией всех потребителей.

В конечном счете на одном кристалле получается схема из полупроводниковых транзисторов и других деталей, электрически изолированных друг от друга и соединенных металлизированными связями. Но физически в микросхеме происходят точно такие же процессы, как и в схеме на основе транзисторов, диодов, конденсаторов, вручную спаянных между собой; различие в масштабах.

Кристалл микропроцессора с миллиардами транзисторов по сравнению со спаянной вручную схемой имеет на много порядков большую вычислительную мощность, а также качественно другие, несравнимые возможности обработки данных. А также намного меньшие размеры и энергопотребление вычислительных устройств, вплоть до карманных смартфонов.

Таким образом, с 1953 г. фундаментальных, качественных изменений в элементной базе компьютеров не произошло. Это все те же твердотельные полупроводниковые электрические схемы, выполненные на транзисторах.

Причем и тогда, и сейчас основным «строительным материалом», из которого создаются микросхемы, служит кремний. А это второй по распространенности химический элемент в земной коре после кислорода. Камень, песок, земля – это в значительной степени кремний. Он дешев и запасы его неограниченны, этот экономический фактор тоже принимается во внимание. Хотя, конечно, в микропроцессорной технике в качестве добавок используются дорогие и редкие химические элементы, которых в земной коре мало.

Факторы, ограничивающие производительность микропроцессоров

Таким образом, наблюдается явное замедление развития микропроцессорной техники. Оно объясняется возникновением нескольких неустранимых факторов; при этом их устранение будет противоречить фундаментальным физическим законам. Рассмотрим три из этих факторов.

1. Выделение тепла. Это первый, а потому самый важный из факторов, с которым столкнулись производители микропроцессорной техники; понятно, что выделение тепла неизбежно, согласно первому началу термодинамики.

Согласно паспортным данным, рабочая температура микропроцессоров компьютеров не должна превышать 60–85°C. За пределами рабочей температуры могут возникнуть разнообразные токи утечки, приводящие к ошибкам в работе программ, к зависанию компьютера, а также к необратимым изменениям. На практике разрушения процессора не происходит, потому что в современных компьютерах установлены термодатчики, и при превышении допустимой температуры процессор автоматически замедляет работу, или выключается.

На рис. 2 показана температурная зависимость электропроводности кремния, германия и лучшего из доступных проводников – меди. Рисунок взят из работы [Tiberius 2019], дополнительно проставлены значения проводимости кремния при комнатной температуре, при средней предельно допустимой температуре процессора и при температуре, когда проводимость за счет собственных электронов кремния становится равной проводимости за счет внесенных примесных атомов, то есть транзистор теряет управляемость [Tiberius 2019].

Как видно из рисунка, регулируемая температура намного ниже той, при которой могут произойти необратимые изменения в кремниевом транзисторе. Но на микроучастках возможны значительные местные перегревы, вплоть до необратимых изменений, а датчик измеряет среднюю температуру материала; поэтому предел рабочей температуры занижается. Очевидно, производители процессоров эмпирически установили допустимые температурные пределы.

Практически все процессоры имеют одинаковый размер 37,5 × 37,5 мм, плюс/минус несколько миллиметров, толщина же кремниевой пластины измеряется в нанометрах [Миллс 2020]. Плоская форма препятствует нагреву. Процессоры имеют более тысячи контактов, через которые тоже отводится тепло. Активно применяются дополнительные средства охлаждения: чипы покры-

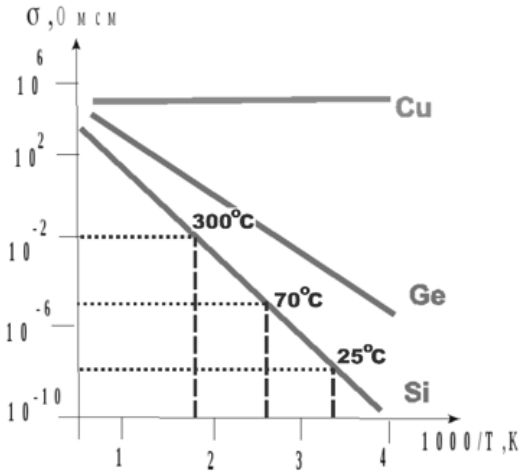

Рис. 2. Температурная зависимость проводимости кремния, германия, меди

вают теплопроводными пастами, используют металлические радиаторы, создающие поток воздуха кулеры, применяют водяное охлаждение, вплоть до помещения в бассейн или озеро.

Но физику не обманешь, и существуют температурные ограничения на размеры компактных мощных вычислительных устройств. А разнесение устройств в пространстве, как будет показано далее, может снизить скорость обработки данных.

2. Скорость света. Скорость света составляет 300 000 км/сек, точнее 299 792 км/сек. Как известно, это предельная скорость распространения частиц и сигналов; ничто не может двигаться с большей скоростью. Электромагнитные волны в вакууме распространяются со скоростью света, при этом обычно считают, что со скоростью света распространяется и электрический сигнал[5].

Однако на самом деле в твердом материале скорость распространения электромагнитных волн снижается. В оптических материалах снижение пропорционально показателю преломления. В книге [Kaiser 2005] приведены данные по скорости распространения электрического сигнала; для различных материалов она составляет от 42 до 99% от скорости света в вакууме.

[5] Конечно, сами электроны и другие носители заряда с такой скоростью не движутся, скорость их перемещения имеет порядок миллиметров в секунду.

Одной из стадий обмена данными является передача сигнала. В случае электромагнитных волн это время передачи сигнала со скоростью света. В земных условиях это время составляет максимум доли секунды, и человек его не замечает. Однако при диалоге с устройствами, запущенными в космос, задержка может составлять минуты и даже часы. Здесь стадия передачи сигнала становится определяющей в информационном обмене.

Для современных процессоров паспортная тактовая частота не превышает 4 Ггц, а время реакции человека на раздражители составляет 0,1–0,2 секунды, то есть 5–10 герц; то есть человек реагирует на сигналы на 8 порядков медленнее компьютера. Можно вывести формулу, связывающую тактовую частоту устройства, скорость распространения электромагнитного сигнала и расстояние, начиная с которого скорость обмена информацией будет определяться скоростью распространения электромагнитного сигнала,

$$L = kc/W, \quad (1)$$

где W – тактовая частота устройства;
c – скорость света;
L – расстояние, начиная с которого скорость обмена информацией определяется распространением сигнала;
k – коэффициент уменьшения скорости распространения электромагнитного сигнала по сравнению с вакуумом.

Для кремния испытания по распространению электрического сигнала не проводились, но большинство испытанных материалов имеет коэффициент скорости 0,6–0,8. Примем для оценочного расчета коэффициент скорости для кремния 0,7.

В результате для процессора с тактовой частотой 4 Ггц получим допустимый размер:

$$L = 0{,}7*3 \cdot 105*105/4 \cdot 109 = 5{,}25 \text{ см}. \quad (2)$$

Как уже указывалось, процессоры таких размеров не выпускаются, они меньше. Но обмен данными процессора с устройствами на материнской плате с такой тактовой частотой невозможен.

Однако и здесь выдерживаются допустимые пределы. Тактовая частота материнских плат меньше, чем у процессоров, она составляет около 1 Ггц. Тогда для обмена данными, согласно формуле, расстояние должно быть не более 21 см. На таком расстоянии вполне можно разместить оперативную память и видеокарту, осуществляя обмен данными с ними через Северный мост. А скорость обмена с прочими устройствами пренебрежимо меньше.

Сказанное не означает, что невозможно создать устройства с тактовой частотой больше 4 Ггц. Вполне можно, если такие устройства будут иметь малую вычислительную мощность, небольшие размеры и будут несильно нагреваться. И устройства со сверхвысокой частотой создаются.

Так, в США в компании Nortrop Grumman создан транзистор на основе фосфида индия (InP), работающий на частоте 1 Тгц [Mei, Yoshida 2015]. Это в сотни раз больше тактовой частоты процессоров для ПК, изделие вошло в книгу рекордов Гиннеса. Однако размеры транзистора составляют всего 25 нанометров. Терагерцевый сигнал поступает в ламповый усилитель; в результате был получен радиопередатчик, излучающий в малоиспользуемом диапазоне частот.

Надо заметить, что сигнал может испускаться с любой частотой. Так, видимый свет имеет частоту излучения 750–790 терагерц, а рентгеновское и гамма-излучение имеют еще большую частоту. Но для обмена информацией требуется время, которое зависит от пути, пройденного электромагнитным сигналом; сверхвысокие частоты здесь бесполезны.

3. Квантовомеханический туннельный эффект. Уменьшение размеров транзисторов положительно сказывается на обоих рассмотренных факторах: снижает выделение тепла и повышает допустимую тактовую частоту обработки данных. Этой возможностью активно пользуются. Технологии совершенствуются, размеры транзисторов уменьшаются. Некоторые из результатов работы по совершенствованию техпроцессов приведены в табл. 1 по данным[6]. В таблице также приведен пересчет размеров на атомы кремния (радиус атома кремния 0,13 нм)[7].

Надо сказать, что эти цифры относятся к разным процессам, а затем стали относиться и к разным объектам. 130 нм и более ранние техпроцессы – это размер транзисторов, величина которых соответствовала линейному разрешению литографического оборудования. Начиная с 45 нм в транзисторах стали использоваться изоляторы на основе не оксида кремния, а оксида гафния, что позволило понизить толщину нанослоя изолятора и избежать туннельного просачивания электронов сквозь изолятор. Следую-

[6] Технологический прогресс в электронной промышленности // Википедия. URL: https://ru.wikipedia.org/wiki/Технологический_процесс_в_электронной_промышленности (дата обращения 5 июля 2022).

[7] Это условная величина, поскольку кремний образует кубическую гранецентрированную решетку алмаза, в состав которой входит много атомов, ее период составляет 0,54 нм.

щим шагом явилось изменение формы транзисторов, что привело к уменьшению плотности тока и снижению сопротивления. А это позволило использовать меньшую силу тока и в конечном счете снизить нагрев материала.

Таблица 1

Техпроцессы нанотехнологий

Год разработки	*Техпроцесс, нм*	*Пересчет на атомы Si*	*Современное применение*
2001	130	500	Применяется редко, для маломощной техники
2006–2007	45/40	170–150	В процессорах Intel, AMD и др.
2009–2012	22/20	84–76	Процессоры Intel 3 и 4 поколений
2017–2018	10	38	Процессоры Intel, Apple, Samsung для мобильной техники
2019–2020	6–5	23–19	Производят TSMS (Тайвань), Samsung, Apple
2022–2023 (план)	3	11	Разрабатывают TSMS, Samsung, Apple
2029 (план)	1,4	5	Планирует Intel

А дальше техпроцесс приобрел черты маркетинга. Оказалось, что реально (и это видно под микроскопом) никому так и не удалось получить транзисторы с размером меньше, чем 25 нм [Tiberius 2019]. Просто за норму техпроцесса стали выдавать другие детали: разрешение фотолитографии, толщину затвора транзистора и даже минимальную ширину дорожки металлизации.

Дело в том, что дальнейшая миниатюризация столкнулась с непреодолимым препятствием уже из области квантовой механики. *С просачиванием электронов сквозь материал за счет туннельного эффекта, который неизбежно возникает согласно соотношению неопределенностей*. А соотношение неопределенностей Гейзенберга – один из фундаментальных законов квантовой механики.

Этот закон, как и многое в квантовой механике, не действует в условиях макромира, поэтому его сложно понять, руководствуясь обычным жизненным опытом. Суть соотношения неопределенностей в том, что в квантовом мире нет точно определенных неком-

мутирующих характеристик, существует предел точности их существования. В частности, существует предел точности определения положения частицы и ее импульса:

$$\Delta x \, \Delta p \geqslant \hbar / 2, \qquad (3)$$

где Δx и Δp – соответственно среднеквадратичные отклонения координаты и импульса частицы, $\hbar$ – приведенная постоянная Планка, $1{,}055 \times 10^{-34}$ дж/сек.

Электрон не находится в определенной точке пространства, имеется электронное облако, внутри которого он существует. Наиболее вероятно нахождение электрона в центре облака, а чем дальше от центра, тем менее вероятно нахождение там электрона. Если на пути электрона возникает барьер, то просто часть электронов оказывается за пределами этого барьера. В реальном мире это будет значить, что часть людей, специально не проходя через стену, просто окажутся на другой стороне стены.

В классический физике объекту для преодоления препятствия надо иметь энергию большую, чем энергия сопротивления препятствия. Пуля может пробить слой дерева, но не пробивает слой металла такой же толщины. Все пули с одинаковой энергией будут вести себя одинаково. А в квантовой механике это вопрос вероятности.

На рис. 3 наглядно показан процесс взаимодействия потока электронов с потенциальным барьером; на рис. 3а электроны еще не достигли барьера; на рис. 3б часть электронов отражается от препятствия, часть проникает сквозь него.

На рис. 3в показано соотношение отраженных и туннелированных электронов. В транзисторе потока электронов нет, каждый электрон взаимодействует с препятствием самостоятельно. Но процесс будет тот же: одни электроны отразятся от барьера, другие пройдут.

Вероятность преодоления энергетического барьера интенсивно снижается с ростом толщины барьера, с ростом энергии, требуемой для его преодоления, с ростом размеров частицы. Для толстого слоя изолятора вероятность туннельного проникновения электрона практически равна нулю. А для ультратонкого слоя она становится значимой.

В наноэлектронных структурах утечка тока может реализоваться при туннельном переходе барьеров между всеми структурами транзистора: сток, исток, затвор, подложка. При переходе технологий от 1 мкм к 0,5 мкм, далее к 100 нм и к менее чем 50 нм доминировали различные токи утечки [Зебрев 2008, с. 226]. Эти процессы моделировали в виде различных барьеров: трапециевидных, треугольных, прямоугольных.

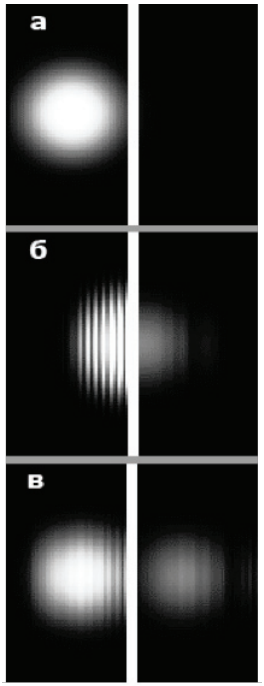

Рис. 3. Отражение и туннелирование потока электронов

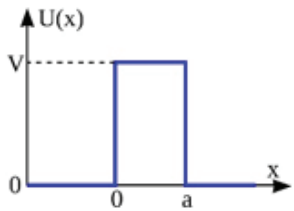

Рис. 4. Прямоугольный барьер

Прямоугольный барьер – самый простой, идеальный случай. Здесь принимается, что энергия барьера одинакова во всей области туннелирования (рис. 4). Общее решение заключается в определении волновых функций частицы (в нашем случае электрона) [Ландау, Лифшиц 2004, с. 108]:

$$\begin{aligned} \psi_1(x) &= e^{ik_1x} + re^{-ik_1x}\ (x < 0); \\ \psi_2(x) &= Ae^{k_2x} + Be^{-k_2x}\ (0 < x < \alpha); \\ \psi_3(x) &= te^{ik_1x}\ (\alpha < x). \end{aligned} \qquad (4)$$

Первая функция описывает движение до барьера, вторая внутри барьера, третья после барьера.

Туннельный эффект наблюдается, если энергия барьера превышает энергию частицы (вариант, невозможный с точки зрения классической физики). Опустим подробности вычислений, они приведены в[8], результат совпадает с [Ландау, Лифшиц 2004, с. 109]. Окончательная формула расчета коэффициента прохождения барьера (коэффициента прозрачности) при $E < U_0$ имеет вид:

$$T = \frac{1}{1 + \frac{U_0^2}{4E(U_0 - E)} \sinh^2(\kappa a)}, \tag{5}$$

здесь U_0 – энергия барьера,
E – энергия частицы (в данном случае электрона),
sinh – гиперболический синус (секанс),
a – толщина барьера,
k – волновое число.

$$\kappa = \sqrt{\frac{2m}{\hbar^2}(U_0 - E)}, \tag{6}$$

здесь m – масса частицы,
$\hbar$ – приведенная постоянная Планка.

Если туннельный эффект нежелателен, то значение T должно быть минимальным. Можно оценить условия, при которых знаменатель будет минимален. На рис. 5 показано значение первого сомножителя в зависимости от соотношения энергии барьера и электрона, в логарифмической и линейной форме. Этот сомножитель будет минимальным, когда энергия электрона будет составлять от 10 до 90% от энергии барьера. В этой области данный сомножитель будет вносить максимальный вклад в туннельный эффект, но это как раз практически используемый диапазон. Ниже энергия электронов такова, что материал будет проявлять свойства изолятора. А выше – транзистор теряет работоспособность, здесь уже электроны могут преодолевать барьер за счет термодинамического разброса энергий.

[8] Туннельный эффект // Википедия. URL: https://ru.wikipedia.org/wiki/Туннельный_эффект (дата обращения 5 июля 2022).

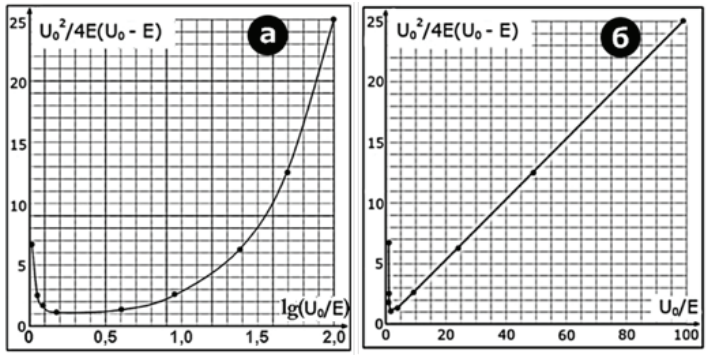

Рис. 5. Зависимость $U_0^2/4E$ (U_0 – E)
от соотношения энергии частицы и барьера:
а – в логарифмической форме; б – в линейной форме

В рассматриваемом диапазоне значение колеблется от 1 до 3, а минимумом этой функции будет соотношение $E/U_0 = 1/2$, при котором:

$$U_0^2/4E(U_0 - E) = 1. \quad (7)$$

Что касается секанса угла (и его квадрата), то они могут принимать значения от единицы до бесконечности. $Sec0^0 = 1$; $Sec90^0 = +\infty$; $Sec180^0 = -1$; $Sec270^0 = -\infty$. Здесь возможен больший разброс.

Минимальное значение $Sec^2ka = 1$ будет при $ka = 0$, что может соблюдаться только при $U_0 = E$. Как уже указывалось, при таких значениях транзистор будет нерабочим, этот вариант рассматривать ни к чему.

В целом же, как показано на рис. 6 [Ильин 2021], параметры, определяющие значение секанса, находятся в довольно сложных соотношениях, в результате чего получаются интересные волнообразные функции. Но мы рассматриваем вариант, где $E/U_0 < 1$, здесь функции проще.

Как можно видеть, туннельные эффекты в наноструктурах разнообразны, а соответствующие расчеты сложны. В самых миниатюрных структурах наиболее актуальным становится устранение утечек тока через окисел кремния, SiO_2. «Экспериментально установлено, что в n-МОПТ (LG = 45 нм) ток в открытом состоянии начинает деградировать при толщине окисла ~ 1.3 нм, в то время как в p-МОПТ подобный эффект наблюдается при dox ~ 1.2 нм.

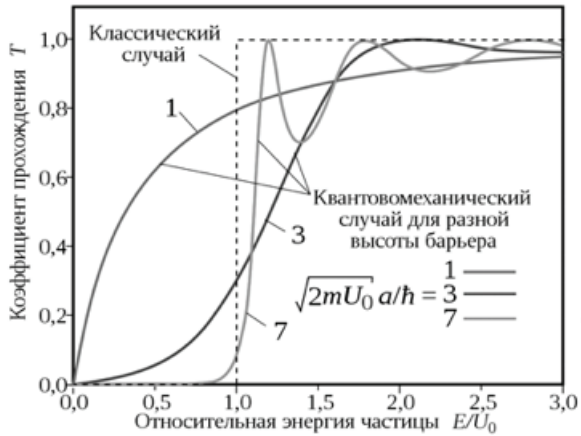

Рис. 6. Коэффициент прохождения прямоугольного барьера с различными параметрами

Уменьшение неоднородности по толщине подзатворного окисла и использование нитридированного окисла может уменьшить эту его толщину еще на 0.1–0.2 нм, т. е. незначительно. Толщина окисла в 90 нм технологии Intel составляет 1.2 нм – это всего 4 атомных слоя решетки окисла кремния» [Зебрев 2008, с. 237].

Проблему решили за счет использования окислов других редких элементов, наилучшим образом себя показали окислы гафния и циркония, HfO_2 и ZrO_2. Толщина меньше не стала, но за счет большей высоты барьера туннелирование снизилось до величины допустимой погрешности.

Проявляется еще один опасный эффект: рост туннельного тока утечки из стока в базу, а также лавинный пробой в этом месте. С этими явлениями можно бороться за счет снижения напряжения питания, но оно и так уже достигло 1 в, ниже нельзя. Можно увеличивать концентрацию примесей (это увеличит барьер), но как раз это и приводит к пробою. Совокупность данных факторов приводит к ограничению ширины базы в районе 25 нм [Tiberius 2019].

Таким образом, за счет квантовомеханических эффектов предельный размер транзисторов на базе кремния остановился на уровне порядка 25 нм. А провозглашение меньших величин – это всего лишь маркетинг.

Поиски новой элементной базы

В связи с описанными проблемами во всем мире ведутся интенсивные поиски принципиально новой элементной базы для вычислительной техники. Создавались опытные образцы:

– оптических компьютеров, с обработкой информации при помощи света;
– молекулярных компьютеров, где под действием электрического сигнала органические молекулы меняют свою структуру;
– ДНК компьютеров с использованием вычислительных возможностей молекул, хранящих генетический код.

По этим направлениям в течение десятилетий велась работа, создавались опытные образцы, но до стадии внедрения в практику так ничего и не дошло.

Перспективным считается создание полевых транзисторов на базе графена. Графен – это специальная двумерная модификация углерода, материал, обладающий многими уникальными свойствами. Работы по созданию графеновых транзисторов начались сравнительно недавно, и пока до стадии практического внедрения не дошли [Шурыгина 2014]. К тому же графеновые наноленты и нанотрубки очень дороги, трудно представить миллиардные тиражи вычислительных устройств на их основе, которые по цене будут конкурентоспособны по сравнению с существующими микропроцессорами. По крайней мере, в обозримом будущем.

До практического использования с недавнего времени дошли квантовые компьютеры. Их возможности уникальны, некоторые задачи они способны решать за разумное время, тогда как самым мощным из существующих ЭВМ для решения тех же задач потребуется время, превышающее возраст Вселенной. К таким задачам относится криптография (вскрытие паролей и шифров), получение новых химических и биологических материалов (поскольку свойства любых веществ в конечном счете определяются квантовыми состояниями входящих в них атомов), искусственный интеллект, логистика, бизнес и пр. (в разных областях приходится выполнять расчеты, превосходящие возможности имеющихся суперкомпьютеров) [Olegbunin 2020].

В мире построено довольно много опытных образцов квантовых компьютеров, доказана их работоспособность. С 2016 г. существуют услуги по предоставлению квантовых облачных вычислений[9], и число таких сервисов постоянно растет.

[9] Undergraduates on a cloud using IBM Quantum Experience // University of Waterloo. URL: https://uwaterloo.ca/institute-for-quantum-computing/news/undergraduates-cloud-using-ibm-quantum-experience (дата обращения 5 июля 2022).

Но освоение квантовых компьютеров сопровождается трудностями. Дело в том, что там не работает привычная математика. Там нет точно определенных значений, там нет битов. Там есть кубиты, каждый из которых принимает значения 0 или 1 с некоторой вероятностью. В квантовом компьютере сразу вычисляются вероятности всех кубитов, число которых может быть весьма большим. Правильным ответом будет тот, вероятность которого максимальна. Если же «правильный ответ» на практике окажется негодным, то можно провести еще один расчет и получить другой «правильный» ответ (это, конечно, сокращенное и упрощенное описание).

Разработан ряд алгоритмов, реализующих квантовые вычисления. Однако они способны решать только узкоспециальные задачи. Поэтому остро стоит проблема расширения возможностей квантовых вычислений, совершенствования «квантовой математики», но не факт, что она будет способна везде заменить обычную математику.

Заключение

Становится все более очевидным, что развитие элементной базы вычислительной техники, основанной на кремниевых транзисторах, серьезно замедлилось. Это позволяет сделать ряд выводов.

1. Теперь сложно говорить о том, что кто-то «отстал навсегда». Процесс, который пребывает в стагнации, догнать намного проще процесса, который экспоненциально развивается.

2. Из философских соображений ясно, что любое интенсивное развитие когда-нибудь заканчивается. Взрыв сверхновой звезды, размножение биологических популяций, великие исторические завоевания – все когда-нибудь заканчивается и сменяется чем-то принципиально новым.

3. Интенсивно ведутся поиски принципиально новой элементной базы. Но пока не найдено полноценной замены полупроводникам на базе кремниевых кристаллов.

Литература

Дудкин 2021a – *Дудкин А.* Таблица характеристик и производительности всех процессоров AMD (2006–2021 гг.). 14 марта 2021 // Alexis Hardware World. URL: https://hww.ru/2021/03/tablica-harakteristik-i-proizvoditelnosti-vseh-processorov-amd-2006-2021-g/ (дата обращения 6 июля 2022).

Дудкин 2021b – *Дудкин А.* Хронология, характеристики и производительность процессоров Intel (2008–2021 гг.). 14 марта 2021 // Alexis Hardware World.

URL: https://hww.ru/2021/03/hronologija-harakteristiki-i-proizvoditelnost-processorov-intel-2008-2021-g/ (дата обращения 6 июля 2022).

Зебрев 2008 – *Зебрев Г.И.* Физические основы кремниевой наноэлектроники: Учеб. пособие. М.: МИФИ, 2008. 288 с.

Ильин 2021 – *Ильин Д.* File:TvsE9.tif // Wikimedia. URL: https://commons.wikimedia.org/w/index.php?curid=106025769 (дата обращения 6 июля 2022).

Казакова 2011 – *Казакова И.А.* История вычислительной техники: Учеб. пособие. Пенза: ПГУ, 2011. 232 с.

Ландау, Лифшиц 2004 – *Ландау Л.Д., Лифшиц Е.М.* Курс теоретической физики: Учеб. пособие для вузов: В 10 т. Т. 3: Квантовая механика (нерелятивистская теория). 6-е изд., испр. М.: ФИЗМАТ ЛИТ, 2004. 800 с.

Миллс 2020 – *Миллс М.* Размеры процессора: почему они не меньше // ITIGIC. 31 августа 2020. URL: https://itigic.com/ru/processor-sizes-why-they-are-not-smaller/ (дата обращения 6 июля 2022).

Родионов 2019 – *Родионов Ю.А.* Технологические процессы в микро- и наноэлектронике: Учеб. пособие. М.; Вологда: Инфра-Инженерия, 2019. 352 с.

Шурыгина 2014 – *Шурыгина В.* «Чудо-материал» – графен новый конкурент на рынке рч-электроники // Электроника. 2014. № 4 (00135). С. 141–148.

Kaiser 2005 – *Kaiser Kennet L.* Transmission lines, matching and crosstalk. London: CRC Press, 2005. 448 p.

Mei, Yoshida et al 2015 – Mei, X., Yoshida, W., Lange, M., Lee, J., Zhou, J., Liu, P.-H., Leong, K., Zamora, A., Padilla, J., Sarkozy, S., Lai, R., Deal, W.R. First Demonstration of Amplification at 1 THz Using 25-nm InP High Electron Mobility Transistor Process // IEEE Electron Device Letters. 2015. Vol. 36 (4). P. 327–329. https://doi.org/10.1109/led.2015.2407193

Tiberius 2019 – *Tiberius.* Технологии микроэлектроники на пальцах: «закон Мура», маркетинговые ходы и почему нанометры нынче не те. Часть 2 // Habr, 19 июня 2019. URL: https://habr.com/ru/post/456298/ (дата обращения 6 июля 2022).

Olegbunin 2020 – *Olegbunin.* Что может квантовый компьютер // Habr, 7 апреля 2020. https://habr.com/ru/company/oleg-bunin/blog/493244/ (дата обращения 5 июля 2022).

References

Dudkin, A. (2021a), “Table of characteristics and performance of all AMD processors (2006–2021)”, *Alexis Hardware World*, March 14, 2021, available at: URL: https://hww.ru/2021/03/tablica-harakteristik-i-proizvoditelnosti-vseh-processorov-amd-2006-2021-g/ (Accessed 6 July 2022).

Dudkin, A. (2021b), “Chronology, characteristics and performance of Intel processors (2008–2021)”, *Alexis Hardware World*, March 14, 2021, available at: URL: https://hww.ru/2021/03/hronologija-harakteristiki-i-proizvoditelnost-processorov-intel-2008-2021-g/ (Accessed 6 July 2022).

Zebrev, G.I. (2008), *Fizicheskie osnovy kremnievoi nanoelektroniki: ucheb. posob.* [Physical foundations of silicon nanoelectronics. Study guide], MEPhI, Moscow, Russia, 288 p.

Ilyin, D. (2021), "File:TvsE9.tif", *Wikimedia*, available at: https://commons.wikimedia.org/w/index.php?curid=106025769 (Accessed 6 July 2022).

Kaiser, Kennet L. (2005), Transmission Lines, Matching and Crosstalk, CRC Press, London, UK, 448 p.

Kazakova, I.A. (2011), *Istoriya vychislitel'noi tekhniki: ucheb. posobie* [History of computer technology. Study guide], PGU, Penza, Russia, 232 p.

Landau, L.D. and Lifshits, E.M. (2004), *Kurs teoreticheskoi fiziki: Ucheb. posob: Dlya vuzov. V 10 t. T. 3: Kvantovaya mekhanika (nerelyativistskaya teoriya)* [Course of theoretical physics. Study guide for universities: In 10 vols. Vol. 3: Quantum mechanics (non-relativistic theory). 6th ed., revsd., FIZMAT LIT, Moscow, Russia, 800 p.

Mei, X., Yoshida, W., Lange, M., Lee, J., Zhou, J., Liu, P.-H., Leong, K., Zamora, A., Padilla, J., Sarkozy, S., Lai, R., & Deal, W.R. (2015), "First Demonstration of Amplification at 1 THz Using 25-nm InP High Electron Mobility Transistor Process", *IEEE Electron Device Letters*, no. 36 (4), pp. 327–329, https://doi.org/10.1109/led.2015.2407193.

Mills, M. (2020), "Processor sizes. Why they are not getting smaller", *ITIGIC,* August 31, 2020. available at: https://itigic.com/ru/processor-sizes-why-they-are-not-smaller/ (Accessed 6 July 2022).

Olegbunin (2020), "What a quantum computer can do", *Habr*, April 7, 2020, available at: https://habr.com/ru/company/oleg-bunin/blog/493244/ (Accessed 5 July 2022).

Rodionov, Yu.A. (2019), *Tekhnologicheskie protsessy v mikro- i nanoelektronike: ucheb. posobie* [Technological processes in micro- and nanoelectronics. Study guide], Infra-Engineering, Moscow; Vologda, 352 p.

Shurygina, V. (2014), " 'Miracle Material' – graphene is a new competitor in the RF electronics market", *Electronics*, no. 4 (00135), pp. 141–148.

Tiberius (2019), "Microelectronics technologies on the fingers: 'Moore's law', marketing moves and why nanometers are not the same today", *Habr*, June 19, 2019, available at: https://habr.com/ru/post/456298/ (Accessed 6 July 2022).

Информация об авторе

Андрей М. Подорожный, кандидат химических наук, доцент, Российский государственный гуманитарный университет, Москва, Россия; 125047, Россия, Москва, Миусская пл., д. 6; poam@mail.ru

Information about the author

Andrey M. Podorozhnyi, Cand. of Sci. (Chemistry), associate professor, Russian State University for the Humanities, Moscow, Russia; bld. 6, Miusskaya Sq., Moscow, Russia, 125047; poam@mail.ru

УДК 004.056+629.7
DOI: 10.28995/2686-679X-2022-3-43-65

Использование беспилотных летательных аппаратов в ОРМ «Наблюдение»: модель угроз безопасности информации

Дмитрий А. Митюшин
Российский государственный гуманитарный университет, Москва, Россия, dalexx@inbox.ru

Аннотация. В настоящее время мобильные роботы и мобильные роботизированные комплексы все чаще применяются не только для решения военных задач, но и для задач, которые в англоязычной прессе называют полувоенными (*paramilitary*). К таким задачам можно отнести задачи, стоящие перед полицией, органами государственной безопасности, гражданской обороны и ликвидации последствий чрезвычайных ситуаций. Разумеется, что правоохранительные структуры не только за рубежом, но и в нашей стране обратили в настоящее время на беспилотные комплексы самое пристальное внимание.

На полицию, в частности, возложен целый ряд задач по защите жизни, здоровья, прав и свобод людей по противодействию преступности, охране общественного порядка, собственности и по обеспечению общественной безопасности. Согласно Федеральному закону «О полиции» полиция обязана использовать в своей деятельности последние достижения науки и техники.

Актуальность статьи связана с тем, что сфера применения мобильных роботов и, в частности, беспилотных летательных аппаратов (БЛА) при решении задач, стоящих перед полицией, будет постоянно расширяться. Совершенствуются вопросы управления как одиночных БЛА, так и их группировкой, уменьшаются массогабаритные характеристики БЛА.

Автор рассматривает три варианта применения беспилотных летательных аппаратов при проведении оперативно-розыскного мероприятия «Наблюдение», которое является одним из наиболее массовых и эффективных для получения оперативно значимой информации. Приводится неформальная модель угроз безопасности информации в контуре «борт – наземный пункт – банк данных» при применении беспилотных летательных аппаратов с учетом особенностей оперативно-розыскной деятельности.

Ключевые слова: беспилотный летательный аппарат, БЛА, полиция, оперативно-розыскная деятельность, защита информации, угрозы безопасности информации, нарушители безопасности информации, модель угроз

Для цитирования: *Митюшин Д.А.* Использование беспилотных летательных аппаратов в ОРМ «Наблюдение»: модель угроз безопасности информации // Вестник РГГУ. Серия «Информатика. Информационная безопасность. Математика». 2022. № 3. С. 43–65. DOI: 10.28995/2686-679X-2022-3-43-65

The use of unmanned aerial vehicles in the OSM "Surveillance". Model of threats for the information security

Dmitrii A. Mityushin
Russian State University for the Humanities, Moscow, Russia, dalexx@inbox.ru

Abstract. At the present time, mobile robots and mobile robotic systems are increasingly being used not only for solving military tasks, but also for the so-called paramilitary ones. They include the tasks facing the police, state security, civil defense and emergency response. Of course, law enforcement agencies not only abroad and in our country have paid close attention to unmanned systems.

In particular, the police are entrusted with a number of tasks to protect the life, health, rights and freedoms of people, to combat crime, protect public order, property and ensure public safety. According to the Federal Law "On Police", the police are obliged to use the latest achievements of science and technology in their activities.

The timeliness of the article is due to the fact that the scope of mobile robots and, in particular, unmanned aerial vehicles (UAV) in solving the issues facing the police will be constantly expanding. The controlling of both single UAVs and their grouping is being improved; the weight and size characteristics of UAVs are decreasing.

The author considers three variants for the use of unmanned aerial vehicles during the operational-search measure "Surveillance" which is one of the most widespread and effective for obtaining operationally significant information. An informal model of threats to the information security in the contour "board – ground point – data bank" is given when using unmanned aerial vehicles, taking into account the specifics of operational-search measures.

Keywords: unmanned aerial vehicle, UAV, police, operational-search measures, information security, information security threats, information security violators, threat model

For citation: Mityushin, D.A. (2022), "The use of unmanned aerial vehicles in the OSM "Surveillance". Model of threats for the information security", *RSUH/RGGU Bulletin. "Informatics. Information security. Mathematics"* Series, no. 3, pp. 43–65, DOI: 10.28995/2686-679X-2022-3-43-65

Введение

В настоящее время комплексы с беспилотными летательными аппаратами (далее – БЛА) все чаще применяются для решения задач, стоящих перед полицией. Некоторое время назад в составе органов внутренних дел РФ были авиационные подразделения, решающие задачи непосредственно в интересах полиции. Однако в 2016 г. Указом Президента России все авиационные подразделения, как органов внутренних дел, так и внутренних войск МВД России, переведены в Федеральную службу войск национальной гвардии (далее – Росгвардия)[1]. С точки зрения автора данной работы, такое решение снизило эффективность деятельности полиции, так как вопросы межведомственного взаимодействия всегда были весьма «болезненными».

Вопросы применения БЛА в деятельности полиции неоднократно рассматривались автором и его коллегами. Так в работах [Митюшин 2010], [Митюшин 2011b] и [Митюшин 2012a] рассматриваются практически все потенциальные возможности БЛА для решения задач, стоящих перед полицией. В работе [Митюшин 2011a] приводятся показатели эффективности комплексов с БЛА при решении полицейских задач. В [Митюшин 2012b] анализируется опыт применения БЛА полицией зарубежных стран. В работе [Митюшин, Казарин 2019] рассматривается отказоустойчивое управление группировкой БЛА на основе разделения секрета. Авторы [Кубасов, Пучков 2012] и [Ананьев, Ерзин, Стафеев 2016] рассматривают применение БЛА для организации подвижной радиосвязи в интересах органов внутренних дел. Авторы [Котарев, Котарева, Александров 2017] и [Савельева, Смушкин 2017] анализируют применение БЛА при осмотре места происшествия и в криминалистической деятельности полиции соответственно. При этом автор статьи неоднократно отмечал, что комплексы с БЛА, применяемые для решения стоящих перед полицией задач, должны

[1] Указ Президента РФ от 05.04.2016 № 157 (ред. от 17.06.2019) «Вопросы Федеральной службы войск национальной гвардии Российской Федерации» // СПС «КонсультантПлюс». URL: http://www.consultant.ru/document/cons_doc_LAW_196284/ (дата обращения 3 апреля 2022).

находится на вооружении соответствующих подразделений полиции, а в ряде случаев, связанных с особой сложностью применяемой техники – в составе отдельных подразделений, подчиненных МВД России или территориальному органу МВД России.

В настоящее время БЛА применяются для наблюдения при охране общественного порядка, при наблюдении за транспортными потоками и фиксации нарушений правил дорожного движения, при проведении спецопераций совместно с подразделениями Росгвардии, при расследовании преступлений и т. д. Упоминается применение БЛА при проведении оперативно-розыскного мероприятия «Наблюдение» [Кудряшов 2018].

Однако при проведении мероприятий, связанных с передачей информации, возникает проблема защиты этой информации, т. е. обеспечения ее целостности, доступности и конфиденциальности.

Вопросами защиты информации в мобильных робототехнических комплексах, к которым относятся БЛА, также занимаются достаточно долго. Наиболее известны работы групп авторов во главе с А.П. Жуком [Жук, Осипов, Гавришев, Бурмистров 2016], И.А. Зикратовым [Зикратов 2017], А.С. Басаном [Басан 2017] и [Басан 2019] и ряд других.

Однако применение БЛА, а также мобильных роботов других типов в ходе проведения оперативно-розыскных мероприятий имеет ряд особенностей, которые влияют на безопасность информации.

Введем некоторые термины и определения, которые понадобятся в дальнейшем.

Термины и определения

Согласно теории оперативно-розыскной деятельности [ОРД 2020] и Федерального закона «Об ОРД»[2]:

оперативно-розыскная деятельность (далее – ОРД) – вид деятельности, осуществляемой гласно и негласно оперативными подразделениями государственных органов, уполномоченных на то указанным Федеральным законом, в пределах их полномочий посредством проведения оперативно-розыскных мероприятий в

[2] Федеральный закон «Об оперативно-розыскной деятельности» от 12.08.1995, № 144-ФЗ (последняя редакция) // СПС «КонсультантПлюс». URL: http://www.consultant.ru/document/cons_doc_LAW_7519/ (дата обращения 3 апреля 2022).

целях защиты жизни, здоровья, прав и свобод человека и гражданина, собственности, обеспечения безопасности общества и государства от преступных посягательств;

оперативно-розыскное мероприятие (далее – ОРМ) – закрепленный в законе структурный элемент ОРД, состоящий из системы действий разведывательно-поискового характера, направленных на решение конкретных тактических задач [ОРД 2020];

наблюдение – ОРМ, которое заключается в визуальном и ином восприятии и фиксации значимых для решения задач ОРД явлений, деяний, событий и процессов; то есть это негласное слежение субъектом ОРД за действиями проверяемых лиц в период их передвижения и нахождения в определенных местах с целью получения оперативно значимой информации и решения частных задач оперативной проверки (разработки).

Приведем еще одно определение ОРМ «Наблюдение» [Алферов, Гришин, Ильин 2016]: *наблюдение* – ОРМ, состоящее в конспиративном (негласном) слежении субъектом ОРД за действиями проверяемых лиц в период их передвижения и нахождения в определенных местах в целях получения оперативно значимой информации и решения частных задач оперативной проверки (разработки).

С точки зрения теории ОРД наблюдение может быть трех видов:
- физическое, т. е. с использованием органов чувств (как правило зрения и слуха);
- электронное, т. е. с использованием технических средств;
- комбинированное.

Объект, представляющий оперативный интерес (далее – ООИ) – любой объект материального мира (в том числе человек), который является источником оперативно значимой информации.

Постановка задачи

Применение БЛА при решении задач полиции отличается от решения военных задач. Иногда такое отличие не очевидно. Прежде всего, это касается условий применения. При охране общественного порядка, контроле транспортных потоков, осмотре места происшествия и решении других задач, БЛА, как и остальная спецтехника, применяется открыто. В то же время при решении большинства задач ОРД и, в частности, при проведении ОРМ «Наблюдение» вся спецтехника применяется негласно, т. е. как от ООИ, так и от окружающих скрываются факт, объект и цель применения специальных технических средств, в том числе и БЛА.

Однако иногда факт применения БЛА скрыть трудно, поэтому возникает дополнительная необходимость в легендировании и зашифровке мероприятия.

Кроме того, в отличие от применения БЛА военного назначения в условиях широкомасштабных боевых действий, при проведении ОРМ у противника отсутствуют средства ПВО и сколь-либо серьезные средства радиоэлектронной борьбы. Однако не стоит забывать о коммерческих комплексах противодействия беспилотным летательным аппаратам.

Применение БЛА в наблюдении за ООИ целесообразно, когда нахождение оперативных сотрудников вблизи объекта невозможно или подступы к нему надежно блокированы, либо когда объект и группу наблюдения разделяют труднопреодолимые преграды [Митюшин 2010] и [Митюшин 2012а]. С развитием миниатюризации возможности по применению БЛА и мобильных роботов других типов будут расширяться.

Необходимо разработать модель угроз безопасности информации, циркулирующей между БЛА, пунктом управления и специализированными базами данных (далее – БД) подразделений оперативно-розыскной информации (далее – ОРИ) МВД России.

Исходные данные

Имеется комплекс БЛА в составе нескольких аппаратов и наземного пункта управления (далее – НПУ), размещенного на мобильной платформе (например, автомобиль-фургон).

БЛА могут быть выполнены по самолетной, вертолетной схемам или по схеме махолета, могут быть замаскированы под птиц или крупных насекомых. Маскировка БЛА существенно снижает вероятность их обнаружения $P_{об}$. С другой стороны, незначительные габаритные размеры БЛА также снижают $P_{об}$.

Поскольку наблюдение с использованием БЛА является электронным, то задачи, решаемые БЛА в ОРМ «Наблюдение», можно свести к следующим:

– ведение негласного аудио и/или видеоконтроля ООИ;
– сброс и размещение спецаппапатуры для контроля ООИ (данная задача интересна, если информация от аппаратуры передается на борт БЛА и с него транслируется на НПУ).

Тип двигателя (электрический или внутреннего сгорания) также влияет на вероятность $P_{об}$ БЛА. У аппарата с электродвигателем акустическая заметность гораздо ниже.

Марка и варианты размещения целевой нагрузки (далее – ЦН) на БЛА в контексте защиты информации объявим несущественными. В качестве допущений предположим следующее:

– ЦН передает аудио- и видеоинформацию с борта БЛА с качеством, достаточным для идентификации ООИ;
– аппаратура аудиоконтроля подавляет акустические помехи со стороны БЛА.

Передача информации на НПУ возможна следующими тремя способами:

– запись информации на борту БЛА и доставка ее после приземления аппарата около НПУ или в иной контролируемой зоне (далее – КЗ) была распространена около 10...15 лет назад, однако с развитием искусственного интеллекта и коммерческих систем противодействия БЛА данный метод может стать снова актуальным;
– передача информации по радиоканалу;
– комбинированный способ.

Основное достоинство первого способа:

– полученная информация наиболее полная, нет потерь, связанных с передачей по радиоканалу;
– информацию сложнее обнаружить коммерческими средствами противодействия.

Недостатки:

– задержка получения информации на время полета, приземления и извлечения накопителя с борта БЛА (иногда это некритично, а иногда может быть очень важным);
– отсутствие контроля за ООИ в режиме реального времени;
– риск неполучения информации в случае падения БЛА, его уничтожения или захвата злоумышленниками;
– накопленная информация может попасть к злоумышленникам (в случае захвата БЛА).

Основное достоинство второго способа – информация передается в реальном масштабе времени. Недостаток – возможность искажения или потери информации, связанной с особенностями передачи по радиоканалу.

Таким образом, наиболее оптимальным является комбинированный способ. В этом случае информация может дублироваться с бортового накопителя в случае ее потери при замираниях сигнала.

Возможны следующие варианты применения:

– одиночный БЛА, когда количество БЛА в воздухе $n = 1$;
– группировка БЛА, при $n > 1$.

В зависимости от конструкции комплекса с БЛА, самих аппаратов, условий применения возможны следующие варианты:

– выделенный канал(ы) связи (рис. 1);
– собственная локальная сеть (рис. 2);
– использование инфраструктуры сети сотовой связи 3G+ (рис. 3).

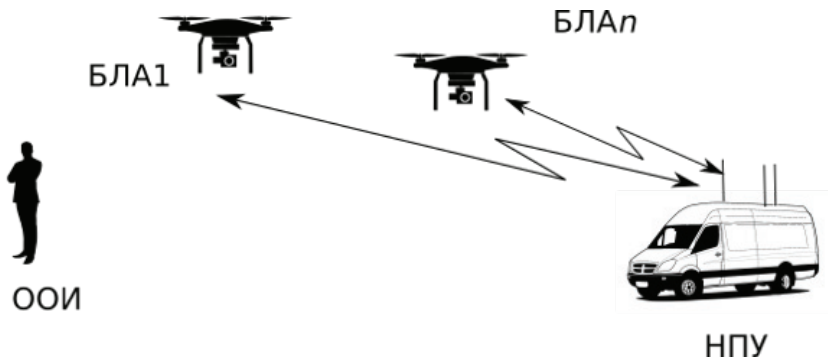

Рис. 1. Вариант применения БЛА
с выделенным каналом связи

При использовании выделенного канала связи обмен информацией между НПУ и БЛА происходит по отдельному каналу. При этом управление БЛА осуществляется по узкополосному двухстороннему командно-контрольному каналу управления (далее – ККУ). Передача оперативно значимой информации с «борта» на НПУ осуществляется по широкополосному каналу.

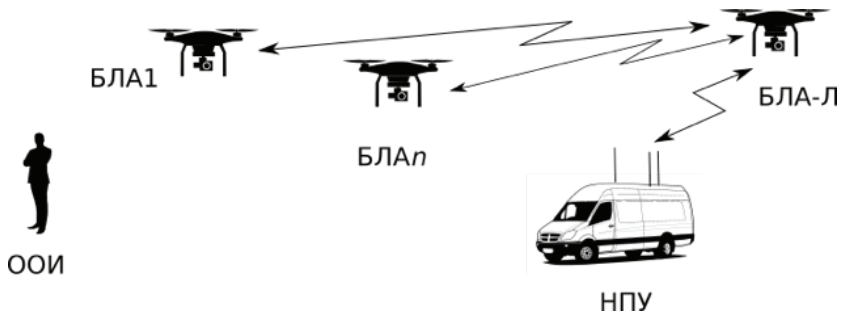

Рис. 2. Вариант применения БЛА
с собственной локальной сетью

На рис. 2 приведен вариант применения БЛА с собственной локальной сетью. В данном случае выделяется БЛА-лидер (БЛА-Л), который осуществляет связь с НПУ. Этот БЛА является своего рода базовой станцией или коммутатором для связи с БЛА, проводящим разведывательно-поисковые мероприятия. ЛВС может быть организована по любым существующим протоколам (например, 802.11, 802.16), либо использовать собственный протокол, что достаточно дорого и чаще всего нецелесообразно.

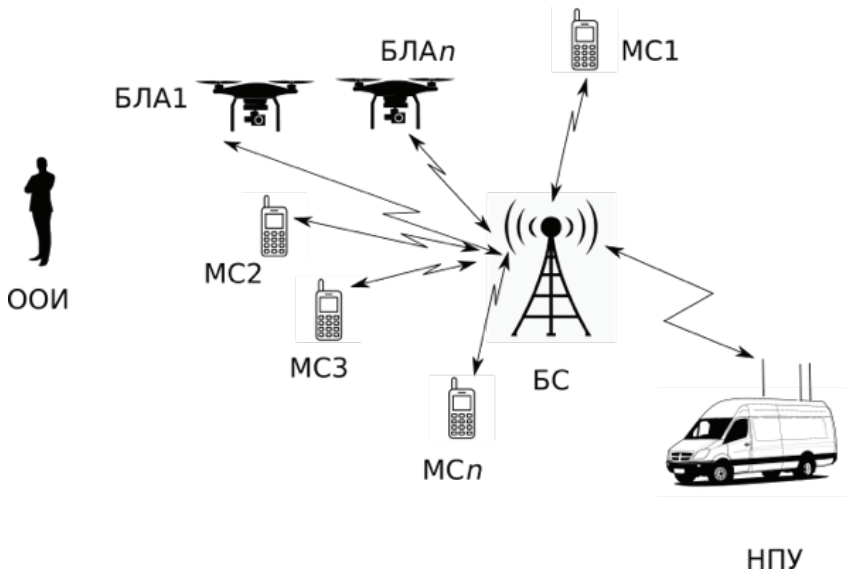

Рис. 3. Вариант применения БЛА
с использование инфраструктуры сети сотовой связи

На рис. 3 представлен вариант применения БЛА с использованием инфраструктуры сети сотовой связи. В данном случае БЛА обмениваются информацией с НПУ, в том числе передача команд и их квитирование осуществляется посредством базовых станций. При этом БЛА или группировка БЛА НПУ являются такими же мобильными станциями (далее – МС), как и обычные сотовые телефоны, которым присвоен собственный IP-адрес в этой сети (как правило, в сети 10.0.0.0/8).

Основной особенностью рассмотренных трех вариантов применения БЛА является то, что создаваемые ими сети являются вре-

менными и время их существования ограничено десятками минут, в перспективе – несколькими часами, а в более отдаленной – сутками.

Еще в качестве допущения, влияющего на эффективность применения БЛА в ходе ОРМ «Наблюдение» и на безопасность передаваемых данных, примем, что в случае обнаружения контролируемым объектом интереса к себе со стороны БЛА или обнаружения НПУ мероприятие считается расшифрованным и сворачивается.

Таким образом, рассмотрим три варианта применения БЛА в ОРМ «Наблюдение»:

Вариант 1 – использование собственных каналов связи.

Вариант 2 – создание и использование собственной локальной сети.

Вариант 3 – использование инфраструктуры сети сотовой связи.

За границы контролируемой зоны примем габариты автофургона.

Информационные потоки

Разделим информационные потоки, циркулирующие в комплексе с БЛА, на *внутренние* и *внешние*. Внутренние информационные потоки – это потоки информации, которые не выходят за пределы элементов комплекса с БЛА, т. е. НПУ и самого аппарата, а также КЗ. Внешние – это потоки информации, которые выходят за пределы элементов комплекса и КЗ. Таким образом, к внешним информационным потокам может относиться:

- получаемая оперативно значимая информация, передаваемая на НПУ или другие БЛА;
- телеметрическая информация и квитанции о выполнении команд, передаваемые с борта БЛА на НПУ;
- видеоинформация, передаваемая с курсовых видеокамер с борта БЛА на НПУ;
- ретранслируемая с борта других БЛА на НПУ информация любого вида;
- командная информация для управления БЛА в пространстве и бортовым оборудованием (включая ЦН), передаваемая с НПУ на борт БЛА;
- навигационная информация и сигналы точного времени, поступающие со спутников на борт БЛА;
- информация о местоположении БЛА, передаваемая на НПУ;
- передаваемая оперативно значимая информация с НПУ в БД подразделения ОРИ.

К внутренним информационным потокам может относиться:

- информация с датчиков (видеокамеры, микрофоны) на блок обработки;

– информация с обработки блока на бортовой накопитель, с накопителя на передатчик;
– командная информация для управления системами БЛА и ЦН в пространстве, передаваемая с НПУ на борт летального аппарата;
– информация инерциальной навигационной системы;
– информация с подсистемы управления НПУ на остальные подсистемы и модули НПУ;
– оперативно значимая информация, поступающая с антенного блока в подсистему обработки данных НПУ;
– с блока обработки на устройства вывода (монитор, наушники и т. д.) и накопитель.

Теперь рассмотрим угрозы безопасности информации при проведении ОРМ «Наблюдение» с использованием БЛА.

Угрозы безопасности информации

При разработке модели угроз используем положения методик ФСТЭК России[3, 4, 5].

Определим следующие угрозы безопасности информации. По одной из классификаций можно выделить три разновидности угроз:

– угрозы утечки информации по техническим каналам утечки;
– угрозы несанкционированного доступа (НСД) к информации, включая угрозы внедрения вредоносных программ;
– угрозы специальных воздействий.

Построим модель нарушителя безопасности информации.

---

[3] Методика определения актуальных угроз безопасности персональных данных при их обработке в информационных системах персональных данных. Утверждена Заместителем директора ФСТЭК России 14 февраля 2008 г. // ФСТЭК России. URL: https://fstec.ru/component/attachments/download/290 (дата обращения 3 апреля 2022).

[4] Методический документ. Методика оценки угроз безопасности информации. Утвержден ФСТЭК России 5 февраля 2021 г. // ФСТЭК России. URL: https://fstec.ru/component/attachments/download/2919 (дата обращения 3 апреля 2022).

[5] Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных (выписка). Утверждена Заместителем директора ФСТЭК России 15 февраля 2008 г. // ФСТЭК России. URL: https://fstec.ru/component/attachments/download/289 (дата обращения 3 апреля 2022).

Модель нарушителя

Рассмотрим две категории нарушителей:

1-я категория – *внешние*, т. е. нарушители, осуществляющие атаку на систему за пределами контролируемой зоны;

2-я категория – *внутренние*, т. е. нарушители, осуществляющие атаку на систему в пределах контролируемой зоны.

К первой категории можно отнести:

– представителей преступных организаций, включая представителей разрабатываемого объекта;
– бывших сотрудников полиции, включая сотрудников подразделения, осуществляющих конкретное мероприятие в данный момент времени;
– посторонних лиц, пытающиеся использовать БЛА в своих целях;
– сотрудников иностранных разведывательных служб.

Внешний нарушитель может иметь возможность:

– осуществить физическое воздействие на БЛА;
– осуществить перехват управления БЛА;
– осуществить несанкционированное воздействие на информацию, циркулирующую в системе.

Ко второй категории можно отнести следующие четыре групп лиц.

Первая группа представляет собой сотрудников подразделения, осуществляющих конкретное ОРМ, которые имеют доступ в автофургон, получают информацию, но не имеют доступа к управлению БЛА.

Лица этой группы:

– имеют доступ к полученной оперативно-значимой информации;
– знают, по меньшей мере, одно легальное имя доступа;
– могут располагать фрагментами информации о топологии группировки БЛА, функционировании БЛА и его ЦН;
– могут располагать именами и вести выявление паролей зарегистрированных пользователей системы и оператора БЛА.

Вторая группа – оператор БЛА. Входит в подразделение, осуществляющее ОРМ, и осуществляет управление БЛА.

Лица данной группы:

– могут иметь доступ к полученной оперативно значимой информации;
– имеют информацию об используемых коммуникационных протоколах и их сервисах;
– знают, по меньшей мере, одно легальное имя доступа;
– могут изменять конфигурацию технических средств информационной системы, вносить в нее программно-аппаратные

закладки и обеспечивать съем информации, используя непосредственное подключение к техническим средствам системы.

Третья группа – инженерно-технический персонал по обслуживанию и подбору БЛА; может входить в подразделение, осуществляющее ОРМ.

Лица данной группы:

- могут иметь доступ к фрагментам полученной оперативно значимой информации;
- имеют информацию об используемых коммуникационных протоколах и их сервисах;
- знают, по меньшей мере, одно легальное имя доступа;
- могут изменять конфигурацию БЛА, вносить в нее программно-аппаратные закладки и обеспечивать съем информации, используя непосредственное подключение к техническим средствам БЛА.

Четвертая группа – разработчики БЛА и программного обеспечения к ним. Лица данной группы непосредственно в ОРМ не участвуют, тем не менее они:

- обладают информацией об алгоритмах и программах обработки информации в БЛА и на НПУ;
- обладают возможностями внесения ошибок, недекларированных возможностей, программных закладок, вредоносных программ в программное обеспечение комплекса с БЛА на стадии разработки, внедрения и сопровождения;
- могут располагать любыми фрагментами информации о вариантах применения БЛА, топологии построения локальной сети, технических средствах обработки и защиты информации как на борту БЛА, так и на НПУ.

Теперь перейдем к модели угроз безопасности информации.

Модель угроз

Для разработки модели необходимо определить перечень угроз, уровень исходной защищенности комплекса с БЛА и вероятность реализации угроз.

Уровень исходной защищенности зависит от ООИ, его квалификации, наличии у него службы контрразведки (в том числе технической и информационной), от места применения (в городе, сельской местности, в лесу, в горах и т. д.).

Считаем, что комплекс применяется в городских условиях, включая все три рассмотренных выше варианта передачи информации. ООИ – представители организованной преступной группировки.

В случае варианта 1 осуществляется шифрование оперативно значимой информации, передаваемой с борта БЛА на НПУ.

В случае вариантов 2 и 3 используется VPN-канал.

Повышение уровня исходной защищенности осуществляется организационными мерами, в первую очередь обеспечение конспиративности проведения мероприятия. Считаем, что все меры конспирации приняты.

С учетом положений методики[6] ФСТЭК России под уровнем исходной защищенности комплекса с БЛА будем понимать обобщенный показатель, зависящий от технических и эксплуатационных характеристик комплекса (табл. 1). В таблице уровни защищенности обозначены символами «В» – высокий, «С» – средний и «Н» – низкий.

Таблица 1

Уровень исходной защищенности комплекса с БЛА при проведении ОРМ «Наблюдение» для трех вариантов применения

Технические и эксплуатационные характеристики комплекса с БЛА	*Уровень защищенности*								
	Вариант 1			*Вариант 2*			*Вариант 3*		
	В	*С*	*Н*	*В*	*С*	*Н*	*В*	*С*	*Н*
1. Устойчивость к внешним воздействиям на БЛА	+			+			+		
2. Устойчивость к внешним воздействиям на НПУ	+			+			+		
3. Целостность системных компонентов БЛА	+			+			+		
4. Целостность системных компонентов НПУ		+			+			+	
5. Контроль управления полетом БЛА		+			+		+		
6. Устойчивость связи БЛА с НПУ		+			+		+		
7. НСД к информации в контуре «борт-НПУ»			+		+		+		
8. НСД к информации в контуре «НПУ-ПОРИ»		+			+		+		

[6] Методика определения актуальных угроз безопасности персональных данных при их обработке в информационных системах персональных данных. Утверждена Заместителем директора ФСТЭК России 14 февраля 2008 г. // ФСТЭК России URL: https://fstec.ru/component/attachments/download/290 (дата обращения 12 сентября 2021).

Таким образом, для первого и второго вариантов применения БЛА, уровень исходной защищенности – средний, так как более 70% характеристик соответствуют уровню не ниже среднего, а для третьего варианта применения – высокий[7]. Следовательно, показатель исходной защищенности Y_1 для первого и второго вариантов применения $Y_1^1 = Y_1^2 = 5$, а для третьего $Y_1^3 = 0$.

Вероятность реализации угрозы Y_2 оценивается по четырехбалльной шкале, где:

0 – угроза маловероятна;

2 – низкая вероятность угрозы;

5 – средняя вероятность угрозы;

10 – высокая вероятность угрозы.

Данные значения приведены в табл. 2.

Коэффициент реализуемости указанных угроз Y будет определяться соотношением[8]:

$$Y = (Y_1 + Y_2)/20$$

По значению коэффициента Y формируется вербальная интерпретация (ВИ) реализуемости угрозы следующим образом:

если $0 \leqslant Y \leqslant 0{,}3$, то возможность реализации угрозы признается низкой («Н»);

если $0{,}3 < Y \leqslant 0{,}6$, то возможность реализации угрозы признается средней («С»);

если $0{,}6 < Y \leqslant 0{,}8$, то возможность реализации угрозы признается высокой («В»);

если $Y > 0{,}8$, то возможность реализации угрозы признается очень высокой («ОВ»).

В табл. 3 приведены значения коэффициента реализуемости угроз и их актуальность *Ак* («+» – актуальна, «–» – неактуальна), интерпретируемая с учетом положений методики[9].

[7] Там же.

[8] Там же

[9] Там же.

Таблица 2

Вероятность реализации угроз при проведении ОРМ «Наблюдение» для трех вариантов применения БЛА

Угрозы безопасности информации	*Вариант 1*		*Вариант 2*		*Вариант 3*	
	Y_2	*Вероятность угрозы*	Y_2	*Вероятность угрозы*	Y_2	*Вероятность угрозы*
1. Угрозы утечки информации по техническим каналам утечки						
1.1. Утечка по виброакустическому каналу	0	маловероятна	0	маловероятна	0	маловероятна
1.2. Утечка по оптическому и электронно-оптическому каналам	0	маловероятна	0	маловероятна	0	маловероятна
1.3. Утечка по ПЭМИН	0	маловероятна	0	маловероятна	0	маловероятна
2. Угрозы НСД к информации, включая угрозы внедрения вредоносных программ (внешний нарушитель)						
2.1. Угроза перехвата трафика «борт-НПУ»	10	высокая	10	высокая	2	низкая
2.2. Угроза перехвата трафика «НПУ-ПОРИ»	5	средняя	2	низкая	0	маловероятна
2.3. Подмена доверенного объекта сети	0	маловероятна	5	средняя	2	низкая
2.4. Внедрение и действие вредоносных программ	0	маловероятна	2	низкая	2	низкая
2.6. Внедрение ложного объекта сети	0	маловероятна	5	средняя	0	маловероятна
2.7. Отказ в обслуживании	0	маловероятна	2	низкая	2	низкая
2.8. Удаленный запуск приложений	0	маловероятна	2	низкая	2	низкая

ISSN 2686-679X • RSUH/RGGU BULLETIN. "Information Science. Information Security. Mathematics" Series, 2022, no. 3

3. Угрозы НСД к информации, включая угрозы внедрения вредоносных программ (внутренний нарушитель)						
3.1. Ошибки персонала	2	низкая	2	низкая	2	низкая
3.2. Доступ к ПО БЛА и НПУ	2	низкая	2	низкая	2	низкая
3.3. Установка вредоносного ПО на элементах комплекса	2	низкая	2	низкая	2	низкая
3.4. Запуск вредоносного ПО на элементах комплекса	2	низкая	2	низкая	2	низкая
3.5. Организация утечки полученной оперативно значимой информации	5	средняя	5	средняя	5	средняя
4. Угрозы специальных воздействий						
4.1. Обнаружение БЛА в полете	5	средняя	5	средняя	2	низкая
4.2. Огневое поражение БЛА в полете	2	низкая	2	низкая	2	низкая
4.3. Радиоэлектронное подавление канала «борт-НПУ»	5	средняя	5	средняя	0	маловероятна
4.4. Перехват управления БЛА	5	средняя	5	средняя	2	низкая

Таблица 3

Реализуемость угроз и их актуальность при проведении ОРМ «Наблюдение» для трех вариантов применения БЛА

Угрозы безопасности информации	Вариант 1			Вариант 2			Вариант 3		
	Y	ВИ	Ак	Y	ВИ	Ак	Y	ВИ	Ак
1. Угрозы утечки информации по техническим каналам утечки									
1.1. Утечка по виброакустическому каналу	0,25	Н	–	0,25	Н	–	0	Н	–
1.2. Утечка по оптическому и электронно-оптическому каналам	0,25	Н	–	0,25	Н	–	0	Н	–
1.3. Утечка по ПЭМИН	0,25	Н	–	0,25	Н	–	0	Н	–
2. Угрозы НСД к информации, включая угрозы внедрения вредоносных программ (внешний нарушитель)									
2.1. Угроза перехвата трафика «борт-НПУ»	0,75	В	+	0,75	В	+	0,1	Н	–
2.2. Угроза перехвата трафика «НПУ-ПОРИ»	0,5	С	+	0,35	С	–	0	Н	–
2.3. Подмена доверенного объекта сети	0,25	Н	–	0,5	С	+	0,1	Н	–
2.4. Внедрение и действие вредоносных программ	0,25	Н	–	0,35	С	–	0,1	Н	–
2.5. Навязывание ложного маршрута сети	0,25	Н	–	0,35	С	–	0,1	Н	–
2.6. Внедрение ложного объекта сети	0,25	Н	–	0,5	С	+	0	Н	–
2.7. Отказ в обслуживании	0,25	Н	–	0,35	С	–	0,1	Н	–
2.8. Удаленный запуск приложений	0,25	Н	–	0,35	С	–	0,1	Н	–

ISSN 2686-679X • RSUH/RGGU BULLETIN. "Information Science. Information Security. Mathematics" Series, 2022, no. 3

3. Угрозы НСД к информации, включая угрозы внедрения вредоносных программ (внутренний нарушитель)										
3.1. Ошибки персонала	0,35	С	–	0,35	С	–	0,1	Н	–
3.2. Доступ к ПО БЛА и НПУ	0,35	С	–	0,35	С	–	0,1	Н	–
3.3. Установка вредоносного ПО на элементах комплекса	0,35	С	–	0,35	С	–	0,1	Н	–
3.4. Запуск вредоносного ПО на элементах комплекса	0,35	С	–	0,35	С	–	0,1	Н	–
3.5. Организация утечки полученной оперативно-значимой информации	0,5	С	+	0,5	С	+	0,25	Н	–
4. Угрозы специальных воздействий									
4.1. Обнаружение БЛА в полете	0,5	С	+	0,5	С	+	0,1	Н	–
4.2. Огневое поражение БЛА в полете	0,35	С	–	0,35	С	–	0,1	Н	–
4.3. Радиоэлектронное подавление канала «борт-НПУ»	0,5	С	+	0,5	С	+	0	Н	–
4.4. Перехват управления БЛА	0,5	С	+	0,5	С	+	0,1	Н	–

Выводы

Таким образом, создана модель угроз безопасности информации в контуре «борт – наземный пункт – банк данных» при применении БЛА в ходе проведения ОРМ «Наблюдение». Ряд полученных показателей имеют низкие значения реализации. В основном это связано с особенностями проведения ОРД. Кроме того, данная модель является достаточно общей. В реальности возможны отклонения в ту или иную сторону в зависимости от характеристик комплекса с БЛА, вариантов и способов применения, типа ООИ, условий применения, подготовленности злоумышленника и ряда других факторов.

Литература

Алферов, Гришин, Ильин 2016 – *Алферов В.Ю., Гришин А.И., Ильин Н.И.* Правовые основы оперативно-розыскной деятельности: учеб. пособие для студентов, обучающихся по специальности 40.05.01 «Правовое обеспечение национальной безопасности» (специализация «Уголовно-правовая») и специальности 40.05.02 «Правоохранительная деятельность» (специализация «Административная деятельность») / Под общ. ред. В.В. Степанова. 3-е изд., испр. и доп. Саратов: Саратовский социально-экономический институт (филиал) РЭУ им. Г.В. Плеханова, 2016. 296 с.

Ананьев, Ерзин, Стафеев 2016 – *Ананьев А.В., Ерзин И.Х., Стафеев М.А.* Обоснование рационального выбора беспилотного летательного аппарата для построения аэромобильной сети связи // Фундаментальные исследования. 2016. № 12–2. С. 251–255.

Басан 2017 – *Басан А.С., Басан Е.С.* Модель угроз для систем группового управления мобильными роботами // VIII Всероссийская научная конференция «Системный синтез и прикладная синергетика»: Сб. науч. трудов. Ростов н/Д.: Южный федеральный университет, 2017. С. 205–212.

Басан 2019 – *Басан Е.С., Басан А.С., Макаревич О.Б., Бабенко Л.К.* Исследование влияния активных сетевых атак на группу мобильных роботов // Вопросы кибербезопасности. 2019. № 1 (29). С. 35–44.

Жук 2016 – *Жук А.П., Осипов Д.Л., Гавришев А.А., Бурмистров В.А.* Анализ методов защиты от несанкционированного доступа беспроводных каналов связи робототехнических систем // Наукоемкие технологии в космических исследованиях Земли. 2016. Т. 8. № 2. С. 38–42.

Зикратов 2017 – *Зикратов И.А., Викснин И.И., Зикратова Т.В., Шлыков А.А., Медведков Д.И.* Модель безопасности мобильных мультиагентных робототехнических систем с коллективным управлением // Научно-технический

вестник информационных технологий, механики и оптики. 2017. Т. 17. № 3. С. 439–449. DOI: 10.17586/2226-1494-2017-17-3-439-449

Котарев, Котарева, Александров 2017 – *Котарев С.Н.*, *Котарева О.В.*, *Александров А.Н.* Использование беспилотных летательных аппаратов для обеспечения безопасности на объектах транспорта // Вестник Восточно-сибирского института Министерства внутренних дел России. 2017. № 4 (83). С. 199–204.

Кубасов, Пучков 2012 – *Кубасов И.А.*, *Пучков Г.Ю.* Анализ технических решений в области организации оперативной радиосвязи и особенности использования беспилотных летательных аппаратов в интересах органов внутренних дел Российской Федерации // Известия ЮФУ. Технические науки. 2012. № 3 (128). С. 41–48.

Кудряшов 2018 – *Кудряшов А.Б.* Применение полицией российских беспилотных летательных аппаратов // Актуальные проблемы науки и практики: Сб. науч. тр. Хабаровск, 2018. С. 201–204.

Митюшин 2010 – *Митюшин Д.А.* Роль и место систем и комплексов с беспилотными летательными аппаратами в деятельности органов внутренних дел // Вестник Московского университета МВД России. 2010. № 12. С. 123–127.

Митюшин 2011a – *Митюшин Д.А.* Вопросы оценки эффективности комплексов и систем с беспилотными летательными аппаратами министерства внутренних дел // Специальная техника. 2011. № 5. С. 40–46.

Митюшин 2011b – *Митюшин Д.А.* Вопросы применения комплексов с БЛА в деятельности органов внутренних дел РФ // Специальная техника. 2011. № 1. С. 26–30.

Митюшин 2012a – *Митюшин Д.А.* Беспилотные системы и комплексы в деятельности полиции. Воронеж: Кварта, 2012. 183 с.

Митюшин 2012b – *Митюшин Д.А.* Опыт применения беспилотных комплексов и систем в деятельности полиции зарубежных стран // Специальная техника. 2012. № 2. С. 9–19.

Митюшин, Казарин 2019 – *Митюшин Д.А.*, *Казарин О.В.* Об отказоустойчивом управлении группировкой беспилотных летательных аппаратов // Материалы XXIV научно-практической конференции «Комплексная защита информации». Витебск: Витебский государственный технологический университет, 2019. С. 89–96.

ОРД 2020 – Оперативно-розыскная деятельность органов внутренних дел. Общая часть. Омск: Омская академия МВД России, 2020. 206 с. URL: https://ordrf.ru/uchebnik/index.html (дата обращения 6 июля 2022).

Савельева, Смушкин 2017 – *Савельева М.В.*, *Смушкин А.Б.* Тактико-технические возможности беспилотных летательных аппаратов в криминалистической деятельности // Проблемы уголовного процесса, криминалистики и судебной экспертизы. 2017. № 2 (10). С. 9–14.

References

Alferov, V.Yu., Grishin, A.I. and Il'in, N.I. (2016), *Pravovyye osnovy operativno-rozysknoj deyatel'nosti: ucheb. posobiye dlya studentov, obuchayushchikhsya po spetsial'nosti 40.05.01 "Pravovoye obespecheniye natsional'noi bezopasnosti* (*spetsializatsiya "Ugolovno-pravovaya"*) *i spetsial'nosti 40.05.02 "Pravookhranitel'naya deyatel'nost'"* (*spetsializatsiya "Administrativnaya deyatel'nost'"*) [Legal foundations of operational-search measures. Study guide for students of specialization 40.05.01 "Legal support of national security" (specialization "Criminal Law") and 40.05.02 "Law enforcement" (specialization "Administrative activity")], Saratov Socio-Economic Institute (branch) of the Plekhanov Russian University of Economics, Saratov, Russia.

Ananev, A.V., Erzin, I.Kh. and Stafeev, M.A. (2016), "Justification in the rational choice of the unmanned aerial vehicle for design of the aeromobile communication network", *Fundamental Research*, vol. 12–2, pp. 251–255.

Basan, A.S. and Basan, E.S. (2017), "Threat model for the group control systems of mobile robots", *VIII Vserossiiskaya nauchnaya konferentsiya "Sistemnyi sintez i prikladnaya sinergetika": Sbornik nauchnyh trudov*, [8th All-Russian Sc. Conf. "System Synthesis and Applied Synergetics". Collection of Proceedings], Southern Federal University, Rostov-on-Don, Russia, pp. 205–212.

Basan, E.S., Basan, A.S., Makarevich, O.B. and Babenko, L.K. (2019), "Studying the impact of active network attacks on a mobile robots group", *Cybersecurity issues*, vol. 1 (29), pp. 35–44.

Kotarev, S.N., Kotareva, O.V. and Alexandrov, A.N. (2017), "The use of unmanned aerial vehicles to ensure safety on transport", *Bulletin of the East Siberian Institute of the Ministry of Internal Affairs of Russia*, vol. 4 (83). pp. 199–204.

Kubassov, I.A. and Puchkov, G.Yu. (2012), "Analysis of engineering solutions in the field of operative communications and specifics of unmanned aerial vehicles (UAVs) use in the interests of internal affairs bodies of the Russian Federation", *Bulletin of Southern Federal University. Technical Sciences*, vol. 3 (128). pp. 41–48.

Kudryashov, A.B. (2018), "The Use of Russian Unmanned Aerial Vehicles by Police Activities", *Aktual'nyye problemy nauki i praktiki. Sbornik nauchnykh trudov* [Current Issues of Science and Practice. Collection of Scientific Papers], Far Eastern Law Institute of the Ministry of Internal Affairs of the Russian Federation, Khabarovsk, Russia, pp. 201–204.

Mityushin, D.A. (2010), "The role and the place of the unmanned aerial vehicle systems and complexes in activities of internal affairs bodies", *Vestnik of Moscow University of the Ministry of Internal Affairs of Russia*, vol. 12, pp. 123–127.

Mityushin, D.A. (2011a), "Issues of using systems and complexes with the unmanned aerial vehicles by internal affairs bodies of the Russian Federation", *Spetsialnaya Tekhnika*, vol. 1, pp. 26–30.

Mityushin, D.A. (2011b), "Matters of efficiency evaluation of unmanned aerial vehicle systems and complexes of the ministry of internal affairs", *Spetsialnaya Tekhnika*, vol. 5, pp. 40–46.

Mityushin, D.A. (2012a), *Bespilotnyye sistemy i kompleksy v deyatel'nosti politsii* [Unmanned Systems and Complexes in the Activities of the Police], Kvarta, Voronezh, Russia.

Mityushin, D.A. (2012b), "The experience of using the unmanned aerial systems and complexes in the police's activities of the foreign countries", *Spetsialnaya Tekhnika*, vol. 2, pp. 9–19.

Mityushin, D.A. and Kazarin, O.V. (2019), "On Fault-Tolerant Control of a Grouping of Unmanned Aerial Vehicles", *Materialy XXIV nauchno-prakticheskoi konferentsii "Kompleksnaya zashchita informatsii"* [Proceedings of the 25th Sc. and Prac. Conf. "Comprehensive data protection"], Vitebsk State Technological University, Vitebsk, Republic of Belarus, pp. 89–96.

Operativno-rozysknaya deyatel'nost' organov vnutrennikh del. Obshchaya chast' (2020), [Operational-search activity of the Internal Affairs Authorities. General part], Omsk Academy of the MIAs of Russia, Omsk, Russia, available at: https://ordrf.ru/uchebnik/index.html (Accessed 6 July 2022).

Savel'eva, M.V. and Smushkin, A.B. (2017), "Tactical and technical capabilities of unmanned aerial vehicles in criminalistical activity", *Criminal procedure, Criminalistics and Judicial examination Problems*, vol. 2 (10), pp. 9–14.

Zhuk, A.P., Osipov, D.L., Gavrishev, A.A. and Burmistrov, V.A. (2016), "Analysis of methods of protection against unauthorized access to wireless communication channels of robotic systems", *H&ES Research*, vol. 8, no. 2, pp. 38–42.

Zikratov, I.A., Viksnin, 1.1., Zikratova, T.V., Shlykov, A.A. and Medvedkov, D.I. (2017), "Security model of mobile multi-agent robotic systems with collective control", *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, vol. 17, no. 3, pp. 439–449, DOI: 10.17586/2226-1494-2017-17-3-439-449

Информация об авторе

Дмитрий А. Митюшин, кандидат технических наук, Российский государственный гуманитарный университет, Москва, Россия; 125047, Россия, Москва, Миусская пл., д. 6; dalexx@inbox.ru

Information about the author

Dmitrii A. Mityushin, Cand. of Sci. (Engineering), Russian State University for the Humanities, Moscow, Russia; bld. 6, Miusskaya Sq., Moscow, Russia, 125047; dalexx@inbox.ru

УДК 004.056
DOI: 10.28995/2686-679X-2022-3-66-84

О формировании баз прецедентов для решения задач информационной безопасности

Эркин Р. Наврузов
Национальный университет Узбекистана имени Мирзо Улугбека; Ташкент, Республики Узбекистан, erkinbek0989@gmail.com

Аннотация. Рассматриваются вопросы формирования баз прецедентов, связанные с проблемой «больших данных» в области информационной безопасности. Эта проблема выражается в отсутствии возможностей использования классических методов анализа данных с целью принятия обоснованных решений. Как правило, к числу основанных требований к математическому и программному обеспечению информационной безопасности относят минимизацию затрат вычислительных ресурсов и высокие показатели обобщающей способности по результатам машинного обучения. Для учета этих требований предложено использовать несколько критериев. Применение критериев связано с формированием латентного признакового пространства с меньшей размерностью, чем исходное и исследование структуры отношений объектов в нем. Анализ многообразия структуры отношений объектов производится в пространствах, сформированных по парам классов (тип DDOS атак, нормальный трафик). Предложено правило для включения объекта в состав базы прецедентов. Эффективность использования базы прецедентов демонстрируется (по тестовой выборке) на алгоритмах распознавания k ближайших соседей, Random Forest и SVM. Один из типов DDOS-атак тестовой выборки не представлен на обучении. Описание объектов задано без отбора и с отбором информативных признаков, а также по результатам нормирования данных и без него.

Ключевые слова: латентные признаки, база прецедентов, большие данные, селекция обучающих выборок

Для цитирования: *Наврузов Э.Р.* О формировании баз прецедентов для решения задач информационной безопасности // Вестник РГГУ. Серия «Информатика. Информационная безопасность. Математика». 2022. № 3. С. 66–84. DOI: 10.28995/2686-679X-2022-3-66-84

On forming the precedent bases for solving problems of the information security

Erkin R. Navruzov
Mirzo Ulugbek National University of Uzbekistan, Tashkent, Uzbekistan, erkinbek0989@gmail.com

Abstract. The issues of the formation of bases of precedents related to the one of "big data" in the field of information security are considered. That manifests itself in the lack of opportunities for using classical methods of the data analysis in order to make informed decisions. As a rule, among the basic requirements for the mathematical and software information security include minimizing the cost of computing resources and high rates of generalizing ability based on the results of machine learning. To record those requirements, it is proposed to use several criteria. The application of criteria is associated with the formation of a latent feature space with a smaller dimension than the original one and the study of the structure of object relations in it. The analysis of the structure diversity of objects relations is carried out in the spaces formed by pairs of classes (type of DDOS attacks, normal traffic). A rule is proposed for including an object in the database of precedents. The effectiveness of using the precedent base is demonstrated (based on a test sample) using the k nearest neighbor recognition algorithms, Random Forest, and SVM. One of the types of DDOS attacks of the test sample is not presented in the training. Descriptions of objects are given without selection and with selection of informative features, according to the results of data normalization and without it.

Keywords: latent features, precedent database, big data, selection of training samples

For citation: Navruzov, E.R. (2022), "On forming the precedent bases for solving problems of the information security", *RSUH/RGGU Bulletin. "Information Science. Information Security. Mathematics" Series,* no. 3, pp. 66–84, DOI: 10.28995/2686-679X-2022-3-66-84

Введение

Классический подход к селекции обучающих выборок сводится к повышению обобщающий способности алгоритмов. В задачах информационной безопасности (ИБ) основные трудности процесса селекции связаны с проблемами Big Data. В классическом варианте при селекции используются такие понятия, как шумовые объекты и шумовые признаки. Разработано множество эвристических кри-

териев для удаления объектов и признаков, попадающих под определение шумовых. Практическое использование этих критериев не всегда распространяется на данные по ИБ.

Основными целями селекции по формированию баз прецедентов в ИБ является достижение максимальной точности классификации при минимальных затратах вычислительных ресурсов. Как желательное свойство, определяется сохранение особенностей известных типов DDOS-атак в этих базах. Одним из способов достижения этого свойства является исследование структуры отношений объектов методами кластерного анализа.

Известно, что как неотъемлемый атрибут кластеризации рассматривается мера расстояния между объектами и признаками. Выбор меры расстояния – это отдельная трудноформализуемая и реализуемая проблема из-за «проклятия» размерности, разнотипности и разномасштабности измерений, наличия пропусков в данных и т. д.

Альтернативой близости объектов в «сыром» пространстве служит близость в латентном, построенном на базе «сырого». Считается, что из-за низкой размерности латентного пространства в нем можно решать задачи группировки с использованием метрических алгоритмов.

В основе выбора латентного пространства лежит утверждение (гипотеза) о том, что объекты, близкие в «сыром» пространстве, будут близкими и в латентном. Выбор размерности латентного пространства может производиться алгоритмическим путем, либо лицом, принимающим решение (ЛПР).

При разработке методов машинного обучения необходимо принимать решение о том, какие признаки следует использовать в качестве входных данных для распознающего алгоритма. Как правило, отбор признаков или выбор новых признаков при формировании пространства для описания предшествует процессу обучения и снижает его размерность.

Одним из способов снижения размерности пространства является использование латентных признаков. Отметим некоторые особенности, которые нужно учитывать при формировании латентного признакового пространства:

- «сырые» признаки измеряются в одной шкале (интервальной или номинальной);
- «сырые» признаки являются разнотипными;
- методы снижения размерности используют (не используют) классификацию объектов;
- методы снижения размерности являются линейными (нелинейными).

В данной работе рассматривается разнотипное признаковое пространство, на описание объектов в котором введено разделение на классы. Для снижения размерности «сырые» признаки разбиваются на непересекающиеся группы, по каждой из которой формируются латентные признаки.

Процесс разбиения может производиться:

– эвристическим путем по заданным ЛПР числом групп и количеством признаков в них;
– по правилам иерархической агломеративной группировки.

Правила иерархической группировки различаются по:

– выбору признака-кандидата на включение его в группу;
– реализации «жадной» (не «жадной») стратегии оптимизации.

Общим для алгоритмов иерархической группировки является следующее:

– процесс вычисления латентного признака реализуется через линейное отображение сырых признаков на числовую ось;
– число групп изначально неизвестно;
– слабо учитывается свойство робастности пространства;
– допускается введение регуляризаторов на число групп.

Латентные признаки могут представлять комбинации из номинальных и количественных признаков. Для формирования процесса принятия решения было разработано несколько способов получения латентных признаков из исходных разнотипных признаков.

Природа среды (функция плотности распределения) описаний DDOS-атак считается неизвестной. По этой причине необходим выбор метода вычисления латентных признаков с учетом робастности пространства.

Базовыми понятиями при формировании латентных признаков является метод вычисления обобщённых оценок и устойчивость разнотипных признаков. Формальное описание метода (стохастического) появилось в [Игнатьев 2011], модифицированной детерминистический алгоритм в [Игнатьев, Рахимова 2021], [Мирзаев 2021]. Модификация связана с вычислением показателя устойчивости сырых (количественных и номинальных) признаков по значениям функции принадлежности к классам. Эти значения определяются алгоритмическим путем и служат для нелинейных преобразований признаков. Происходит выбор нового пространства, вычисление обобщенных оценок в котором реализуется детерминистическим алгоритмом.

При синтезе латентных признаков из исходных могут применяться правила иерархической агломеративной группировки. В основе доказательства единственности числа групп (латентных

признаков) и состава исходных признаков в них лежит принцип динамического программирования.

В [Саидов 2017] были предложены правила иерархической агломеративной группировки исходных признаков для нелинейного отображения их значений в описании объектов на числовую ось. В основе правил для попарного объединения признаков при группировке лежит вычисление компактности объектов классов по латентному признаку в границах непересекающихся интервалов.

Отбор прецедентов для машинного обучения является трудоемкой процедурой. Например, в методе SVM в качестве таковых выбирается множество граничных объектов в общем-то расширенном признаковом пространстве. Расширение пространства происходит за счет использования ядерных функций. Граничные объекты и шумовые объекты в их составе не являются претендентами на улучшение обобщающей способности алгоритмов, так как это могут сделать объекты-эталоны на некотором удалении от границы.

Вариант с поиском объектов-эталонов через решение задачи с минимальным покрытием [Ignatyev 2018] выглядит привлекательным только для малых выборок. Есть идея реализации построения базы прецедентов для ИБ в несколько этапов. Последовательность этапов может быть следующей:

– формирование наборов латентных признаков по парам (тип DDOS-атак, нормальный трафик);
– разбиение на кластеры в латентном пространстве;
– отбор представителей по каждому кластеру.

На объединении результатов отбора представителей по каждому типу DDOS-атак формируется разбиение на классы (тип DDOS-атак, нормальный трафик). Проводится сокращение числа объектов аналогично описанной выше последовательности этапов по каждому типу.

Формирование базы прецедентов по DDOS-атакам предлагается проводить с использованием методов группировки по заданному число групп и по наборам латентных признаков. Унифицированные наборы признаков определяется по двум вариантам разбиения объектов на классы:

– тип DDOS-атак, нормальный трафик;
– объединение 12 типов DDOS-атак, нормальный трафик.

Смысл унификации заключается в равенстве мощности наборов латентных признаков и использовании упорядочения «сырых» признаков по множеству пар (тип DDOS-атак, нормальный трафик), которые в общем-то отличаются друг от друга, как и составы наборов (групп) «сырых» признаков для синтеза латентных.

Необходимо принимать следующие решения:
- какое число групп использовать по указанным выше двум вариантам?
- как обосновывать зависимость (независимость) числа групп от длины обучающей выборки?
- какие объекты оставлять (удалять) для наполнения базы прецедентов?

Предлагается для решения проблем Big Data использовать метрические алгоритмы распознавания. Выбор метрических алгоритмов связан с:
- отказом от использования методов кросс-валидации для вычисления обобщающей способности;
- селекцией обучающих выборок через их минимальное покрытие объектами-эталонами и отбором информативных наборов признаков;
- формированием латентного признакового пространства и использованием мер расстояния между объектами;
- методом нормирования данных.

Проблема формирования латентного признакового пространства связана с выбором критериев и алгоритмов по их оптимизации [Ignatyev 2018]. Как правило, алгоритмы оптимизации используют «жадную» стратегию поиска решения. «Жадная» стратегия хорошо проявляется на высоких результатах обучения, но резко ухудшает показатели обобщающей способности на тестовых выборках. Указанные особенности машинного обучения демонстрируются на алгоритме иерархической агломеративной группировки бинарных признаков. Стремление к минимизации эмпирического риска на алгоритмах распознавания приводит к проблеме переобучения.

Доказательство эффективности учета робастности при формировании латентных признаков приводится в [Ignatyev 2021]. Для решения проблемы робастности предлагается использовать упорядочение «сырых» признаков по их устойчивости. Устойчивость рассматривается как универсальный показатель для сравнения номинальных и количественных признаков. С помощью этого показателя можно осуществлять отбор информативных наборов, руководствуясь следующим утверждением: «Вероятность вхождения признака с относительно малым значением устойчивости в состав набора ничтожно мала».

Анализ визуального представления объектов с описанием типов DDOS-атак и нормального трафика указывает на отсутствие приемлемого с точки зрения классификации «зазора» между ними. Атаки маскируются под нормальный трафик, тем самым усиливая

свою неразличимость. Есть проблемы в оценке реальной плотности распределения данных, которая в общем-то считается неизвестной. Для решения этой проблемы предлагается использовать нормирование признаков с помощью параметров распределения в виде границ между двумя классами на числовой оси.

Структура размещения описаний типов DDOS-атак и нормального трафика в признаковом пространстве приводит к идее о целесообразности использования локальных метрик для всех объектов или их части в виде объектов-эталонов минимального покрытия.

Постановка задачи

Рассматривается множество из $E_0 = \{S_1,...,S_m\}$ объектов, разбитое на $l + 1$ непересекающихся подмножеств (классов) K_0, $K_1,...,K_l$, из которых $K_1,...,K_l$ представляют описание типов DDOS атак, K_0 – нормальной трафик. Каждый объект описывается набором признаков $X(n)$, ξ – из которых измеряются в номинальной, n-ξ – в качественных шкалах.

Считается, что на $X(n)$ определена процедура упорядочения признаков по значениям их устойчивости на каждой паре (K_0, K_i), $i = 1,...,l$ и формирования наборов $X^i(n_1)$, ..., $X^i(n_t)$ для вычисления латентных признаков $Y^i(t) = (y^i_1, ..., y^i_t)$, $n_1 + ... + n_t \leqslant n$ по ним.

Требуется:

- на каждой паре (K_0, K_i) по $Y^i(t)$ провести разбиение объектов на заданное число непересекающихся групп $G^i_1,..., G^i_p$;
- выделить эталоны по каждой из групп $G^i_1,..., G^i_p$;
- сформировать объединение эталонов в качестве обучающей выборки.

Процесс объединения эталонов можно считать как полной, так и частичной селекций обучающих выборок. Отсутствие полной определенности связано с использованием моделей алгоритмов распознавания. Например, для алгоритма «ближайший сосед» процесс селекции может быть продолжен через поиск минимального покрытия выборки объектами-эталонами.

Заметим, что число «сырых» признаков в наборах $X^i(n_1)$, ..., $X^i(n_t)$ совпадает, различается лишь их состав. На содержимое состава влияет последовательность сырых признаков, упорядоченных по значениям их устойчивости.

Формирование наборов латентных признаков связано с:

- нелинейным преобразованием значений сырых признаков в унифицированное представление в {1,2};

– вычислением значений устойчивости по каждому признаку;
– группировкой непересекающихся наборов сырых признаков по упорядоченной последовательности значений их устойчивости;
– вычислением латентных признаков по каждому набору.

Группировка признаков

Критерии применяются для анализа многообразия отношений значений количественных признаков объектов на числовой оси и предобработки данных. Особенности предобработки данных заключаются в нелинейных преобразованиях разнотипных (качественных и номинальных) признаков через значения функции принадлежности объектов к классам. Поиску экстремумов критериев предшествует упорядочение значений признаков по неубыванию.

Пусть для значений признака $x_c \in X\ (n)$ в описании объектов $K_0 \cup K_j, j = 1,...,l$ построена упорядоченная по неубыванию последовательность

$$r_1,...,r_i,...,r_h,\ h = |\ K_0 \cup K_j\ |. \qquad (1)$$

В качестве границ двух непересекающихся интервалов $[\pi_1;\ \pi_2]$, $(\pi_2;\ \pi_3]$, определяемых по (1), используются $\pi_1 = r_1$, $\pi_2 = r_i$, $1 < i < h$, $\pi_3 = r_h$. Интервалы $[\pi_1;\ \pi_2]$ и $(\pi_2;\ \pi_3]$ идентифицируются, соответственно, как первый и второй. Вес признака у объектов классов по (1) вычисляется [Игнатьев 2011] как максимум произведения внутриклассового сходства и межклассового различия по критерию

$$\left(\frac{\sum_{d=1}^{2}\sum_{i=1}^{2}\left(u_{a[i]}^{d}-1\right)u_{a[i]}^{d}}{\sum_{i=1}^{2}\left|K_{a[i]}\right|\left(\left|K_{a[i]}\right|-1\right)}\right)\left(\frac{\sum_{d=1}^{2}\sum_{i=1}^{2}u_{a[i]}^{d}\left(\left|K_{a[i]}\right|-u_{a[3-i]}^{d}\right)}{2\left|K_0\right|\left|K_j\right|}\right)\rightarrow \max_{\pi_1<\pi_2<\pi_3}, \qquad (2)$$

где $u_{a[i]}^{d}\left(u_{a[3-i]}^{d}\right)$ – количество значений признака x_c у объектов из класса $K_{a[i]}$, $(K_{a[3-i]})$ в d-ом интервале, $a\ [1] = 0$, $a\ [2] = j$. Множество допустимых значений критерия (2) принадлежит (0;1] и используется для оценки объектов классов на числовой оси. Если в каждом интервале содержатся все значения признака объектов из одного класса, то его вес равен 1.

Граница между классами (порог) для количественного признака x_a вычисляется как

$$\Gamma_a = \frac{\pi_2 + \eta}{2}, \tag{3}$$

где η – ближайшее к π_2 значение из интервала $(\pi_2; \pi_3]$, определяемого по (2). Считается, что природа среды данных при вычислении порога по (3) неизвестна. Значение (2) интерпретируется как мера компактности объектов выборки из двух классов на числовой оси. В данной работе эта мера применяется для оценки наборов исходных признаков по значениям, формируемым по ним латентных признаков.

Преобразование количественных признаков в градации номинальных по (1) позволяет упростить поиск схожих объектов по обучающей выборке Ω_j (K_0, K_j) с помощью критерия

$$\left| \frac{d_0^i(u,v)}{|K_t|} - \frac{d_j^i(u,v)}{|K_j|} \right| \to \max, \tag{4}$$

где $d_0^i(u,v)$ $\left(d_j^i(u,v)\right)$ – количество объектов класса K_0 (K_j) в интервале $[r_u, r_v]^i$, $i = 1,..., \tau_c$, τ_c, – число непересекающих интервалов (градаций признака) $x_c \in X(n)$.

Заменим значения признака $x_c \in X(n)$ у объектов из Ω_j на номера интервалов $[r_u, r_v]^\mu$, $\mu = 1,...,\tau_c$. Тогда $d_{0c}(\mu) = d_0^\mu(u, v)$, $d_{jc}(\mu) = d_j^\mu(u, v)$. Значение функции принадлежности признака $x_c \in X(n)$ к K_0 определяется как

$$f_c(\mu) = \frac{d_{0c}(\mu) / |K_0|}{d_{0c}(\mu) / |K_0| + d_{jc}(\mu) / |K_j|}. \tag{5}$$

Обозначим $z_c(\mu) = d_{0c}(\mu) + d_{jc}(\mu)$. Тогда вычисление значения устойчивости признака $x_c \in X(n)$ будет таким

$$g_c = \frac{1}{m} \sum_{\mu=1}^{\tau_c} \begin{cases} f_c(\mu) \cdot z_c(\mu), f_{ci} > 0.5, \\ \left(1 - f_c(\mu) \cdot z_c(\mu)\right), f_{ci} < 0.5 \end{cases}. \tag{6}$$

Заметим, что вариант с $f_c(\mu) = 0.5$ в (6) не рассматривается. Вероятность такого события в задачах Big Data близка к нулю.

Значения устойчивости (6) необходимы для преобразования исходных «сырых» признаков в значения бинарных. Обоснование такого преобразовании приводится в [Ignatiev 2021] для решения задачи снижения размерности пространства. Бинарные признаки используются для вычисления значений обобщенных оценок объектов.

Граница между объектами классов по (5) для $x_c \ X\ (n)$ определяется как

$$G_c = (q1 + q2)/2, \tag{7}$$

где $q2 = \max\{f_c(\mu) \mid 0.5 - f_c(\mu) > 0, \mu = 1, ..., p_c\}$,
$q1 = \min\{f_c(\mu) \mid 1 - f_c(\mu) < 0.5, \mu = 1, ..., p_c\}$.

Вес признака x_c для объектов, значения которых представлены через нелинейные преобразования (6), можно вычислить по (2) либо через градации из {1, 2} в номинальной шкале. При вычислении значения градации a_{ic}, $c \in D$, $D = \{1, \ldots, n\}$ для объекта $S_i = \{x_{iu}\}_{u \in D}$ с использованием (7) и учетом шкал измерений рассматривается одно из двух условий: $x_{ic} \in [r_u; r_v]^{\mu}$ либо $x_{ic} = \mu$. Проверка условий необходима для выбора значений функции принадлежности $f_c(\mu)$ для вычисления a_{ic} как

$$a_{ic} = \begin{cases} 1, f(\mu) < G_c, \\ 2, f(\mu) > G_c. \end{cases}$$

Обозначим через g_{1c}^{j}, g_{2c}^{j} количество значений градации $j \in \{1,2\}$ признака $x_c \in X\ (n)$ в описании объектов соответственно класса K_1 и K_2. Межклассовое различие по признаку x_c определяется как величина

$$\lambda_c = 1 - \frac{\sum_{j=1}^{2} g_{1c}^{j} g_{2c}^{j}}{|K_1||K_2|}. \tag{8}$$

Степень однородности (мера внутриклассового сходства) β_c значений градаций признака по классам K_1, K_2 вычисляется по формуле

$$\beta_c = \frac{\sum_{j=1}^{2} g_{1c}^{j}\left(g_{1c}^{j}-1\right)+ g_{2c}^{j}\left(g_{2c}^{j}-1\right)}{|K_1|\left(|K_1|-1\right)+|K_2|\left(|K_2|-1\right)}. \quad (9)$$

С помощью (8), (9) вес признака $x_c \in X(n)$ в номинальной шкале аналогично (2) определяется как произведение внутриклассового сходства и межклассового различия

$$w_c = \beta_c \lambda_c. \quad (10)$$

Множество допустимых значений весов признаков, вычисляемых по (10), принадлежит интервалу (0;1].

Для вычисления обобщенных оценок объектов [Ignatiev 2021] на E_0 используются вклады градаций признаков. Вклад градации $j \in \{1, 2\}$ признака $x_c \in X(n)$ определяется как

$$\eta_c(j) = w_c\left(\frac{\alpha_{cj}^{1}}{|K_1|}-\frac{\alpha_{cj}^{2}}{|K_2|}\right), \quad (11)$$

где $\alpha_{cj}^{1}, \alpha_{cj}^{2}$ – количество значений градации j признака x_c соответственно в классах K_1 и K_2, w_c – вес признака x_c по (10). Обобщенная оценка объекта $S_r \in E_0$ по описанию в номинальной шкале измерений $S_r = \{a_{ri}\}_{i\in D}$ и вкладам (11) вычисляется как:

$$R(S_r) = \sum_{i\in D}\eta_i(a_{ri}). \quad (12)$$

Базовым понятием для формирования информативных наборов признаков является устойчивость признака (6). Множество допустимых значений (6) принадлежат (0,5; 1]. Устойчивость $g_c = 1$, если по границе (7) объекты без ошибок разделяются на классы K_0 и K_j.

Значение устойчивости (6) можно использовать для сокращения размерности признакового пространства. Одним из вариантов выбора латентного признакового пространства является:

– упорядочение признаков из $X(n)$ по (6) на $_j(K_0, K_j)$

$$x_1^j, \ldots, x_n^j; \quad (13)$$

– разбиение (13) на заданное число непересекающихся групп $X^j(n_1), \ldots, X^j(n_t)$;

– формирование набора $Y(t) = (y_1^j,\ldots, y_t^i)$ методом обобщенных оценок по (12) на $X^j(n_1),\ldots, X^j(n_t)$.

Вычислительный эксперимент

Для создания базы прецедентов использовался набор данных CICDDOS 2019 [Sharafaldin, Lashkari, Hakak, Ghorbani 2019], содержащий описание 12 типов DDOS-атак и нормального трафика через протоколы прикладного уровня TCP/UDP.

Из набора данных сформированы 12 выборок по парам классов (нормальный трафик, тип DDOS-атак). По каждой выборке нормальный трафик представляли одни и те же объекты. На 12 выборках проведено нелинейное преобразование значений сырых признаков объектов в {1, 2} и упорядочение их по значению устойчивости (6) в виде последовательности (13). Порядок следования первых 10 элементов исходной последовательности показан в табл. 1.

Таблица 1

Порядок следования признаков по устойчивости (6)

№	*Тип DDOS атаки*	*Номера элементов последовательности*									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	DNS	38	7	14	8	53	52	69	3	64	37
2	LDAP	38	7	8	53	52	14	40	39	6	4
3	MSSQL	38	7	8	53	14	52	40	39	1	18
4	NetBIOS	38	7	8	53	79	37	52	63	4	64
5	SNMP	38	7	8	53	52	40	39	14	6	4
6	UDP-Lag	79	20	37	23	62	2	66	1	18	21
7	WebDDoS	66	69	34	55	1	2	62	35	39	67
8	SYN	66	73	74	76	18	15	16	21	75	72
9	NTP	39	6	8	53	38	7	40	52	14	22
10	SSDP	38	7	8	53	39	40	52	6	9	23
11	UDP	38	7	8	53	39	40	52	22	18	6
12	TFTP	6	39	38	7	40	8	53	52	23	18

Последовательность (13) по каждой выборке из пары классов $(K_0, K_j), j = 1,\ldots,12$ была разбита на 8 равных по мощности групп. Из этих групп сформированы наборы латентных признаков для каждого из 12 типов DDOS-атак. Результаты разбиения объектов по каждой паре (K_0, K_j) на 1000 групп алгоритмом *k-means* из [Scikit-learn 2019] показаны в табл. 2. Принцип подсчета количества групп был следующим. Если число представителей из K_0 (K_j) в группе было больше 50%, то группу относили к нормальному трафику (типу DDOS-атак).

Таблица 2

Разбиение типов DDOS-атак на группы методом *k-means*

№	*Тип DDOS-атак*	*Число объектов*	*Количество групп*	
			нормального трафика	*DDOS-атак*
1	DNS	158 846	661	339
2	LDAP	79 506	763	237
3	MSSQL	244 291	800	200
4	NetBIOS	68 652	729	251
5	SNMP	170 967	644	356
6	UDP-Lag	139 720	784	216
7	WebDDoS	51 049	972	28
8	SYN	206 154	679	321
9	NTP	150 635	593	407
10	SSDP	150 635	13	983
11	UDP	150 635	792	208
12	TFTP	150 635	810	190
	Итого	1 595 977	8240	3736

Из каждой группы (см. табл. 2) в обучающую выборку выбиралось по одному представителю по отношению близости его к центру группы. Близость определялась по евклидовой метрике, представитель был из доминантного класса в этой группе.

Для исследования отношений объектов рассматривалось нормирование по интервалам $[\pi_1; \pi_2]$ $(\pi_2; \pi_3]$, вычисляемых по (2) для $x \in X(n)$ как

$$x^* = \frac{x - \pi_2}{\pi_3 - \pi_1}. \qquad (14)$$

Визуальное представление в R^2 методом t-sne [Scikit-learn 2019] базы прецедентов без нормирования, с нормированием в [0;1] и по (14) демонстрируется соответственно на рис. 1, 2, 3.

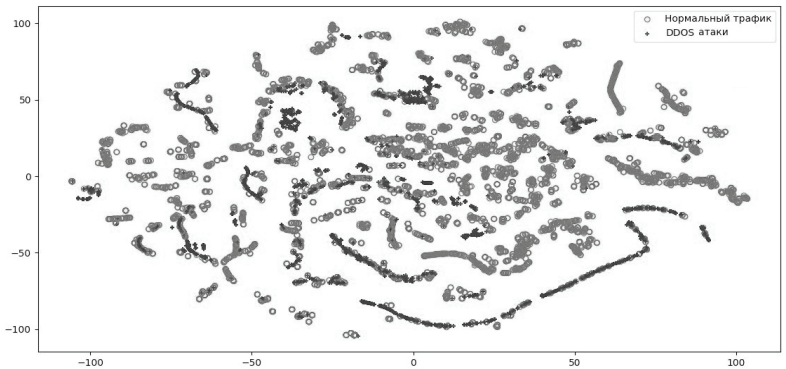

Рис. 1. Визуальное представление выборки прецедентов без нормирования

На рис. 1 видно, что DDOS-атаки маскируются над нормальный траффик.

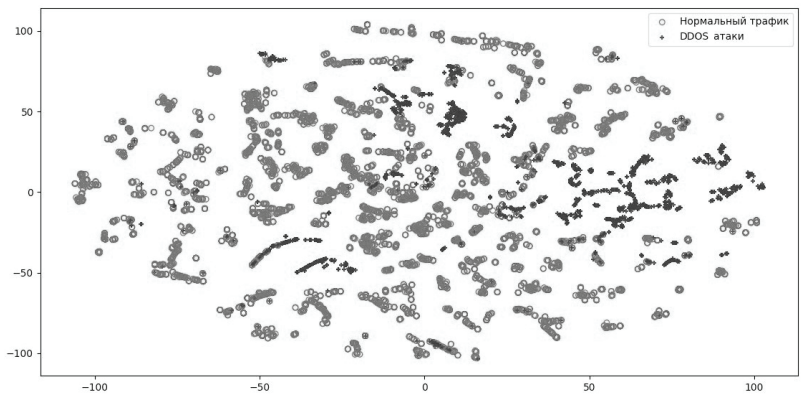

Рис. 2. Визуальное представление выборки прецедентов после нормирования в [0;1]

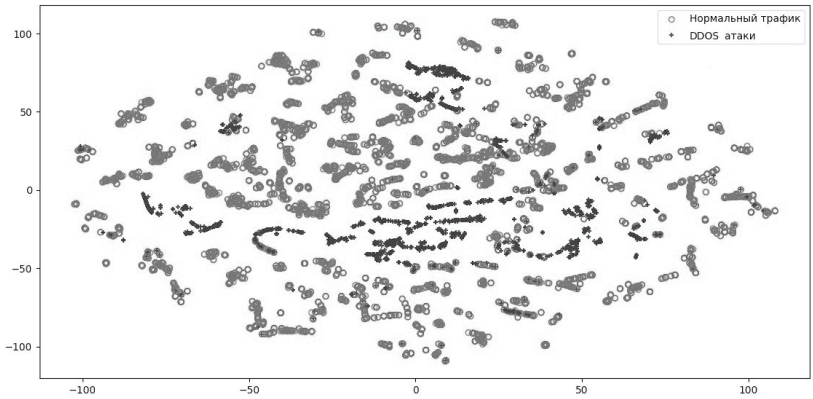

Рис. 3. Визуальное представление выборки прецедентов после нормирования по (14)

Эффективность формирования базы прецедентов оценивалась по тестовой выборке, состав который представлен в табл. 3. Смысл предобработки выборки сводился к удалению повторяющихся объектов.

Таблица 3

Состав тестовой выборки

№	*Тип DDOS атак*	*Количество объектов*	
		исходное	*после предобработки*
1	Benign	56 965	45 424
2	LDAP	1 915 122	17 894
3	MSSQL	5 787453	272 282
4	NetBIOS	3 657497	10 420
5	Portmap	186 960	1638
6	Syn	4 891 500	449 963
7	UDP	3 867 155	1 440 231
8	UDP_Lag	1873	455
	Итого	*20 364 525*	*2 238 307*

Далее будем использовать следующие обозначения для наборов признаков:

– $X(n)$ – исходный;
– $D(\tau), D(\tau) \subset X(n), < n$ – упорядоченный по значениям устойчивости (6);
– $X'(n)$ – нормированный по в [0;1];
– $D'(\tau), D'(\tau)\, X'(n), \subset < n$ – упорядоченный по значениям устойчивости (6);
– $F(n)$ – нормированный по (14);
– $T(\tau), T(\tau) \subset F(n), \tau < n$ – упорядоченный по значениям устойчивости (6).

Для тестирования использовался весь набор исходных признаков и 40 информативных в 3-х вариантах: без нормирования; с нормированием в [0;1] и по (14). В число информативных включены 40 признаков из упорядоченной по убыванию последовательности значений устойчивости (6) для всей базы прецедентов.

Результаты распознавания по алгоритмам *k* ближайших соседей (KNN), Random Forest и SVM [Scikit-learn 2019] приводятся в табл. 4, 5, 6.

Таблица 4

Результаты распознавания
по алгоритму *k* ближайших соседей (KNN)

№	*Тип DDOS-атак*	*Наборы признаков*					
		$X(n)$	$D(\tau)$	$X'(n)$	$D'(\tau)$	$F(n)$	$T(\tau)$
1	Benign	0.9587	0.9569	0.9542	0.9565	0.9542	0.9565
2	LDAP	0.8687	0.8713	0.9953	0.9951	0.9953	0.9951
3	MSSQL	0.7652	0.7772	0.9987	0.9989	0.9987	0.9989
4	NetBIOS	0.9113	0.9199	0.9479	0.9249	0.9478	0.9247
5	Portmap	0.7131	0.7198	0.8846	0.84432	0.8846	0.8443
6	Syn	0.7174	0.8311	0.9614	0.9911	0.9614	0.9911
7	UDP	0.9931	0.9935	0.9998	0.9997	0.9998	0.9997
8	UDP_Lag	0.6989	0.7077	0.8527	0.85055	0.8527	0.8505

Результаты эксперимента из табл. 4 указывают на зависимость алгоритма KNN от размерности и нормирования признакового пространства. Снижение размерности приводит к увеличению обобщающей способности алгоритма.

Таблица 5

Результаты распознавания
по алгоритму Random Forest (RM)

№	*Тип DDOS-атак*	*Наборы признаков*					
		$X(n)$	$D(\tau)$	$X'(n)$	$D'(\tau)$	$F(n)$	$T(\tau)$
1	Benign	0.9872	0.9926	0.9894	0.9901	0.9909	0.9908
2	LDAP	0.9918	0.9927	0.9915	0.9919	0.9921	0.9917
3	MSSQL	0.9991	0.9991	0.9990	0.9990	0.9991	0.9991
4	NetBIOS	0.9730	0.9879	0.9746	0.9753	0.9914	0.9752
5	Portmap	0.7552	0.8449	0.7503	0.7527	0.8486	0.7521
6	Syn	0.8795	0.9902	0.4536	0.7534	0.6627	0.1857
7	UDP	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
8	UDP_Lag	0.7341	0.7604	0.7362	0.7363	0.7626	0.7406

Особенности реализации алгоритма Random Forest (см. табл. 5) проявляются в отсутствии чувствительности результатов от размерности и способов нормирования признакового пространства.

Таблица 6

Результаты распознавания по алгоритму SVM

№	*Тип DDOS-атак*	*Наборы признаков*					
		$X(n)$	$D(\tau)$	$X'(n)$	$D'(\tau)$	$F(n)$	$T(\tau)$
1	Benign	0.9964	0.9887	0.9965	0.9887	0.9964	0.9887
2	LDAP	0.9823	0.9824	0.9823	0.9802	0.9823	0.9824
3	MSSQL	0.9983	0.9983	0.9983	0.9983	0.9983	0.9983
4	NetBIOS	0.9627	0.9538	0.9626	0.9538	0.9627	0.9538
5	Portmap	0.6996	0.7021	0.6996	0.7021	0.6996	0.7021
6	Syn	0.8941	0.6462	0.8888	0.6462	0.8941	0.6462
7	UDP	0.9999	0.9998	0.9998	0.9998	0.9998	0.9998
8	UDP_Lag	0.6395	0.6549	0.6396	0.6549	0.6396	0.6549

Эффект по точности распознавания (см. табл. 6) алгоритм SVM получен без снижения размерности пространства. Практической пользы от нормирования данных нет.

Заключение

Разработана методика анализа структуры отношений типов DDOS-атак с использованием значений устойчивости признаков по парам классов «нормальный трафик» и «тип DDOS-атак». Показана эффективность выбора латентного признакового пространства и мер расстояния между объектами в нем для решения задач Big Data.

Предложены методика формирования латентного признакового пространства с меньшей размерностью, чем исходное, и исследование структуры отношений объектов в нем. Результатам исследования является база данных прецедентов для обнаружения DDOS-атак.

Литература

Игнатьев 2011 – *Игнатьев Н.А.* Вычисление обобщенных показателей и интеллектуальный анализ данных // Автоматика и телемеханика. 2011. № 5. С. 183–190.

Игнатьев, Рахимова 2021 – *Игнатьев Н.А., Рахимова М.А.* Формирование и анализ наборов информативных признаков объектов по парам классов // Искусственный интеллект и принятие решений. 2021. № 4. С. 18–26. DOI: http://dx.doi.org/10.14357/20718594210402.

Мирзаев 2021 – *Мирзаев А.И.* О выборе пространства для описания объектов при машинном обучении на больших выборках данных // Проблемы вычислительной и прикладной математики. 2021. № 6 (36). С. 120–127.

Саидов 2017 – *Саидов Д.Ю.* Информационные модели на основе нелинейных преобразований признакового пространства в задачах распознавания: Дис. ... д-ра философии по физ.-мат. наукам. Ташкент, 2017.

Ignatyev 2018 – Ignatyev *N.A.* Structure Choice for Relations between Objects in Metric Classification Algorithms // Pattern Recognition and Image Analysis. 2018. Vol. 28 (4). P. 590–597. DOI: https://doi.org/10.1134/S1054661818040132.

Ignatyev 2021 – Ignatyev *N.A.* On Nonlinear Transformations of Features Based on the Functions of Objects Belonging to Classes // Pattern Recognition and Image Analysist. 2021. Vol. 31 (2). P. 197–204. DOI: http://dx.doi.org/10.1134/S1054661821020085.

Sharafaldin, Lashkari, Hakak, Ghorbani 2019 – *Sharafaldin I., Lashkari A.H., Hakak S., Ghorbani A.A.* Developing Realistic Distributed Denial of Service (DDOS) Attack Dataset and Taxonomy // The IEEE (53rd) International Carnahan Conference on Security Technology (ICCST2019), Chennai, India, October 1–3, 2019. New York, NY: IEEE, 2019. P. 1–8. DOI: http://dx.doi.org/10.1109/CCST.2019.8888419.

Scikit-learn 2019 – User guide released 0.21.3, 2019 // Scikit-learn. URL: https://scikit-learn.org/ (дата обращения 6 июня 2022).

References

Ignatyev, N.A. (2011), "Calculations of generalized indices and the intelligent data analysis", *Automation and Remote Control*, vol. 5, pp. 183–190.

Ignatyev, N.A. (2018), "Structure Choice for Relations between Objects in Metric Classification Algorithms", *Pattern Recognition and Image Analysis,* vol. 28, no. 4, pp. 590–597. DOI: https://doi.org/10.1134/S1054661818040132.

Ignatyev, N.A. (2021), "On Nonlinear Transformations of Features Based on the Functions of Objects Belonging to Classes", *Pattern Recognition and Image Analysist*, vol. 31, no. 2, pp. 197–204. DOI: http://dx.doi.org/10.1134/S1054661821020085.

Ignatyev, N.A. and Rakhimova, M.A. (2021), "Formation and analysis of sets of informative features of objects by pairs of classes", Artificial Intelligence and Decision Making, vol. 4, pp. 18–26, DOI: http://dx.doi.org/10.14357/20718594210402.

Mirzayev, A.I. (2021), "On the choice of space for describing objects in machine learning on large data samples", *Problems of computational and applied mathematics*, vol. 6 (36), pp. 120–127.

Saidov, D.Yu. (2017), *Informatsionnie modeli na osnove nelineynykh preobrazovaniy priznakovogo prostranstva v zadachax raspoznavaniya* [Information Models Based on Nonlinear Transformations of Feature Space in Recognition Problems], Ph.D. Thesis, Tashkent, Uzbekistan.

Scikit-learn (2019), User guide released 0.21.3, 2019, available at: URL: https://scikit-learn.org/ (Accessed 6 June 2022).

Sharafaldin, I., Lashkari, A.H., Hakak, S. and Ghorbani, A.A. (2019), "Developing Realistic Distributed Denial of Service (DDOS) Attack Dataset and Taxonomy", *The IEEE (53rd) International Carnahan Conference on Security Technology (ICCST2019)*, Chennai, India, October 1–3, 2019, IEEE, New York, NY, USA, pp. 1–8. DOI: http://dx.doi.org/10.1109/CCST.2019.8888419.

Информация об авторе

Эркин Р. Наврузов, докторант, Национальный университет Узбекистана им. Мирзо Улугбека, Ташкент, Республика Узбекистан; 100174, Республика Узбекистан, Ташкент, ул. Университетская, д. 4; erkinbek0989@gmail.com

Information about the author

Erkin R. Navruzov, doctoral student, Mirzo Ulugbek National University of Uzbekistan, Tashkent, Uzbekistan; bld. 4, University str., Tashkent, Uzbekistan, 100174; erkinbek0989@gmail.com

УДК 004.056.53
DOI: 10.28995/2686-679X-2022-3-85-95

Роль профайлинга в обеспечении информационной безопасности

Ирина А. Русецкая
Российский государственный гуманитарный университет, Москва, Россия, irkom@mail.ru

Аннотация. Статья посвящена анализу места и значения профайлинга при решении проблем обеспечения информационной безопасности. Рассматриваются понятие и истоки возникновения профайлинга, основные его составляющие. На фоне выделения таких сфер применения профайлинга в современной практике, как криминалистическая, кадровая и коммерческая, рассматривается роль профайлинга в области информационной безопасности. Проводится анализ задач, которые могут помочь решить использование методик профайлинга, а также инструментарий, применяемый для этого. Автор анализирует возможные источники получения данных для составления профайла в коммерческих организациях. В статье рассматриваются вопросы, ответы на которые могут получить сотрудники служб информационной безопасности компаний, воспользовавшись методиками профайлинга. Особое внимание автор уделяет автоматизированным методикам профайлинга, использующимся сегодня и в ближайшей перспективе. Рассматриваются примеры возможных критериев оценки пользовательского поведения при применении автоматизированных методик профайлинга. В статье анализируются преимущества автоматизированного профайлинга перед традиционным, примеры существующих в этой области практических программных разработок, а также основные проблемы, возникающие при разработке и внедрении в практику коммерческих и государственных организаций автоматизированных методик профайлинга, которые помогают решать проблемы обеспечения информационной безопасности.

Ключевые слова: профайлинг, информационная безопасность, DLP-системы, поведенческий анализ, профайл

Для цитирования: *Русецкая И.А.* Роль профайлинга в обеспечении информационной безопасности // Вестник РГГУ. Серия «Информатика. Информационная безопасность. Математика». 2022. № 3. С. 85–95. DOI: 10.28995/2686-679X-2022-3-85-95

The role of profiling in ensuring information security

Irina A. Rusetskaya
Russian State University for the Humanities, Moscow, Russia, irkom@mail.ru

Abstract. The article deals with the analysis of the place and importance of profiling in solving the issues of ensuring information security. It considers the concept and origins of profiling and its main components. The role of profiling in the field of information security is considered against the background of emphasizing of such areas of the profiling application in modern practice as the forensic, personnel and commercial. The analysis of tasks that can help to solve the use of profiling techniques, as well as the tools used for that, is carried out. The author analyzes possible sources of obtaining data for compiling a profile in commercial organizations. The article discusses questions that can be answered by employees of the information security services of companies using profiling techniques. The author also pays special attention to automated profiling techniques used today and supposed to be used in the near future. Examples of possible criteria for evaluating the user behavior when applying automated profiling techniques are considered. The article analyzes the advantages of automated profiling over traditional profiling, examples of practical software developments existing in that area, as well as the main issues that arise in the development and implementation of automated profiling techniques in the practice of commercial and government organizations that help in solving the information security issues.

Keywords: profiling, information security, DLP systems, behavioral analysis, profile

For citation: Rusetskaya, I.A. (2022), "The role of profiling in ensuring information security", *RSUH/RGGU Bulletin. "Information Science. Information Security. Mathematics" Series*, no. 3, pp. 85–95, DOI: 10.28995/2686-679X-2022-3-85-95

Введение

Анализ инцидентов информационной безопасности в настоящее время показывает, что, несмотря на активное и все больше расширяющееся применение средств информационных технологий, одним из важнейших факторов уязвимости является так называемый человеческий фактор. Он предполагает, что едва ли не подавляющее большинство угроз защищаемой информации

исходит от людей, являющихся участниками информационных процессов и субъектами информационной среды. Так, если речь идет о коммерческих организациях, то это сотрудники, клиенты, конкуренты, представители сторонних фирм, взаимодействующие в информационном поле, и др. Обязательной составляющей обеспечения безопасности организации является работа с сотрудниками на всех этапах: при подборе и приеме персонала, текущей работе, а также при увольнении, а иногда и после увольнения работников. При этом необходимо учитывать, что негативное воздействие на защищаемую информацию может быть как умышленным, злонамеренным, так и случайным, ошибочным.

Одним из направлений работы с персоналом на всех трех выделенных выше этапах является профайлинг, под которым можно понимать применение совокупности методов оценки вербальных и невербальных высказываний, а также прогнозирование поведения человека.

Целью данной работы является анализ современных подходов к использованию профайлинга для решения задач обеспечения информационной безопасности организаций.

Место и значение профайлинга при решении проблем обеспечения информационной безопасности

Идеи, лежащие в основе профайлинга, возникли не одно столетие назад. Изначально они не были оформлены в качестве научной дисциплины и базировались на околонаучных методах, как, например, популярная в XVIII–XIX вв. в Европе и России физиогномика. Методики профайлинга начали формироваться психиатрами и криминалистами, среди которых одним из самых известных является Чезаре Ломброзо (1835–1909 гг.), предложивший рекомендации по определению врожденной предрасположенности к преступлениям на основе анализа некоторых внешних признаков человека.

Дальнейшее научное развитие профайлинга также связано с областями психиатрии, психологии и криминалистики; в рамках исследований в этих областях в 1970-х гг. появился сам термин «профайлинг».

Методы профайлинга начали интересовать специалистов служб, обеспечивающих общественную, транспортную безопасность, борьбу с терроризмом и пр. Так, профайлинг прочно вошел в арсенал средств работы сотрудников и экспертов правоохрани-

тельных органов и спецслужб в рамках так называемого криминалистического профайлинга.

Сегодня методики профайлинга используются в различных видах деятельности. Принято выделять три ставших традиционными сферы его практического использования:

– криминалистическая;
– кадровая;
– коммерческая.

Про криминалистический профайлинг было сказано выше.

Кадровый профайлинг предполагает использование соответствующих методов при подборе персонала компаний, расстановке кадров и работе с ними.

Профайлинг в коммерческой сфере проявляет себя в различных направлениях деятельности: банковской, страховой, маркетинга, бизнес-коммуникаций и т. п., где важно провести оценку надежности участников коммерческих отношений.

В данном исследовании будет рассматриваться тема профайлинга в сфере информационной безопасности, которой в последние годы исследователями уделяется все большее внимание [Арутюнов 2022].

Данное направление развивается на стыке кадрового и коммерческого, а также отчасти криминалистического профайлинга, так как включает в себя как задачи работы с персоналом, связанные с обеспечением информационной безопасности, оценку рисков информационной безопасности, возникающих в процессе функционирования коммерческих организаций различных направлений деятельности по вине «человеческого фактора», так и выявление мошеннических действий.

В рамках обеспечения информационной безопасности коммерческих организаций использование профайлинга позволяет решать следующие задачи:

– оценивать уровень надежности и лояльности сотрудников и клиентов с точки зрения обеспечения информационной безопасности компании;
– верифицировать достоверность предоставленной информации;
– оценить степень соблюдения сотрудником требований политики информационной безопасности, принятой в организации;
– прогнозировать поведение сотрудников и клиентов, которое может влиять на целостность, доступность и конфиденциальность информации, которой они обладают;
– выявить сильные и слабые стороны поведения сотрудников, влияющие на информационную безопасность компании;

– оказывать помощь в расследовании возникших инцидентов в области информационной безопасности, а также осуществлять их профилактику;
– определять и оценивать риски в области информационной безопасности, связанные с «человеческим фактором».

Инструменты профайлинга, используемые в сфере обеспечения информационной безопасности, включают в себя:
– анализ речи, устной и письменной;
– анализ мимики, жестов, интонаций, тембра голоса и прочих невербальных характеристик речи;
– выявление типа характера;
– анализ стратегий мышления и поведения.

Для составления профайлинга в качестве источников информации используются следующие типы данных:
– данные письменных опросов и анкетирования;
– данные устных собеседований;
– данные документов;
– данные из открытых источников;
– данные, полученные при использовании в организации автоматизированных модулей (DLP-систем).

Результатом использования методов профайлинга может стать получение сотрудниками службы информационной безопасности ответов на следующие вопросы:
– стоит ли предоставлять конкретному сотруднику доступ к конфиденциальной информации;
– какие изменения в поведении сотрудника стоит считать критичными в плане вероятности возникновения инцидента в области информационной безопасности;
– какие индивидуальные меры профилактики (проведение бесед, обучения, проверок и пр.) инцидентов стоит применять для конкретного работника;
– как оценить риски информационной безопасности, исходящие от конкретного работника;
– кто из работников стал причиной возникновения инцидента или его участником;
– каков социально-психологический климат в коллективе и каковы риски его влияния на корпоративную безопасность и т. д.

Развитие цифровых технологий, возможности компьютерной обработки больших объемов информации и формализованных подходов к обеспечению информационной безопасности определило автоматизацию решения многих задач в этой сфере, начиная от криптографии и заканчивая профайлингом [Русецкая 2021].

Помимо классических методик профайлинга, используемых экспертами-профайлерами, в настоящее время все более актуальными становятся автоматизированные методики.

Одним из первых подходов к автоматизации решения такой задачи профайлинга, как верификация информации, стало создание полиграфа в различных его вариантах. Создание полиграфа как инструмента для «детекции лжи», а также оценка достоверности его показаний имеют долгую историю, которая может стать темой отдельного большого исследования.

Современные методики профайлинга предполагают использование автоматизированных методов контроля содержания, контекста, составление шаблонов регулярных выражений и «цифровых отпечатков» («отпечатков браузера») [Пучков 2017].

К примерам критериев оценки пользовательского поведения при применении автоматизированных методик профайлинга могут относиться следующие:

– выбор надежных паролей и их смена в соответствии с требованиями политики безопасности компании;
– данные о попытках несанкционированного доступа к конфиденциальной информации;
– анализ запросов к установленным средствам защиты;
– анализ участия в зарегистрированных инцидентах информационной безопасности [Муравьев 2018];
– анализ оценки изменения рабочего поведения;
– выявление в тексте определенных слов или словосочетаний и др.

Автоматизированный профайлинг имеет следующие преимущества перед традиционным:

– экономия времени и кадровых ресурсов;
– возможность постоянного мониторинга изменений ключевых показателей оценки поведения сотрудников в режиме реального времени;
– возможность дистанционного мониторинга вне зависимости от местонахождения сотрудника во время рабочего дня, что особенно актуально в условиях удаленной работы;
– устранение фактора субъективности при оценке поведения работников экспертами-профайлерами.

Таким образом, внедрение средств автоматизации профайлинга может быть эффективным инструментом решения ряда задач обеспечения информационной безопасности фирмы.

Исследования и практические разработки в этом направлении активно ведутся в настоящее время. Так, российские IT-разработчики в 2018 г. выпустили модуль «КИБ СёрчИнформ

ProfileCenter», который представляет собой особый компонент, позволяющий автоматизировать профайлинг для предотвращения угроз, связанных с действиями внутренних нарушителей [Бируля 2018]. В качестве научной основы при создании модуля использовались методы психолингвистики.

В качестве источника данных используются неформальные тексты, взятые из переписки сотрудников по корпоративной почте, в корпоративных мессенджерах, посты, оставленные с рабочей станции в социальных сетях, которые обрабатываются в DLP-системе.

Однако автоматизация профайлинга сталкивается с рядом сложностей.

Одной из основных является необходимость использования легально полученного контента.

С одной стороны, необходимо учитывать требования законодательства по сохранению неприкосновенности частной жизни, с другой стороны, согласно установленным в РФ правовым нормам, в частности Трудового и Гражданского кодексов РФ, работодатель имеет право на контроль переписки и переговоров сотрудников в целях защиты информации, обладателем которой является. Просмотр контента, полученного с предоставленных работникам аккаунтов, должен осуществляться гласно, с письменного разрешения работника и фиксироваться в нормативно-методических регламентах компании. Сотрудникам фирмы не следует использовать во время делового общения как конфиденциальную информацию организации, если это использование не является санкционированным, так и информацию личного характера [Кундышева, Русецкая 2019].

Второй сложностью является необходимость использования для создания полноценного психологического портрета человека анализа его разноплановых характеристик: мимики, голоса, трафика, «клавиатурного почерка» и т. д. Задачи автоматизации анализа подобных характеристик в комплексе решаются достаточно сложно, так как в отличие от задач традиционного профайлинга должны быть основаны на нетестовых методиках и относятся к числу слабо формализуемых задач[1].

Для автоматизации отдельных элементов профайлинга предлагаются различные практические решения. Так, возможностями распознавания эмоций и оценки голосовых модуляций обладают

[1] *Филатов А.* Пять стадий принятия неизбежного, или как мы разрабатывали программу для автоматизированного профайлинга // Habr, 17 апреля 2020. URL: https://habr.com/ru/company/searchinform/blog/497814/?ysclid=l5b4kizfhw288071069 (дата обращения 15 июля 2022).

облачная платформа Microsoft Azure, когнитивная система IBM Watson [Бируля 2017].

Проблемы обработки визуальной информации, распознавания образов, а также оценки эмоционального состояния на основе анализа видеоизображения лиц решают, например, продукты FaceReader и eMotion Software нидерландских компаний Noldus Information Technology и Visual Recognition соответственно, программные комплексы FaceSecurity и MMER_FEASy – the FacE Analysis System немецких компаний – разработчиков Cognitec и MMER-Systems, продукты Affective Computing Research Group (США) и др.

С точки зрения обеспечения информационной безопасности такие системы могут выполнять, например, следующие задачи: идентификация и аутентификация лиц в целях контроля доступа, отслеживание передвижений лиц с помощью систем видеонаблюдения и т. д.

Подобные программные решения могут использоваться, в том числе, для обеспечения информационной безопасности как коммерческими компаниями, так и государственными службами.

Так, власти Москвы с начала 2020 г. используют систему видеонаблюдения с функцией распознавания лиц, разработанную российской компанией NtechLab, которая позволяет установить личность человека, его пол и примерный возраст [Фокс-Брюстер 2020].

Реализуется также идея автоматизации речевого профайлинга, который предполагает зачитывание контрольного текста испытуемыми и анализ акустических и фонетических свойств речи и их зависимости от текущего психоэмоционального состояния [Савченко, Акатьев 2017].

Третья проблема заключается в недостаточности для проведения поведенческого анализа сотрудников материала, предоставляемого различными системами. Например, для анализа бывает недостаточно данных мониторинга или DLP-систем в случае, если репрезентативно представлены только деловые элементы переписки и письменных переговоров, которые не дают возможности делать выводы о ценностях, настроении, психологическом состоянии, личностных качествах сотрудника и т. п.

Заключение

Итак, в данной статье были выделены основные идеи, лежащие в основе применения профайлинга для нужд обеспечения информационной безопасности, связанные с решением проблем, обусловленных наличием так называемого человеческого фактора,

учет которого является одним из основополагающих для защиты информации.

Инструменты профайлинга на основе анализа различных типов данных, таких, например, как данные представленных документов, анкет, устных бесед с сотрудниками, данные, полученные из Interneta, и данные используемых DLP-систем, могут позволить сотрудникам служб информационной безопасности решать важные задачи обеспечения защиты информации. К числу таких задач относятся оценка степени надежности работников, достоверности предоставляемой ими информации, степени соблюдения норм и регламентов по защите информации, рисков, определяемых «человеческим фактором», а также прогнозирование поведения работников, которое может негативно повлиять на состояние безопасности информации организации.

В настоящее время особый интерес вызывают решения, позволяющие автоматизировать методики профайлинга, применяемые в сфере информационной безопасности, которые могут позволить экономить время сотрудников служб информационной безопасности при помощи непрерывного дистанционного мониторинга важных критериев оценки поведения сотрудников, а также корректировать субъективную оценку поведения работников, которую дают эксперты-профайлеры.

Несмотря на описанные в статье этические, юридические, организационные и технические сложности, которые могут возникать при внедрении в организациях элементов автоматизированного профайлинга, разработка автоматизированных систем профайлинга и их применение в сфере информационной безопасности представляется эффективным. Свидетельством этого являются используемые сегодня практические программные решения, делающие возможным поведенческий анализ на основе изучения текстовых сообщений, распознавания эмоционального состояния и оценки голосовых изменений на основе анализа голосовых сообщений и видеоизображений, методик автоматизации речевого профайлинга и т. д.

Литература

Арутюнов 2022 – *Арутюнов В.В.* Особенности кластера знаний о результативности и востребованности итогов исследований в области профайлинга // Информационная безопасность: вчера, сегодня, завтра: Сб. ст. по материалам Междунар. научно-практич. конф. Москва, 14 апреля 2022 г. М.: РГГУ, 2019. С. 7–13.

Бируля 2018 – *Бируля И.* Нетехнические методы защиты информации: профайлинг на службе ИБ. 19 апреля 2018 // Anti-Malware Journal. URL: https://www.anti-malware.ru/practice/methods/information-security-profiling (дата обращения 15 июля 2022).

Бируля 2017 – *Бируля И.* Обзор технологий профайлинга // Журнал RUБЕЖ, 17 октября 2017. URL: https://ru-bezh.ru/ivan-birulya/18589-obzor-texnologij-profajlinga?ysclid=l5b4swo9q8362330709 (дата обращения 13 июля 2022).

Кундышева, Русецкая 2019 – *Кундышева И.Р., Русецкая И.А.* Правовые аспекты использования DLP -систем в организациях // Информационная безопасность: вчера, сегодня, завтра: Сб. ст. по материалам Междунар. научно-практич. конф. Москва, 23 апреля 2019 г. М.: РГГУ, 2019. С. 175–180.

Муравьев 2018 – *Муравьев Н.С.* Профилактика инцидентов информационной безопасности на основе профилирования пользователей: программно-технический аспект // Вестник УрФО. Безопасность в информационной сфере. 2018. № 1 (27). С. 66–70. URL: http://www.elibrary.ru/item.asp?id=35214406&ysclid=l5jfd2ixxq986689247 (дата обращения 17 июля 2022).

Пучков 2017 – *Пучков И.И.* Коммерческий профайлинг в DLP-системах // Молодой ученый. 2017. № 51 (185). С. 75–77. URL: https://moluch.ru/archive/185/47448/ (дата обращения 13 июля 2022).

Русецкая 2021 – *Русецкая И.А.* Криптография: от прошлого к будущему // Вестник РГГУ. Серия «Информатика. Информационная безопасность. Математика». 2021. № 4. С. 47–57.

Савченко, Акатьев 2017 – *Савченко В.В., Акатьев Д.Ю.* Информационная технология речевого профайлинга // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. 2017. № 9 (258). С. 157–165. URL: https://elibrary.ru/item.asp?id=29810974 (дата обращения 13 июля 2022).

Фокс-Брюстер 2020 – *Фокс-Брюстер Т.* Как Москва получила за $3,2 млн лучшую в мире систему распознавания лиц // Forbes, 01 февраля 2020 г. URL: https://www.forbes.ru/tehnologii/392303-kak-moskva-poluchila-za-32-mln-luchshuyu-v-mire-sistemu-raspoznavaniya-lic?ysclid=l5uwmq9w2f282567852 (дата обращения 21 июля 2022).

References

Arutyunov, V.V. (2022), “Features of the knowledge cluster about the effectiveness and demand for research results in the field of profiling”, *Information security: yesterday, today, tomorrow: Sat. Art. according to Proceedings of the International scientific and practical. conf*., Moscow, April 14, 2022. RGGU, Moscow, Russia, pp. 7–13.

Birulya, I. (2018), “Non-technical methods of information security. Profiling in the service of information security”, April 19, 2018, *Anti-Malware Journal*, available at: https://www.anti-malware.ru/practice/methods/information-security-profiling (Accessed 15 July 2022).

Birulya, I. (2017), "Overview of profiling technologies", *Journal RUBEZH*, October 17, 2017, available at: https://ru-bezh.ru/ivan-birulya/18589-obzor-texnologij-profajlinga?ysclid=l5b4swo9q8362330709 (Accessed 13 July 2022).

Kundysheva, I.R. and Rusetskaya, I.A. (2019), "Legal aspects of using DLP systems in organizations", *Information security: yesterday, today, tomorrow: Sat. Art. according to Proceedings of the International scientific and practical. conf.*, Moscow, April 23, 2019. RGGU, Moscow, Russia, pp. 175–180.

Murav'ev, N.S. (2018), "Prevention of the information security incidents based on user profiling. Software and hardware aspect", *Bulletin of the Ural Federal District. Security in the information sphere*, 2018, no. 1 (27), pp. 66–70, available at: www.elibrary.ru/item.asp?id=35214406&ysclid=l5jfd2ixxq986689247 (Accessed 17 July 2022).

Puchkov, I.I. (2017), "Commercial profiling in DLP systems", *Young scientist*, 2017, no. 51 (185), pp. 75–77, available at: https://moluch.ru/archive/185/47448/ (Accessed 13 July 2022).

Rusetskaya, I.A. (2021), "Cryptography. From the past to the future", *RSUH/RGGU Bulletin. "Information Science. Information Security. Mathematics" Series*, Moscow, vol. 4, pp. 47–57.

Savchenko, V.V. and Akat'ev, D.Yu (2017), "Information technology of speech profiling", *Scientific Bulletin of the Belgorod State University. Series: Economy. Informatics*, no. 9 (258), pp. 157–165, available at: https://elibrary.ru/item.asp?id=29810974 (Accessed 13 July 2022).

Fox Brewster, T. (2020), "How Moscow got the world's best facial recognition system for $3.2 million", *Forbes,* February 01, 2020, available at: https://www.forbes.ru/tehnologii/392303-kak-moskva-poluchila-za-32-mln-luchshuyu-v-mire-sistemu-raspoznavaniya-lic?ysclid=l5uwmq9w2f282567852 (Accessed 21 July 2022).

Информация об авторе

Ирина А. Русецкая, кандидат исторических наук, доцент, Российский государственный гуманитарный университет, Москва, Россия; 125047, Россия, Москва, Миусская пл., д. 6; irkom@mail.ru

Information about the author

Irina A. Rusetskaya, Cand. of Sci. (History), associate professor, Russian State University for the Humanities, Moscow, Russia; bld. 6, Miusskaya Sq., Moscow, Russia, 125047; irkom@mail.ru

УДК 519.2
DOI: 10.28995/2686-679X-2022-3-96-120

Эмпирическое исследование мощности статистического критерия Колмогорова–Смирнова в задачах проверки гипотез о законе распределения

Вячеслав Ю. Синицын
Российский государственный гуманитарный университет, Москва, Россия, fpmrggu@yandex.ru

Екатерина С. Ступакова
Российский государственный гуманитарный университет, Москва, Россия, stupakova13@gmail.com

Аннотация. Статистические критерии, которые используют одновыборочную статистику А.Н. Колмогорова и двухвыборочную статистику Н.В. Смирнова, почти сто лет широко применяются для решения прикладных задач. Эти критерии представлены во многих учебниках и реализованы в компьютерных программных средствах для анализа данных. Цель данной работы состоит в том, чтобы при помощи различных методов перестройки выборочных данных эмпирически оценить мощность критерия Колмогорова–Смирнова на наборе тестовых задач по проверке гипотез о законе распределения, а также исследовать свойства оценок мощности критерия при различных типах ресамплинга. Для получения различных оценок мощности критерия Колмогорова–Смирнова при решении тестовых задач применялись классический бутстреп, непараметрический бутстреп, параметрический бутстреп, бутстреп со случайным слагаемым и ресамплинг без возвращений. По результатам многократного решения тестовых задач были вычислены медианы p-значений, медианы оценок мощности критерия Колмогорова–Смирнова, а также были найдены медианы смещений оценок на тестовых задачах и медианы всех попарных разностей различных оценок мощности критерия. Были исследованы законы распределения для рассмотренных оценок мощности критерия Колмогорова–Смирнова, для смещений оценок мощности и всех попарных разностей оценок мощности. Для тестовых задач, где рассматривались реальные данные, вычисление оценок мощности без ресамплинга невозможно, но несмещенные оценки мощности можно прогнозировать как полусумму двух других оценок мощности,

одна из которых получена при помощи непараметрического бутстрепа, а другая при помощи ресамплинга без возвращений. Методы перестройки выборочных данных и программные средства оценивания мощности статистического критерия, разработанные в рамках данного исследования, являются универсальными и могут быть полезны для анализа свойств оценок мощности других одновыборочных статистических критериев, а также для развития классической методологии проверки статистических гипотез.

Ключевые слова: проверка статистических гипотез, статистический критерий Колмогорова–Смирнова, мощность статистического критерия, ресамплинг

Для цитирования: Синицын В.Ю., Ступакова Е.С. Эмпирическое исследование мощности статистического критерия Колмогорова–Смирнова в задачах проверки гипотез о законе распределения // Вестник РГГУ. Серия «Информатика. Информационная безопасность. Математика». 2022. № 3. С. 96–120. DOI: 10.28995/2686-679X-2022-3-96-120

Empirical study of the power of the Kolmogorov–Smirnov statistical test in problems of testing hypotheses about the distribution law

Vyacheslav Yu. Sinitsyn
Russian State University for the Humanities, Moscow, Russia, fpmrggu@yandex.ru

Ekaterina S. Stupakova
Russian State University for the Humanities, Moscow, Russia, stupakova13@gmail.com

Abstract. Statistical tests that use A.N. Kolmogorov one-sample statistics and N.V. Smirnov two-sample statistics have been widely used for solving applied problems for almost a hundred years. Those criteria are present in many textbooks and implemented in computer software for data analysis. The purpose of the work is to empirically estimate the power of the Kolmogorov–Smirnov criterion on a set of test problems to check hypotheses about the distribution law, as well as to investigate the properties of estimates of the power of the criterion for various types of resampling. To obtain various estimates of the power of the Kolmogorov–Smirnov criterion in solving test problems, the classical bootstrap, nonparametric bootstrap, parametric bootstrap, bootstrap with a random term, and resampling without returns were used. Based on the results of multiple solution of test problems, medians of

estimates of the power of the Kolmogorov–Smirnov criterion were calculated, and medians of biases of estimates on test problems and medians of all pairwise differences of various estimates of the power of the criterion were found. The distribution laws for the considered Kolmogorov–Smirnov power estimates, for the biases of the power estimates, and for all pairwise differences of the power estimates were investigated. For test problems where real data were considered, calculating power estimates without resampling is impossible, but unbiased power estimates can be predicted as the half-sum of two other power estimates, one obtained with a nonparametric bootstrap and the other obtained with resampling without returns. Methods for restructuring sample data and software for estimating the power of a statistical test developed within the study are universal and can be used for analyzing the properties of power estimates of other one-sample statistical tests, as well as for developing the classical methodology for checking statistical hypotheses.

Keywords: statistical hypotheses testing, Kolmogorov–Smirnov statistical test, statistical test power, resampling

For citation: Sinitsyn, V.Yu. and Stupakova, E.S. (2022), "Empirical study of the power of the Kolmogorov–Smirnov statistical test in problems of testing hypotheses about the distribution law", *RSUH/RGGU Bulletin. "Information Science. Information Security. Mathematics" Series*, no. 3, pp. 96–120, DOI: 10.28995/2686-679X-2022-3-96-120

Введение

В 1933 г. А.Н. Колмогоров предложил знаменитый теперь статистический критерий для проверки согласия эмпирического распределения с заданным теоретическим законом распределения. А.Н. Колмогоров нашел и исследовал предельное распределение статистики критерия [Kolmogoroff 1933] и успешно применил свой критерий для подтверждения законов Менделя [Колмогоров 1940]. В 1939 г. Н.В. Смирнов предложил и подробно исследовал двухвыборочную статистику, которая применяется сейчас для проверки однородности двух независимых выборок [Смирнов 1939]. Упомянутые выше работы А.Н. Колмогорова и Н.В. Смирнова послужили важным этапом в развитии методов непараметрической математической статистики. Статистические критерии, которые используют одновыборочную статистику А.Н. Колмогорова и двухвыборочную статистику Н.В. Смирнова, широко применяются для решения прикладных задач. Эти критерии представлены во многих учебниках [Прохоров, Пономаренко 2019] [Кремер 2019] и реализованы в компьютерных программных средствах для ста-

тистического анализа данных SPSS, Statistica, R [Статистический анализ 2010] [Практикум по математической статистике 2021]. Традиционно в инструментах для компьютерной проверки статистических гипотез критерии А.Н. Колмогорова и Н.В. Смирнова бывают объединены в рамках одной вычислительной процедуры или функции под названием «критерий Колмогорова–Смирнова», что терминологически некорректно, но общепринято. В данном тексте мы используем традиционный термин «критерий Колмогорова–Смирнова», подразумевая реализацию в виде функции ks.test() в языке программирования R набора непараметрических критериев А.Н. Колмогорова, Н.В. Смирнова и вспомогательных алгоритмов применения соответствующих статистик, а также процедур статистического моделирования.

Цель данной работы состоит в том, чтобы при помощи различных методов перестройки выборочных данных эмпирически оценить мощность критерия Колмогорова–Смирнова на наборе тестовых задач по проверке гипотез о законе распределения, а также исследовать смещение оценок мощности критерия при различных типах ресамплинга.

Тестовые задачи

Для исследования мощности статистического критерия Колмогорова–Смирнова в данной работе рассматривались двенадцать тестовых задач по проверке статистических гипотез о нормальном законе распределения с фиксированными параметрами. В задачах 1–7, 11 данные, при помощи которых проверялась нулевая гипотеза, получались как случайная выборка из генеральной совокупности, не соответствующей гипотезе H0. Задачи 1–7, 11 формулировались таким образом, чтобы критерию Колмогорова было нелегко отвергнуть H0 по причине «близости» законов распределения данных и генеральной совокупности нулевой гипотезы. В задачах 8–10 использовались реальные, а не симулированные данные о 30 психологических характеристиках студентов РГГУ, полученные в рамках кросскультурных исследований [Воронов, Зайчикова, Синицын 2000] [Синицын, Кашпарова 2004]. Законы распределения данных, использованных в задачах 8–10, хорошо изучены [Практикум по математической статистике 2021]. В задаче 12 данные получались как случайная выборка из генеральной совокупности, соответствующей нулевой гипотезе.

Задача 1

Проверить нулевую гипотезу о том, что закон распределения генеральной совокупности не отличается от нормального с математическим ожиданием 172 и стандартным отклонением, равным 7. Кратко такую проверяемую гипотезу можно обозначить следующим образом H0: X ~ N(172, 7). Данные для проверки нулевой гипотезы представляют собой смесь из двух нормальных законов распределения с разными математическими ожиданиями 168 и 175 и одинаковыми стандартными отклонениями, равными 7. На языке R это можно представить кодом

```
x_data <- c(rnorm(100,mean=168,sd=7),
rnorm(200,mean=175,sd=7))
```

Задача является трудной для статистического критерия Колмогорова, так как законы распределения данных и генеральной совокупности нулевой гипотезы различаются мало.

Задача 2

Проверить нулевую гипотезу H0: X ~ N (172, 7). Данные для проверки нулевой гипотезы представляют собой смесь из двух нормальных законов распределения. Код для получения данных

```
x_data <- c(rnorm(100,mean=172,sd=5),
rnorm(200,mean=172,sd=11))
```

Задача является трудной для статистического критерия Колмогорова, так как законы распределения данных и генеральной совокупности нулевой гипотезы различаются мало.

Задача 3

Проверить нулевую гипотезу H0: X ~ N (172, 7). Данные для проверки нулевой гипотезы берутся из нормального закона распределения с параметрами 172 и 8 с последующим округлением до целых. Код для получения данных

```
x_data <- round(rnorm(300,mean=172,sd=8))
```

Задача 4

Проверить нулевую гипотезу H0: X ~ N (507, 15.8). Данные для проверки нулевой гипотезы берутся как случайная выборка объема 300 из биномиального закона с параметрами n = 1000 и p = 0.507. Код для получения данных

```
x_data <- rbinom(300,1000,0.507)
```

Задача является трудной для статистического критерия Колмогорова, так как согласно теореме Муавра–Лапласа, законы распределения данных и генеральной совокупности нулевой гипотезы различаются мало.

Задача 5

Проверить нулевую гипотезу H0: X ~ N(20, 6). Данные для проверки нулевой гипотезы берутся как случайная выборка объема 300 из закона распределения Хи-квадрат с 20 степенями свободы. Код для получения данных

```
x_data <- rchisq(300,20)
```

Задача является трудной для статистического критерия Колмогорова, так как, согласно Центральной предельной теореме, законы распределения данных и генеральной совокупности нулевой гипотезы различаются мало.

Задача 6

Проверить нулевую гипотезу H0: X ~ N(56, 12). Данные для проверки нулевой гипотезы берутся из логнормального закона распределения с параметрами 4 и 0.2. Код для получения данных

```
x_data <- rlnorm(300,4,0.2)
```

Задача 7

Проверить нулевую гипотезу H0: X ~ N (50, 7). Код для данных

```
x_data <- rpois(300,50)
```

Задача 8

Для каждой из 30 подшкал психометрического теста NEO PI-R [Воронов, Зайчикова, Синицын 2000] [Синицын, Кашпарова 2004] проверить нулевую гипотезу о том, что закон распределения генеральной совокупности не отличается от нормального закона с фиксированными параметрами m и sd. Для проверки нулевой гипотезы берутся данные для всех респондентов, а параметры m и sd равны среднему арифметическому и выборочному стандартному отклонению для данных (по каждой подшкале свои значения).

Задача 9

Аналогично задаче 8, но для проверки нулевой гипотезы берутся данные только для респондентов девушек.

Задача 10

Аналогично задаче 8, но данные берутся только для юношей.

Задача 11

Проверить нулевую гипотезу H0: X ~ N(172, 7). Данные для проверки нулевой гипотезы представляют собой случайную выборку с возвращениями объема 300 из фиксированной выборки x0 объема 300 из нормального закона с параметрами 172 и 7. Код для получения данных

```
x_data <- sample(x0,300,repl=TRUE)
```

Таким образом, данные представляют собой псевдовыборку, полученную при помощи классической бутстреп-процедуры из некоторой конкретной выборки, соответствующей нулевой гипотезе.

Задача 12

Проверить нулевую гипотезу H0: X ~ N(172, 7). Данные для проверки нулевой гипотезы представляют собой случайную выборку объема 300 из нормального закона с параметрами 172 и 7. Код для получения данных

```
x_data <- rnorm(300,mean=172,sd=7)
```

Таким образом, данные представляют собой случайную выборку объема 300, соответствующую нулевой гипотезе. Критерий Колмогорова, отвергая H0 в такой ситуации, совершает ошибку первого рода.

Решение тестовых задач

Однократное решение тестовых задач 1–7, 11 и 12 при помощи статистического критерия Колмогорова–Смирнова выполнялось следующим образом. Сначала в соответствии с условием задачи генерировалась случайная выборка x_data объема 300. На основании этого фиксированного набора данных при помощи критерия Колмогорова–Смирнова проверялась нулевая гипотеза и вычислялось p-значение. Находились пять различных оценок мощности критерия при различных типах перестройки данных и оценка мощности без ресамплинга. Вычислялись смещения оценок мощности критерия, полученных при помощи ресамплинга, по отношению к оценке мощности, найденной без ресамплинга. Вычислялись также

все попарные различия оценок мощности, полученных при помощи разных методов ресамплинга.

Для получения пяти различных оценок мощности критерия Колмогорова–Смирнова при решении тестовых задач применялись различные типы перестройки выборки x_data: классический бутстреп, непараметрический бутстреп, параметрический бутстреп, бутстреп со случайным слагаемым и ресамплинг без возвращений.

Классическая бутстреп-процедура, при которой из выборки случайно извлекались с возвращениями элементы в количестве, равном объему исходных данных, была реализована в виде функции r1_data.

При непараметрической бутстреп-процедуре выборочная квантильная функция, построенная по набору данных x_data и сглаженная кубическими сплайнами, применялась к элементам случайной выборки такого же объема, как x_data, взятой из равномерного закона распределения на отрезке [0;1]. Непараметрическая бутстреп-процедура была реализована в виде функции r2_data, которая обеспечивает практически совпадающие законы распределения для псевдовыборки r2_data(x_data) и исходной выборки x_data.

При параметрической бутстреп-процедуре для выборки x_data сначала вычислялись среднее арифметическое mx и выборочное стандартное отклонение sdx, а затем генерировалась псевдовыборка как случайная выборка из нормального закона распределения с параметрами mx и sdx. Важно отметить, что такой тип перестройки выборочных данных x_data является корректным только в ситуации, когда закон распределения данных x_data близок к некоторому нормальному закону распределения. Параметрическая бутстреп-процедура была реализована в виде функции r3_data.

При бутстреп-процедуре со случайным слагаемым сначала формировалась псевдовыборка с помощью классической бутстреп-процедуры, а затем к каждому элементу такой псевдовыборки добавлялось малое случайное слагаемое, взятое из нормального закона распределения с математическим ожиданием 0 и стандартным отклонением, равным 0.1. Бутстреп со случайным слагаемым был реализован в виде функции r4_data.

Пятый тип перестройки выборочных данных – ресамплинг без возвращений, при котором из выборки x_data случайно извлекались без возвращений элементы в количестве, равном 90% объема исходных данных. Ресамплинг без возвращений реализован в виде функции r5_data.

Для исследования смещения оценок мощности критерия Колмогорова–Смирнова в задачах 1–7, 11, 12 применялась специальная процедура перестройки выборочных данных, которую условно

можно назвать «без ресамплинга». Процедура перестройки данных без ресамплинга реализована в виде функции r0_data (в каждой задаче своей) и обеспечивает генерирование независимой от x_data псевдовыборки r0_data(x_data), которая имеет такой же объем и такой же закон распределения, как исходные данные x_data.

Оценивание мощности критерия при конкретном типе перестройки данных выполнялось путем создания 1000 псевдовыборок из данных x_data, вычисления при помощи критерия Колмогорова–Смирнова 1000 p-значений (для каждой псевдовыборки своего) и определения доли псевдовыборок, для которых p-значения меньше уровня значимости 0.05. Оценивание мощности критерия реализовано в виде функции test_power, интерфейс которой на языке программирования R представлен ниже.

```
test_power <- function(x,tau,r_data,
alpha=0.05,B=1000,n=length(na.omit(x)))
```

Шесть аргументов функции имеют следующую семантику: x – вектор выборочных данных; tau – функция, вычисляющая p-значение по условию задачи; r_data – функция, перестраивающая выборочные данные; alpha – уровень значимости; B – количество псевдовыборок, создаваемых для оценивания мощности критерия; n – объем каждой псевдовыборки.

На языке программирования R код функций, которые использовались для перестройки выборочных данных и оценивания мощности критерия Колмогорова–Смирнова, представлен на листинге 1.

Листинг 1

Набор функций для ресамплинга и оценивания мощности критерия

```
#####################################################
################
### Классический бутстреп
r1_data <- function(x,n=length(x)) { x <- na.omit(x)
sample(x,n,repl=TRUE)
}
### Непараметрический бутстреп
r2_data <- function(x,n=length(x)) { x <- na.omit(x)
 F <- ecdf(x)
 xu <- sort(unique(x)); xu <- c(min(x)-1, xu)
 yu <- sapply(xu,F)
 Qspl <- splinefun(yu,xu)
```

```
 sapply(runif(n),Qspl)
}
### Параметрический бутстреп
r3_data <- function(x,n=length(x)) { x <- na.omit(x)
 mx <- mean(x); sdx <- sd(x)
 rnorm(n,mx,sdx)
}
### Бутстреп со случайным слагаемым
r4_data <- function(x,n=length(x),eps=0.1) { x <- na.omit(x)
 sample(x,n,repl=TRUE) + rnorm(n,0,eps)
}
### Ресамплинг без возвращений
r5_data <- function(x,n=round(0.9*length(x))) { x <- na.omit(x)
 sample(x,n,repl=FALSE)
}
### Оценивание мощности критерия
test_power <- function(x,tau,r_data,
 alpha=0.05,B=1000,n=length(na.omit(x))) {
 x <- na.omit(x); tau.s <- 1:B
 for (i in 1:B) tau.s[i] <- tau(r_data(x=x,n=n))
 power <- sum(tau.s < alpha)/B
}
##########################################################
################
```

Тестовые задачи 1–7, 11, 12 решались по 100 раз каждая. По результатам многократного решения задач были вычислены медианы p-значений, медианы оценок мощности критерия Колмогорова–Смирнова, а также были найдены медианы смещений оценок на тестовых задачах и медианы всех попарных разностей различных оценок мощности критерия.

Результаты вычислений

В табл. 1 для тестовых задач 1–7, 11, 12 представлены медианы p-значений (столбец p_val) и медианы оценок мощности критерия, полученные при различных типах перестройки данных (столбцы power1 – power5), а также медианы оценок мощности, найденных без ресамплинга (столбец power0). Информация, выделенная в ячейках таблиц полужирным курсивом, комментируется далее отдельно.

Таблица 1

Медианы оценок мощности критерия
Колмогорова–Смирнова

	p_val	power0	power1	power2	power3	power4	power5
Задача 1	0.066	0.446	0.673	0.646	0.452	0.633	0.229
Задача 2	0.053	0.463	0.772	0.776	0.926	0.757	0.283
Задача 3	0.064	***0.504***	0.767	0.590	0.330	0.631	0.300
Задача 4	0.387	0.090	0.317	0.267	0.102	0.274	0.000
Задача 5	0.163	0.310	***0.513***	***0.506***	0.129	***0.471***	0.029
Задача 6	0.077	0.420	0.665	0.651	0.146	0.655	0.192
Задача 7	0.148	0.249	***0.533***	0.425	0.104	0.372	0.040
Задача 11	0.264	0.121	0.364	0.357	0.110	0.324	0.003
Задача 12	0.496	0.048	0.270	0.246	0.112	0.225	0.000

В табл. 1 использованы обозначения: powerk – медианы оценок мощности при ресамплинге с номером k (k = 1 – классический бутстреп, k = 2 – непараметрический бутстреп, k = 3 – параметрический бутстреп, k = 4 – бутстреп со случайным слагаемым, k = 5 – ресамплинг без возвращений).

Таблица 2

Медианы смещений оценок мощности
по сравнению с оценками без ресамплинга
для критерия Колмогорова–Смирнова

	s1	s2	s3	s4	s5
Задача 1	0.223	0.203	***0.006***	0.186	–0.206
Задача 2	0.314	0.311	0.464	0.294	–0.169
Задача 3	0.256	0.085	–0.180	0.125	–0.198
Задача 4	0.228	0.172	***0.006***	0.174	–0.082
Задача 5	0.201	0.204	–0.182	0.158	–0.276
Задача 6	0.235	0.241	–0.270	0.220	–0.220
Задача 7	0.275	0.175	–0.152	0.115	–0.214
Задача 11	0.244	0.238	***–0.011***	0.194	–0.111
Задача 12	0.216	0.202	0.067	0.178	***–0.045***

В табл. 2 для тестовых задач 1–7, 11, 12 приведены медианы смещений оценок мощности критерия при различных типах перестройки выборочных данных.

В табл. 2 использованы обозначения: sk – медианы смещений оценок мощности при ресамплинге с номером k (k имеет тот же смысл, как в табл. 1).

В табл. 3 для тестовых задач 1–7, 11, 12 приведены медианы разностей оценок мощности при классическом бутстрепе и оценок мощности при других типах ресамплинга.

Таблица 3

Медианы разностей оценок мощности
при классическом бутстрепе и оценок
при других типах ресамплинга
для критерия Колмогорова–Смирнова

	s1_2	s1_3	s1_4	s1_5
Задача 1	***0.019***	0.171	***0.017***	0.287
Задача 2	***0.006***	–0.130	***0.016***	0.391
Задача 3	0.099	0.374	0.129	0.428
Задача 4	***0.032***	0.182	0.054	0.296
Задача 5	***–0.005***	0.363	***0.025***	0.261
Задача 6	***–0.007***	0.468	***0.012***	0.273
Задача 7	***0.027***	0.384	0.156	0.391
Задача 11	***0.007***	0.240	***0.043***	0.318
Задача 12	***0.003***	0.137	***0.025***	0.244

В табл. 3 использованы обозначения: s1_k – медианы разностей оценок мощности при классическом бутстрепе и ресамплинге с номером k (k = 2, 3, 4, 5 имеет такой же смысл, как в табл. 1).

В табл. 4 для тестовых задач 1–7, 11, 12 приведены медианы попарных разностей оценок мощности при разных типах ресамплинга.

Таблица 4

Медианы попарных разностей оценок мощности
при разных типах ресамплинга
для критерия Колмогорова–Смирнова

	s2_3	s2_4	s2_5	s3_4	s3_5	s4_5
Задача 1	0.154	***–0.002***	0.257	–0.153	0.072	0.258
Задача 2	–0.132	***0.010***	0.382	0.141	0.536	0.381
Задача 3	0.292	***0.026***	0.246	–0.248	***–0.011***	0.286
Задача 4	0.167	***0.025***	0.229	–0.134	0.074	0.243
Задача 5	0.380	***0.034***	0.274	–0.338	0.058	0.226
Задача 6	0.471	***0.023***	0.302	–0.450	***–0.048***	0.271
Задача 7	0.310	0.132	0.297	–0.226	***0.044***	0.238
Задача 11	0.229	***0.031***	***0.031***	–0.197	0.074	0.265
Задача 12	0.130	***0.018***	0.225	–0.108	0.088	0.210

В табл. 4 использованы обозначения: sj_k – медианы разностей оценок мощности при j-ом и k-ом типах ресамплинга (j = 2, 3, 4; k = 3, 4, 5 имеют такой же смысл, как k в табл. 1).

В задачах 8–10 использовались не симулированные данные, а реальные данные о 30 психологических характеристиках студентов РГГУ. Поэтому оценивание мощности критерия Колмогорова–Смирнова без ресамплинга в этих задачах не производилось.

В табл. 5 для тестовой задачи 8 представлены p-значения (столбец p_val) и оценки мощности критерия, полученные при различных типах перестройки данных (столбцы power1 – power5) для всех 30 психологических характеристик. Нумерация типов ресамплинга, как в табл. 1.

Таблица 5

Оценки мощности критерия Колмогорова–Смирнова
в задаче 8

	subscale	p_val	power1	power2	power3	power4	power5
1	N1_Тревожность	0.159	0.722	0.603	0.040	0.416	0.059
2	N2_Враждебность	0.013	0.900	0.844	0.051	0.610	0.764
3	N3_Депрессивность	0.087	0.769	0.470	0.037	0.422	0.117
4	N4_Застенчивость	0.219	0.734	0.548	0.044	0.325	0.007
5	N5_Импульсивность	0.091	0.897	0.683	0.044	0.457	0.136
6	N6_Уязвимость	0.083	0.767	0.609	0.042	0.357	0.167
7	E1_Доброжелательность	0.000	0.996	0.735	0.036	0.921	1.000
8	E2_Общительность	0.001	0.967	0.443	0.049	0.806	0.999
9	E3_Настойчивость	0.132	0.772	0.621	0.048	0.376	0.057
10	E4_Активность	0.104	0.834	0.543	0.048	0.454	0.120
11	E5_Непоседливость	0.040	0.802	0.371	0.053	0.403	0.407
12	E6_Жизнерадостность	0.054	0.857	0.364	0.054	0.401	0.334
13	O1_Фантазия	0.001	0.992	0.547	0.045	0.899	1.000
14	O2_Эстетичность	0.000	0.993	0.566	0.047	0.919	1.000
15	O3_Чувства	0.025	0.915	0.536	0.048	0.567	0.694

Окончание табл. 5

	subscale	p_val	power1	power2	power3	power4	power5
16	O4_Действия	0.037	0.919	0.818	0.042	0.433	0.487
17	O5_Идеи	0.009	0.931	0.565	0.060	0.608	0.938
18	O6_Ценности	0.014	0.961	0.507	0.040	0.511	0.772
19	A1_Доверие	0.000	1.000	0.860	0.043	0.984	1.000
20	A2_Прямота	0.062	0.831	0.497	0.055	0.470	0.198
21	A3_Альтруизм	0.007	0.982	0.693	0.050	0.748	0.948
22	A4_Уступчивость	0.019	0.903	0.435	0.055	0.567	0.777
23	A5_Скромность	0.082	0.817	0.508	0.042	0.429	0.196
24	A6_Отзывчивость	0.026	0.949	0.653	0.053	0.544	0.761
25	C1_Компетентность	0.020	0.917	0.656	0.041	0.520	0.740
26	C2_Организованность	0.027	0.892	0.310	0.039	0.543	0.710
27	C3_Ответственность	0.035	0.965	0.723	0.039	0.577	0.594
28	C4_Целеустремленность	0.110	0.855	0.644	0.054	0.416	0.136
29	C5_Самодисциплина	0.111	0.760	0.712	0.041	0.374	0.042
30	C6_Осмотрительность	0.072	0.829	0.523	0.050	0.468	0.207

Таблица 6

Разности оценок мощности при классическом бутстрепе
и оценок при других типах ресамплинга
для критерия Колмогорова–Смирнова в задаче 8

	subscale	s1_2	s1_3	s1_4	s1_5
1	N1_Тревожность	0.119	0.682	0.306	0.663
2	N2_Враждебность	0.056	0.849	0.290	0.136
3	N3_Депрессивность	0.299	0.732	0.347	0.652
4	N4_Застенчивость	0.186	0.690	0.409	0.727
5	N5_Импульсивность	0.214	0.853	0.440	0.761
6	N6_Уязвимость	0.158	0.725	0.410	0.600
7	E1_Доброжелательность	0.261	0.960	0.075	–0.004
8	E2_Общительность	0.524	0.918	0.161	–0.032
9	E3_Настойчивость	0.151	0.724	0.396	0.715
10	E4_Активность	0.291	0.786	0.380	0.714
11	E5_Непоседливость	0.431	0.749	0.399	0.395
12	E6_Жизнерадостность	0.493	0.803	0.456	0.523
13	O1_Фантазия	0.445	0.947	0.093	–0.008
14	O2_Эстетичность	0.427	0.946	0.074	–0.007
15	O3_Чувства	0.379	0.867	0.348	0.221
16	O4_Действия	0.101	0.877	0.486	0.432
17	O5_Идеи	0.366	0.871	0.323	–0.007
18	O6_Ценности	0.454	0.921	0.450	0.189
19	A1_Доверие	0.140	0.957	0.016	0.000
20	A2_Прямота	0.334	0.776	0.361	0.633
21	A3_Альтруизм	0.289	0.932	0.234	0.034
22	A4_Уступчивость	0.468	0.848	0.336	0.126

Окончание табл. 6

	subscale	s1_2	s1_3	s1_4	s1_5
23	A5_Скромность	0.309	0.775	0.388	0.621
24	A6_Отзывчивость	0.296	0.896	0.405	0.188
25	C1_Компетентность	0.261	0.876	0.397	0.177
26	C2_Организованность	0.582	0.853	0.349	0.182
27	C3_Ответственность	0.242	0.926	0.388	0.371
28	C4_Целеустремленность	0.211	0.801	0.439	0.719
29	C5_Самодисциплина	0.048	0.719	0.386	0.718
30	C6_Осмотрительность	0.306	0.779	0.361	0.622

В табл. 6 для каждой из 30 подшкал опросника NEO PI-R из задачи 8 приведены разности оценок мощности при классическом бутстрепе и оценок мощности при других типах ресамплинга. Нумерация типов ресамплинга как в табл. 3.

В табл. 7 для каждой из 30 подшкал опросника NEO PI-R из тестовой задачи 8 приведены попарные разности оценок мощности при разных типах ресамплинга. Нумерация типов ресамплинга как в табл. 4.

Таблица 7

Попарные разности оценок мощности при разных типах ресамплинга для критерия Колмогорова–Смирнова в задаче 8

	subscale	s2_3	s2_4	s2_5	s3_4	s3_5	s4_5
1	N1_Тревожность	0.563	0.187	0.544	–0.376	–0.019	0.357
2	N2_Враждебность	0.793	0.234	0.080	–0.559	–0.713	–0.154
3	N3_Депрессивность	0.433	0.048	0.353	–0.385	–0.080	0.305
4	N4_Застенчивость	0.504	0.223	0.541	–0.281	0.037	0.318
5	N5_Импульсивность	0.639	0.226	0.547	–0.413	–0.092	0.321
6	N6_Уязвимость	0.567	0.252	0.442	–0.315	–0.125	0.190
7	E1_Доброжелательность	0.699	–0.186	–0.265	–0.885	–0.964	–0.079
8	E2_Общительность	0.394	–0.363	–0.556	–0.757	–0.950	–0.193
9	E3_Настойчивость	0.573	0.245	0.564	–0.328	–0.009	0.319
10	E4_Активность	0.495	0.089	0.423	–0.406	–0.072	0.334
11	E5_Непоседливость	0.318	–0.032	–0.036	–0.350	–0.354	–0.004
12	E6_Жизнерадостность	0.310	–0.037	0.030	–0.347	–0.280	0.067
13	O1_Фантазия	0.502	–0.352	–0.453	–0.854	–0.955	–0.101
14	O2_Эстетичность	0.519	–0.353	–0.434	–0.872	–0.953	–0.081
15	O3_Чувства	0.488	–0.031	–0.158	–0.519	–0.646	–0.127

Окончание табл. 7

	subscale	s2_3	s2_4	s2_5	s3_4	s3_5	s4_5
16	O4_Действия	0.776	0.385	0.331	–0.391	–0.445	–0.054
17	O5_Идеи	0.505	–0.043	–0.373	–0.548	–0.878	–0.330
18	O6_Ценности	0.467	–0.004	–0.265	–0.471	–0.732	–0.261
19	A1_Доверие	0.817	–0.124	–0.140	–0.941	–0.957	–0.016
20	A2_Прямота	0.442	0.027	0.299	–0.415	–0.143	0.272
21	A3_Альтруизм	0.643	–0.055	–0.255	–0.698	–0.898	–0.200
22	A4_Уступчивость	0.380	–0.132	–0.342	–0.512	–0.722	–0.210
23	A5_Скромность	0.466	0.079	0.312	–0.387	–0.154	0.233
24	A6_Отзывчивость	0.600	0.109	–0.108	–0.491	–0.708	–0.217
25	C1_Компетентность	0.615	0.136	–0.084	–0.479	–0.699	–0.220
26	C2_Организованность	0.271	–0.233	–0.400	–0.504	–0.671	–0.167
27	C3_Ответственность	0.684	0.146	0.129	–0.538	–0.555	–0.017
28	C4_Целеустремленность	0.590	0.228	0.508	–0.362	–0.082	0.280
29	C5_Самодисциплина	0.671	0.338	0.670	–0.333	–0.001	0.332
30	C6_Осмотрительность	0.473	0.055	0.316	–0.418	–0.157	0.261

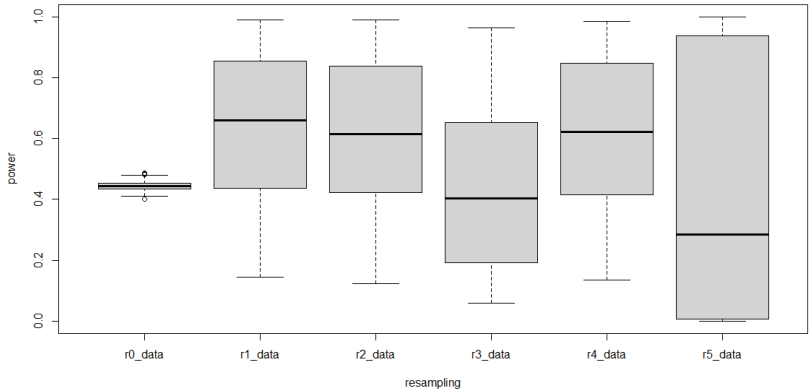

Рис. 1. Законы распределения оценок мощности при разных типах ресамплинга

По результатам многократного решения тестовых задач 1–7, 11, 12 в рамках каждой задачи были найдены 95% доверительные интервалы p-значений, доверительные интервалы оценок мощности критерия Колмогорова–Смирнова, а также доверительные интервалы смещений оценок и всех попарных разностей различных оценок мощности критерия. Кроме того, в задачах 1–7, 11, 12 при помощи одновыборочного критерия Уилкоксона были выполнены проверки статистических гипотез о том, что медианы оценок мощности не отличаются от 0.5, а также о том, что медианы попарных разностей различных оценок мощности критерия равны 0. Критерий Уилкоксона применялся аналогично в задачах 8–10 для проверки статистических гипотез о равенстве нулю медиан попарных разностей оценок мощности критерия Колмогорова–Смирнова.

Были исследованы законы распределения и построены соответствующие графики для всех оценок мощности критерия Колмогорова–Смирнова, для смещений оценок мощности и всех попарных разностей оценок мощности. На рис. 1 в качестве примера представлены boxplot графики законов распределения оценок мощности критерия Колмогорова–Смирнова при различных типах перестройки выборочных данных в тестовой задаче 1.

Обсуждение

Свойства различных оценок мощности критерия Колмогорова–Смирнова, которые проявляются в тестовых задачах 1–7, 11, 12, хорошо видны на рис. 1. Оценка мощности, полученная без ресамплинга (тип перестройки данных r0_data), имеет всегда гораздо меньшую вариабельность, чем другие оценки мощности. В тестовой задаче 1 эмпирически получен для оценки мощности без ресамплинга 95% доверительный интервал (0.414; 0.481), а для медианы этой оценки доверительный интервал (0.442; 0.448). Если тип ресамплинга – классический бутстреп (r1_data), непараметрический бутстреп (r2_data), параметрический бутстреп (r3_data) или бутстреп со случайным слагаемым (r4_data), то соответствующие доверительные интервалы для оценок мощности и медиан оценок мощности имеют приблизительно равную длину для разных типов ресаплинга, которая примерно в 10 раз больше, чем в случае без ресамплинга. Например, доверительный интервал для оценки мощности, найденной при помощи непараметрического бутстрепа, равен (0.217; 0.974), а доверительный интервал для медианы этой оценки имеет вид (0.567; 0.669). Ресамплинг без возвращений (r5_data) дает оценку мощности с самой большой дисперсией. В тестовой задаче 1 получен 95% доверительный интервал для такой оценки (0.000; 1.000), а для медианы этой оценки доверительный интервал (0.308; 0.500). Длины этих доверительных интервалов приблизительно 1 и 0.2, что типично для ресамплинга без возвращений и в других тестовых задачах 1–7, 11, 12. Плотность вероятности оценки мощности при ресамплинге без возвращений, как правило, двухмодовая с одной модой вблизи 0, а другой модой вблизи 1. Если доминирует значение плотности вероятности для моды вблизи 0, то при перестройке выборочных данных стабильно воспроизводится статистическое решение в пользу нулевой гипотезы. Если доминирует мода вблизи 1, то при ресамплинге нулевая гипотеза стабильно отвергается в пользу альтернативы. Таким образом, ресамплинг без возвращений способен обеспечить контроль воспроизводимости результатов статистического вывода при перестройке выборочных данных. Такой контроль имеет важное значение, поскольку позволяет расширить возможности классической методологии проверки статистических гипотез.

В результате применения критерия Уилкоксона на уровне значимости 0.05 были обнаружены полезные эмпирические закономерности для различных оценок мощности статистического критерия Колмогорова–Смирнова. Представленные в табл. 1 медианы оценок мощности обычно статистически значимо отлича-

ются от 0.5. Исключения составляют случаи, выделенные в табл. 1 полужирным курсивом. Важно отметить, что если медиана оценки мощности в рамках некоторой тестовой задачи статистически значимо меньше 0.5, то при перестройке выборочных данных чаще воспроизводится решение в пользу нулевой гипотезы. Если медиана оценки мощности статистически значимо больше 0.5, то чаще принимается решение в пользу альтернативной гипотезы. Медианы смещений оценок мощности, представленные в табл. 2, по критерию Уилкоксона статистически значимо положительные для всех тестовых задач 1–7, 11, 12, если оценки мощности получены при помощи классического бутстрепа, непараметрического бутстрепа или бутстрепа со случайным слагаемым. Положительные медианы смещений имеют величины приблизительно от 0.1 до 0.3 в зависимости от типа ресамплинга и от задачи. Медианы смещений оценок мощности статистически значимо отрицательные для всех тестовых задач, представленных в табл. 2, кроме задачи 12, если для перестройки данных применялся ресамплинг без возвращений. Отрицательные медианы смещений имеют величины приблизительно от –0.3 до –0.1 в зависимости от задачи. Информация о медианах разностей различных оценок мощности, размещенная в табл. 3 и табл. 4, свидетельствует о том, что оценки мощности, найденные при помощи классического бутстрепа, непараметрического бутстрепа и бутстрепа со случайным слагаемым, различаются слабо и эти различия не являются статистически значимыми по критерию Уилкоксона. Отметим, что положительные смещения трех близких оценок мощности, названных выше, во многих ситуациях могут быть хорошо скомпенсированы отрицательным смещением оценки мощности, полученной при помощи ресамплинга без возвращений. По этой причине предлагается следующее эвристическое правило для прогноза медианы несмещенной оценки мощности, полученной без ресамплинга. Медиана оценки мощности без ресамплинга примерно равна полусумме медиан двух других оценок мощности, одна из которых получена при помощи непараметрического бутстрепа, а другая при помощи ресамплинга без возвращений. Легко убедиться, используя табл. 1, что сформулированное эвристическое правило хорошо работает для тестовых задач 1–3, 5–7, когда p-значение приблизительно от 0.05 до 0.2. В ситуации, когда p-значение больше 0.2, в качестве прогноза медианы оценки мощности без ресамплинга можно успешно использовать медиану оценки мощности, полученной при помощи параметрического бутстрепа, как это видно в табл. 1 для задач 4, 11 и 12. Для тестовых задач 8–10, где рассматривались реальные данные, вычисление оценок мощности без ресамплинга невозможно, но несмещенные

оценки мощности можно прогнозировать как полусумму двух других оценок мощности, одна из которых получена при помощи непараметрического бутстрепа, а другая при помощи ресамплинга без возвращений. Можно легко проверить, что такой прогноз несмещенной оценки мощности критерия Колмогорова–Смирнова, выполненный в тестовой задаче 8, очень хорошо согласуется с p-значениями, приведенными в табл. 5. Аналогичный результат получен в тестовых задачах 9–10.

Методы перестройки выборочных данных и программные средства оценивания мощности статистического критерия, разработанные в рамках данного исследования, являются универсальными и могут быть полезны для анализа свойств оценок мощности других одновыборочных статистических критериев, а также для развития классической методологии проверки статистических гипотез.

Заключение

В работе при помощи методов перестройки выборочных данных выполнено эмпирическое исследование пяти различных оценок мощности статистического критерия Колмогорова–Смирнова на тестовых задачах по проверке гипотез о законе распределения. Приближенно найдены смещения оценок мощности и сформулировано эвристическое правило для прогнозирования несмещенной оценки мощности критерия Колмогорова–Смирнова.

Литература

Воронов, Зайчикова, Синицын 2000 – *Воронов А.Я., Зайчикова Е.А., Синицын В.Ю.* Некоторые кросс-культурные психологические различия российских и американских студентов // Ценностная и социальная идентичность российской гуманитарной интеллигенции: тезисы Всерос. конф. М.: РГГУ, 2000. С. 120–123.

Колмогоров 1940 – *Колмогоров А.Н.* Об одном новом подтверждении законов Менделя // Доклады АН СССР. 1940. Т. 27. № 1. С. 38–42.

Кремер 2019 – *Кремер Н.Ш.* Теория вероятностей и математическая статистика. М.: Юрайт, 2019.

Практикум по математической статистике 2021 – Практикум по математической статистике в вычислительной среде R: Учебное пособие / В.Ю. Синицын. Казань: Бук, 2021.

Прохоров, Пономаренко 2019 – *Прохоров Ю.В., Пономаренко Л.С.* Лекции по теории вероятностей и математической статистике. М.: Юрайт, 2019.

Синицын, Кашпарова 2004 – *Синицын В.Ю., Кашпарова В.С.* Психометрические

свойства теста NEO PI-R и его применение для исследования психологического портрета студента // Психолого-педагогические исследования в системе образования: II Всерос. конф. М., Челябинск, 2004. С. 31–34.

Смирнов 1939 – *Смирнов Н.В.* Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках // Бюллетень МГУ. 1939. Т. 2. № 2. С. 3–14.

Статистический анализ 2010 – Статистический анализ данных в системе R: Учебное пособие / Ред. А.Г. Буховец и др. Воронеж: ВГАУ, 2010.

Kolmogoroff 1933 – *Kolmogoroff A.* Sulla determinazione empirica di una legge di distribuzione // Guornale dell' Instituto Italiano degli Attuari. 1933. Vol. 4. № 1. P. 83–91.

References

Bukhovets, A.G. et al. (eds.) (2010), *Statisticheskii analiz dannykh v sisteme R* [Statistical data analysis in the R system. Teaching aid], VSAU, Voronezh, Russia.

Kolmogoroff, A. (1933), "Sulla determinazione empirica di una legge di distribuzione", *Guornale dell' Instituto Italiano degli Attuari*, vol. 4, no. 1, pp. 83–91.

Kolmogorov, A.N. (1940), "On a new confirmation of Mendel's laws", *Transactions of the USSR Academy of Sciences,* vol. 27, no. 1, pp. 38–42.

Kremer, N.Sh. (2019), *Teoriya veroyatnostei i matematicheskaya statistika* [Theory of Probability and Mathematical Statistics], Yurait, Moscow, Russia.

Prokhorov, Yu.V. and Ponomarenko, L.S. (2019) *Lektsii po teorii veroyatnostei i matematicheskoi statistike* [Lectures on the theory of probability and mathematical statistics], Yurait, Moscow, Russia.

Sinitsyn, V.Yu. and Kashparova, V.S. (2004), "Psychometric characteristics of the NEO PI-R test and its application for studying the psychological portrait of a student", *Psikhologo-pedagogicheskie issledovaniya v sisteme obrazovaniya* [Psychological and pedagogical research in the education system], *Proc. of the 2nd All-Russian Conference,* Moscow, Chelyabinsk, Russia, pp. 31–34.

Sinitsyn, V.Yu. (2021), *Praktikum po matematicheskoi statistike v vychislitel'noi srede R* [Practical course of mathematical statistics in the computing environment R], Buk, Kazan', Russia.

Smirnov, N.V. (1939), "Evaluation of the discrepancy between empirical distribution curves in two independent samples", *Bulletin of Moscow State University*, vol. 2, no. 2, pp. 3–14.

Voronov, A.Ya., Zaichikova, E.A. and Sinitsyn, V.Yu. (2000), "Some cross-cultural psychological differences between Russian and American students", *Tsennostnaya i sotsial'naya identichnost' rossiiskoi gumanitarnoi intelligentsii* [The value and social identity of the Russian humanitarian intelligentsia], *Proc. of the All-Russian Conference,* Moscow, Russia, pp. 120–123.

Информация об авторах

Вячеслав Ю. Синицын, кандидат физико-математических наук, доцент, Российский государственный гуманитарный университет, Москва, Россия; 125047, Россия, Москва, Миусская пл., д. 6; fpmrggu@yandex.ru

Екатерина С. Ступакова, студент, Российский государственный гуманитарный университет, Москва, Россия; 125047, Россия, Москва, Миусская пл., д. 6; stupakova13@gmail.com

Information about the authors

Vyacheslav Yu. Sinitsyn, Cand. of Sci. (Physics and Mathematics), associate professor, Russian State University for the Humanities, Moscow, Russia; bld. 6, Miusskaya Sq., Moscow, Russia, 125047; fpmrggu@yandex.ru

Ekaterina S. Stupakova, student, Russian State University for the Humanities, Moscow, Russia; bld. 6, Miusskaya Sq., Moscow, Russia, 125047; stupakova13@gmail.com

Дизайн обложки
Е.В. Амосова

Корректор
О.К. Юрьев

Компьютерная верстка
Н.В. Москвина

Подписано в печать 26.09.2022.
Формат 60×90 1/16.
Уч.-изд. л. 7,0. Усл. печ. л. 7,6.
Тираж 1050 экз. Заказ № 1598

Издательский центр
Российского государственного
гуманитарного университета
125047, Москва, Миусская пл., 6
www.rsuh.ru