DOI 10.17587/issn.1684-6400

ISSN 1684-6400

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 27

2021

№ 3

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

САПР

КОМПЬЮТЕРНАЯ ГРАФИКА

МЕТОДЫ ПРОГРАММИРОВАНИЯ

ОПЕРАЦИОННЫЕ СИСТЕМЫ И СРЕДЫ

ТЕЛЕКОММУНИКАЦИИ И ВЫЧИСЛИТЕЛЬНЫЕ СЕТИ

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

НЕЙРОСЕТИ И НЕЙРОКОМПЬЮТЕРЫ

СТРУКТУРНЫЙ СИНТЕЗ

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ

ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

ОПТИМИЗАЦИЯ И МОДЕЛИРОВАНИЕ

ИТ В ОБРАЗОВАНИИ

ГИС

Рисунки к статье П. А. Шиловских

«ПРОТОКОЛ ЦИФРОВОЙ ПОДПИСИ НА ОСНОВЕ КРИПТОКОДОВЫХ КОНСТРУКЦИЙ»

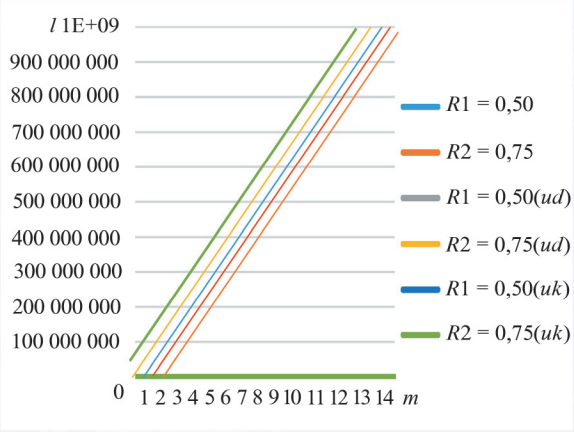


Рис. 6. Зависимость сложности формирования криптограммы в различных *GF*(2*m*)

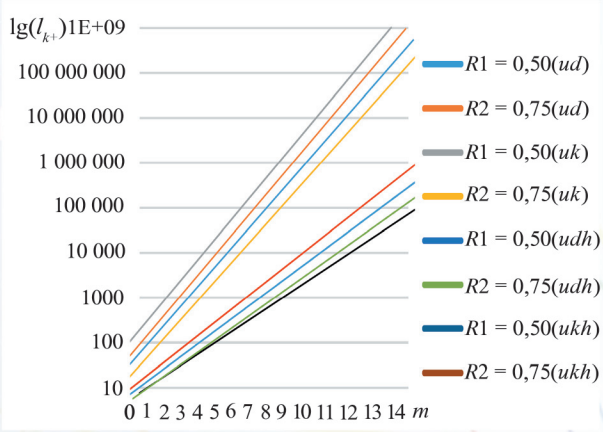


Рис. 9. Зависимость сложности формирования криптограммы в различных *GF*(2*m*)

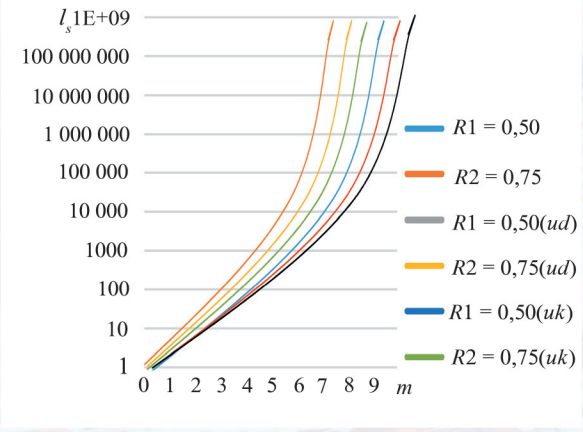


Рис. 7. Зависимость сложности раскодирования кодограммы от мощности поля

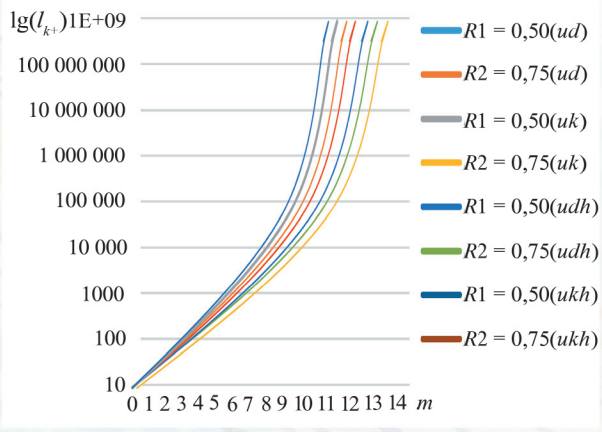


Рис. 10. Зависимость сложности раскодирования в различных *GF*(2*m*)

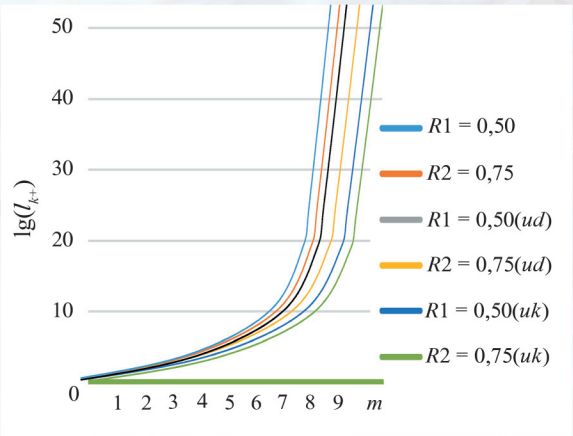


Рис. 8. Зависимость сложности взлома на основе перестановочного декодирования от мощности поля

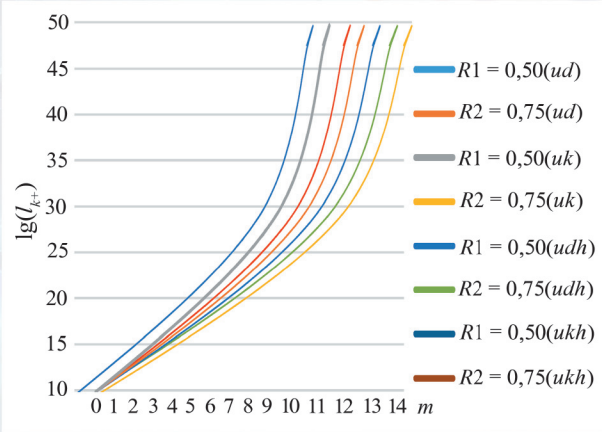


Рис. 11. Зависимость сложности взлома ГККК над *GF*(2*m*) (перестановочное декодирование)

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 27
2021
№ 3

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г. DOI 10.17587/issn.1684-6400

УЧРЕДИТЕЛЬ

Издательство "Новые технологии"

СОДЕРЖАНИЕ

Информация о журнале доступна по сети Internet по адресу **http://novtex.ru/IT.**
Журнал включен в систему Российского индекса научного цитирования и базу данных RSCI на платформе Web of Science.
Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

INFORMATION TECHNOLOGIES
INFORMACIONNYE TEHNOLOGII

Vol. 27
2021
No. 3

THEORETICAL AND APPLIED SCIENTIFIC AND TECHNICAL JOURNAL

Published since November 1995

DOI 10.17587/issn.1684-6400

ISSN 1684-6400

CONTENTS

Complete Internet version of the journal at site: **http://novtex.ru/IT.**
According to the decision of the Higher Certifying Commission of the Ministry of Education of Russian Federation, the journal is inscribed in "The List of the Leading Scientific Journals and Editions wherein Main Scientific Results of Theses for Doctor's or Candidate's Degrees Should Be Published"

УДК 004.7, 001.83 DOI: 10.17587/it.27.115-124

А. Г. Абрамов[1], канд. физ.-мат. наук, вед. науч. сотр., e-mail: abramov@runnet.ru,
А. А. Гончар[2], зам. директора, e-mail: andrey.gonchar@jscc.ru,
А. В. Евсеев[1], директор, e-mail: evseev@runnet.ru,
Б. М. Шабанов[2], д-р техн. наук, доц., директор, e-mail: shabanov@jscc.ru,
[1] СПбО МСЦ РАН — филиал ФГУ ФНЦ НИИСИ РАН,
[2] МСЦ РАН — филиал ФГУ ФНЦ НИИСИ РАН

Национальная исследовательская компьютерная сеть нового поколения: текущее состояние и концепция развития[1]

Систематизирована актуальная информация о статусе работ по обеспечению текущего функционирования и планах развития Национальной исследовательской компьютерной сети нового поколения (НИКС), созданной в 2019 г. по результатам интеграции отраслевых телекоммуникационных сетей сферы высшего образования и науки России RUNNet и RASNet. Представлены нормативные основания и предпосылки создания в стране единой научно-образовательной сети. Приведены ключевые характеристики ведущих зарубежных национальных научно-образовательных сетей. Обозначены параметры телекоммуникационной инфраструктуры и сетевой связности НИКС по состоянию на текущий момент, примеры сервисов, развиваемых в интересах российского научно-образовательного сообщества. Повышенное внимание уделено представлению основных направлений ускоренного развития НИКС на 2021—2024 гг. и ожидаемых результатов.

Ключевые слова: *национальная исследовательская компьютерная сеть нового поколения, НИКС, национальная научно-образовательная сеть, NREN, телекоммуникационная инфраструктура, сетевая связность, сетевые сервисы, концепция развития*

Введение

Накопленный многолетний мировой опыт создания и развития глобальных телекоммуникационных сетей подчеркивает значимую роль в этих процессах отраслевых сетей сферы науки и образования. В большинстве стран мира в разные периоды времени были созданы и поступательно совершенствуются специализированные отраслевые сети, за которыми закреплено общепринятое наименование — *национальная научно-образовательная сеть* (National Research and Education Network, NREN) [1—3].

Национальная научно-образовательная сеть — информационно-телекоммуникационная сеть, обладающая повышенными характеристиками в сравнении с функционирующими сетями общего пользования, высокопроизводительная инфраструктура масштаба страны, которая эксплуатируется в интересах науки и образования, обеспечивает сетевую связность пользователей, межсетевое взаимодействие с зарубежными NREN и консорциумами с повышенными требованиями к качеству сервиса, доступ пользователей в глобальное ИКТ-пространство, а также является ядром развития и провайдером базовых сетевых сервисов, сервисов коллективного пользования и специализированных научно-образовательных сервисов [4, 5].

К основным задачам, решаемым NREN, принято относить также взаимодействие с региональными научно-образовательными сетями страны, организацию связности с публичными сетями и Интернетом, квалифицированную поддержку пользователей, управление операционными аспектами сервисов и финансами, публичность и продвижение [1, 3—5].

[1]Работа выполнена в МСЦ РАН в рамках государственного задания № 0580-2021-0014.

Сегодня NREN функционируют (находясь на разных ступенях развития) в качестве неотъемлемых частей национальных информационно-телекоммуникационных инфраструктур (ИТКИ) более чем в 140 странах, обычно координируются государственными органами управления наукой и образованием, представляют страну в международных проектах, при реализации которых используются современные средства телекоммуникаций, сетевые технологии, а также специализированные сервисы.

В связи с обсуждаемой проблематикой достаточно упомянуть широко известные в профессиональном сообществе отраслевые сети США (Internet2, https://internet2.edu, и ESnet, https://es.net), Германии (DFN, https://dfn.de), Франции (RENATER, https://www.renater.fr), Нидерландов (SURFnet, https://surf.nl), Италии (GARR, https://garr.it) и др.

Примером реализуемых глобальных международных ИКТ-проектов является инициатива Еврокомиссии — EOSC (European Open Science Cloud, https://eosc-portal.eu), составными компонентами которой являются общеевропейская инфраструктура высокопроизводительных вычислений (PRACE, https://prace-ri.eu), инфраструктура национальных научно-образовательных сетей (GÉANT, https://geant.org) и широкий спектр сервисов, реализуемых в рамках отдельных проектов (EGI, EUDAT и др.).

В качестве другого примера масштабного технологического проекта можно указать на межконтинентальный транснациональный проект GNA (Global Network Architecture, https://gna-re.net), целью которого является развертывание надежной, устойчивой, высокопроизводительной ИТКИ нового поколения для нужд интернационального сообщества науки и образования.

Среди крупных научно-исследовательских проектов, интенсивно использующих средства ИКТ и реализуемых при участии ученых и исследователей из многих стран мира, включая и РФ, можно упомянуть проекты в области физики высоких энергий и физики частиц на БАК в ЦЕРН, ITER, XFEL, ESRF, DESY, FAIR, Belle II, в области астрофизики и спутниковых наблюдений — EUMETSAT, SKA, NOvA, XENON, LIGO и др.

В иерархии телекоммуникационных сетей науки и образования NREN взаимодействует с аналогичными по функциям зарубежными сетями и наднациональными сетевыми консорциумами, управляет научно-образовательной сетью национального масштаба, осуществляет верхнеуровневое присоединение региональных научно-образовательных сетей страны, предоставляет услуги локальным сетям пользователей [5].

Задачи повышения эффективности реализации международных проектов и взаимной кооперации успешно решают наднациональные объединения — консорциумы научно-образовательных сетей GÉANT (Европа), NORDUnet (Скандинавские страны), Asia@Connect и APAN (Азиатско-Тихоокеанский регион), RedClara (Латинская Америка), AfricaConnect (Африка).

Полезно заметить, что GÉANT представляет собой ассоциацию научно-образовательных сетей европейских стран и является координирующим партнером для одноименной общеевропейской научно-образовательной сети, имеющей пропускную способность магистральной инфраструктуры 100 Гбит/с, объединяющей более 50 млн пользователей из 10 тыс. организаций [6, 7]. Согласно Уставу членами ассоциации могут быть только NREN (юридические лица, в уставные цели которых не входит производственная или коммерческая деятельность) или их ассоциации. Подключение к сети GÉANT без членства в ассоциации возможно лишь значимых для международного научного сотрудничества сетей и организаций, нацеленных на поддержку финансируемых государством исследований и образования.

В настоящей статье рассматриваются существенные основания и предпосылки для создания Национальной исследовательской компьютерной сети нового поколения (НИКС) в качестве NREN России, представляется ход реализации мероприятий по созданию НИКС, фиксируется ее текущее состояние, приводятся и обсуждаются вопросы обеспечения функционирования и планы ускоренного развития проекта на 2021—2024 гг.

1. Нормативные основания и предпосылки создания НИКС в России

Достижение отдельных целей нацпроекта "Наука" (в ближайшей перспективе — нацпроекта "Наука и университеты") предусматривает, в том числе, решение задач создания передовой инфраструктуры научных исследований и разработок, инновационной деятельности, включая создание и развитие сети уникальных научных установок (УНУ) класса "мегасайенс", создания научных центров мирового уровня (НЦМУ), на-

учно-образовательных центров мирового уровня (НОЦ), Центров компетенции Национальной технологической инициативы (ЦК НТИ).

В круг задач нацпроекта входит развитие инфраструктуры и поддержка функционирования распределенной сети центров коллективного пользования научно-технологическим оборудованием (ЦКП), поддержка создания и развития востребованных УНУ, обеспечение доступа исследовательских групп к национальным и зарубежным информационным ресурсам и сервисам, а также участия российских ученых и исследовательских групп в международных проектах, предоставляющих доступ к новым компетенциям и ресурсам организации с учетом национальных интересов.

Представляется вполне очевидным, что в качестве одного из системообразующих компонентов ИТКИ страны, вносящего вклад в решение задач, достижение результатов и ключевых показателей нацпроекта, должна выступать соответствующая современному уровню достижений отрасли специализированная инфраструктура сферы науки и образования.

Положением о Минобрнауки России к сфере его деятельности отнесены, в том числе, функции по оказанию государственных услуг и управлению государственным имуществом в сфере высшего образования, научной, научно-технической и инновационной деятельности, включая деятельность *национальной исследовательской компьютерной сети нового поколения.*

Обращаясь кратко к истории развития научно-образовательных сетей в нашей стране, стоит отметить, что с начала 1990-х гг. при господдержке в рамках федеральных целевых и государственных научных программ, межведомственных и ведомственных программ и проектов создавались отдельные отраслевые телекоммуникационные сети, эксплуатировавшиеся в интересах отдельных научно-образовательных сообществ [5, 8—10].

На первых этапах соответствующие проекты были инициированы и выполнялись преимущественно силами крупных научно-исследовательских и образовательных организаций (МГУ им. М. В. Ломоносова, ФГАУ ГНИИ ИТТ "Информика", НИЦ "Курчатовский институт", МСЦ РАН, ИКИ РАН, Университет "ИТМО") при поддержке различных государственных структур (Минсвязи России, Миннауки России, Госкомвуз, РАН и др.) [8].

Научно-образовательное телекоммуникационное пространство страны строилось в виде ряда независимых сетей, в значительной степени базировавшихся на собственной (или арендованной) телекоммуникационной инфраструктуре и, при наличии потребности, в собственных же международных каналах. Некоторые из таких сетей изначально задумывались и проектировались как сети общего назначения, другие — как специализированные проблемно-ориентированные сети [5, 8—10].

Примеры таких сетей федерального и регионального уровней: RUNNet (Russian UNiversity Network), RASNet (Russian Academy of Sciences Network), RSSI (Russian Space Science Internet), RBNet (Russian Backbone Network), RUHEP/Radio-MSU, сеть участников НИЦ "Курчатовский институт" (НИЦ КИ), FREEnet (Москва), RELARN-IP (Москва), ЮМОС (Москва), РОКСОН (Санкт-Петербург), SENet (Республика Татарстан), ПЕРСОНА (Пермский край), сети региональных отделений РАН и т.д. Некоторые сети в той или иной степени функционируют и в настоящее время, другие же выведены из эксплуатации.

Множество разрозненных сетей по своей сути решало аналогичные или близкие задачи для сферы образования и науки, при этом имело место дублирование отдельных функций и параллельное бюджетное финансирование. Попытки создания единой сети национального уровня сложившимся сообществом периодически предпринимались, но необходимого развития соответствующие идеи не получали, в том числе и в силу отсутствия централизованного решения профильных федеральных органов исполнительной власти (ФОИВ).

Такое положение дел, в целом, противоречило мировому опыту создания и эволюции NREN, указывающему на наличие в большинстве стран единой научно-образовательной сети национального уровня, поддерживаемой государством и выполняющей интегрирующие функции в отношении региональных и межрегиональных сетей, включая и проблемно-ориентированные. Можно заметить здесь, что в нескольких странах, в том числе в США и Китае, имеются отдельные крупные, взаимодействующие между собой NREN масштаба страны.

В США (одном из мировых лидеров развития NREN) на протяжении многих лет успешно сосуществуют и сотрудничают друг с другом некоммерческий консорциум компьютерных сетей Internet2 и проблемно-ориентированная сеть ESnet, пропускная способность магистральной инфраструктуры которых составляет

сегодня от 100 до 400 Гбит/с. Целевыми пользователями Internet2 являются университеты, исследовательские институты, некоторые правительственные организации, большинство региональных образовательных сетей. ESnet — сеть, ориентированная на поддержку научных исследований в энергетической отрасли, основными пользователями которой являются национальные лаборатории и технологические центры Минэнергетики США.

Подобная модель представляется вполне применимой для нашей страны в условиях создании единой NREN общего назначения и наличия действующей "энергетической" сети НИЦ КИ, эксплуатируемой в интересах российских организаций, вовлеченных в обработку данных экспериментов на БАК (центров уровня Tier-2), а также участия НИЦ КИ в качестве центра уровня Tier-1 в глобальной грид-инфраструктуре Worldwide LHC Computing Grid (WLCG).

Фактически по состоянию на 2018 г. в нашей стране в полноценно функционирующем виде сохранились только две наиболее крупные научно-образовательные сети общего назначения (созданные в своем первоначальном виде еще в 1994 г.): RUNNet — федеральная университетская компьютерная сеть [4, 5, 11] и RASNet — сеть Российской академии наук.

В результате реорганизации в 2018 г. ведомственного учреждения в сфере информатизации образования и науки — ФГАУ ГНИИ ИТТ "Информика", на протяжении многих лет выполнявшего функции оператора сети RUNNet, и последующего разделения Министерства образования и науки Российской Федерации на два ФОИВ, функции управления сетью были переданы в подведомственную Минпросвещения России организацию, для которой деятельность по эксплуатации университетской сети являлась непрофильной.

В соответствии с решениями, совместно принятыми двумя ФОИВ, МСЦ РАН в качестве администратора и оператора сетей RUNNet и RASNet реализовал в 2019 г. комплекс мероприятий по обеспечению бесперебойного функционирования сетей и созданию на их платформе единой NREN России. В ходе работ функции управления объединенной сетью, а также имущество, обязательства и кадровые ресурсы RUNNet были переданы в МСЦ РАН.

НИКС была создана и функционирует в качестве NREN страны, выполняя традиционные для таких сетей функции, в том числе и на международной арене. НИКС как единая сеть сферы образования и науки нацелена на предоставление научным и образовательным организациям возможностей для выполнения исследований и разработок по приоритетным направлениям научно-технологического развития, участия в российских и международных научных проектах, базирующихся на использовании устойчивой и отвечающей современным требованиям отраслевой сети, интегрированной в инфраструктуру мировых NREN [5].

2. Текущий статус работ по проекту НИКС

В ходе многолетней эксплуатации сетей RUNNet и RASNet сформировалось устойчивое ядро пользователей (организаций науки и высшего образования), активно использующих ее инфраструктуру и сервисы в своей деятельности. По состоянию на 2020 г. НИКС имеет точки присутствия в 34 субъектах РФ, напрямую предоставляя услуги более чем 150 организациям высшего образования и науки, среди которых — федеральные, национальные исследовательские и опорные университеты, ведущие научные центры и институты РАН. К сети подключено 16 крупнейших суперкомпьютерных центров (СКЦ) сферы науки и образования, более 150 ЦКП и более 100 УНУ, другие объекты научной инфраструктуры коллективного пользования.

Функционирование НИКС основано на эксплуатации и развитии глобальной гетерогенной сети передачи данных, обеспечивающей прямое сетевое взаимодействие целевых пользователей, сетевых субъектов и внешних генераторов больших данных. В составе телекоммуникационной инфраструктуры НИКС, как и других сетей аналогичного масштаба и функций, выделяются магистральная (опорная) сетевая инфраструктура, региональная инфраструктура и инфраструктура доступа.

Работу географически распределенной сети обеспечивают 16 маршрутизаторов, 90 узлов связи, выполняющих разные функции и построенных на базе различного телекоммуникационного оборудования. Магистральная инфраструктура НИКС объединяет магистральные телекоммуникационные узлы связи (узлы федерального и регионального уровней и зарубежные), каналы между ними и формирует опорную сеть передачи данных от Амстердама до Хабаровска.

Федеральные узлы связи расположены в Москве и Санкт-Петербурге, объединяются в городах высокоскоростными каналами с полным резервированием. Узлы между городами связаны четырьмя каналами (4 × 10 Гбит/с), организованными на базе физически независимых магистралей. Федеральные узлы связаны магистральными каналами с региональными магистральными узлами и узлами доступа, с зарубежными NREN, с региональными научно-образовательными сетями страны, а также имеют прямую связность с сетями отдельных коммерческих операторов и Интернетом.

Узлы НИКС, расположенные в некоторых крупных городах страны (Екатеринбург, Курган, Нижний Новгород, Новосибирск, Пермь, Самара, Саратов, Томск, Уфа, Челябинск, Хабаровск), представляют собой магистральные узлы с размещенным телекоммуникационным оборудованием и используются для подключения расположенных в субъектах организаций. Каналы связи между узлами имеют пропускную способность преимущественно от 1 до 10 Гбит/с. Сегмент магистральной инфраструктуры сети в европейской части, связывающий между собой Москву, Нижний Новгород, Пермь, Екатеринбург, Курган, Челябинск, Уфу, Самару и Саратов, представляет собой транспортное "кольцо" с пропускной способностью 10 Гбит/с.

Функционирование магистральной инфраструктуры НИКС за границей основывается на устоявшемся сотрудничестве с сетью NORDUnet (для доступа к ресурсам зарубежных NREN) и межсетевом взаимодействии с несколькими Tier-1-операторами (для доступа в Интернет и межсетевого взаимодействия с отдельными NREN за пределами Европы).

Международные узлы расположены на площадках сети NORDUnet (Стокгольм), Национального института ядерной физики и физики высоких энергий (Амстердам) и СКЦ Финляндии (CSC, Хельсинки). НИКС имеет два независимых подключения к зарубежным NREN — в Хельсинки (к GÉANT) и в Стокгольме (к NORDUnet) с пропускной способностью по 10 Гбит/с.

Выстроенная инфраструктура доступа НИКС позволяет поддерживать работоспособность имеющихся сегментов и осуществлять подключение к умеренно утилизированным участкам инфраструктуры новых пользователей.

Наиболее развитые региональные сегменты НИКС расположены в Москве и Санкт-Петербурге, где к сети подключено большинство организаций высшего образования, большое число институтов РАН, региональные научно-образовательные сети, ряд организаций культуры и здравоохранения. Инфраструктура доступа основана на использовании собственных (в Москве, >300 км) или арендуемых (в Санкт-Петербурге) волоконно-оптических линиях связи, протянутых до оборудования конечных пользователей. Пропускная способность опорной инфраструктуры сети в городах составляет 10 Гбит/с, подключение пользователей осуществляется, как правило, каналами от 1 до 10 Гбит/с.

В других городах размещения магистральных узлов сети "последние мили" арендуются организациями-пользователями. В городах, где такие узлы отсутствуют, "последние мили", как правило, арендуются или обеспечиваются магистральными операторами связи, у которых заказываются каналы от телекоммуникационных площадок организаций до узлов операторов. В качестве оборудования доступа в составе региональных узлов НИКС используются коммутаторы с портами пропускной способностью 1 и 10 Гбит/с.

В НИКС реализовано прямое подключение сетей ведущих российских исследовательских центров в области физики высоких энергий — участников НИЦ КИ и ОИЯИ (г. Дубна) с пропускной способностью по 10 Гбит/с.

Сеть участвует в межсетевом обмене трафиком на нескольких узлах (AMS-IX, MSK-IX, SPB-IX, NSK-IX, DATA-IX и др.), имеет некоммерческие пиринговые соединения с большинством крупных российских и рядом зарубежных операторов связи (>30 соединений, >200 Гбит/с в сумме).

3. Вопросы обеспечения функционирования и развития НИКС

Приведем здесь для общего понимания современного состояния некоторые существенные особенности передовых NREN мира [6, 7, 10]. Важной характеристикой NREN, в значительной степени определяющей направления и интенсивность развития таких сетей, является их целевая аудитория. Сегодня большинство образовательных и научных организаций во многих странах подключены не к сетям коммерческих операторов связи, а к своим локальным NREN, эксплуатируя их ресурсы и сервисы.

Согласно актуальным данным GÉANT [7] в большинстве стран ЕС практически 100 % университетов и научных организаций подключены к NREN. В последние годы интенсифицировался процесс подключения к сетям начальных и средних школ (составляют уже более 80 % от общего числа организаций-пользователей некоторых сетей). Среди других типов пользователей можно указать на правительственные учреждения, библиотеки, музеи, больницы, организации дополнительного образования.

Отличительными особенностями NREN являются высокая пропускная способность магистральной инфраструктуры, а также специальные требования к качеству предоставляемых сервисов. На сегодняшний день типичные скорости подключения организаций к инфраструктуре NREN составляют от 1 до 10 Гбит/с, при этом около трети европейских университетов подключены к сетям на скорости 10 Гбит/с и выше.

Типичная пропускная способность магистральной инфраструктуры NREN в разных странах Европы кардинально различается, варьируясь в пределах от 1 до 600 Гбит/с. Вместе с тем, для наиболее развитых сетей (>10) этот показатель превышает 100 Гбит/с. Повсеместное использование технологии DWDM позволяет, при возникновении потребностей, наращивать пропускную способность практически в неограниченных пределах.

Для многих NREN специфично наличие собственных или арендованных внутригородских и межгородских ВОЛС, что позволяет увеличивать емкость магистральной сети до необходимых пользователям объемов и оперативно организовывать выделенные высокоскоростные каналы под проекты. Важное значение для обеспечения надежности предоставляемых NREN сервисов (особенно большого масштаба) имеет топология сети, что обусловливает ориентирование на кольцевые топологии с дублирующими путями.

Наличие современной инфраструктуры NREN и ее интенсивная эксплуатация позволяют оптимизировать затраты целевых пользователей на ИКТ-инфраструктуру, услуги передачи данных, использовать в нуждах сферы науки и образования развиваемые на базе NREN специализированные сервисы.

Механизмы финансирования NREN в разных странах различаются: некоторые из сетей целиком финансируются государством, другие в той или иной степени функционируют за счет своих пользователей, встречается также модель, основанная на нескольких источниках [7]. В любом случае, существенная часть средств прямо или косвенно поступает от государства. Годовой бюджет ведущих NREN европейских стран (в том числе Германии, Великобритании, Нидерландов) в последнее время варьируется в пределах от 40 до 60 млн евро.

В кадровом отношении наиболее укомплектованными в Европе являются NREN Чехии, Великобритании, Хорватии и Нидерландов, чей штат превышает 150 человек [7]. Такой внушительный состав объясняется постоянным расширением сервисного портфеля сетей, что требует привлечения все большего числа квалифицированных сотрудников для их развертывания, внедрения, поддержки и продвижения.

Подчеркнем здесь, что НИКС в своем развитии во многом ориентируется на опыт ведущих мировых NREN и сетевых консорциумов, учитывая, конечно, и российские особенности, в том числе географическую распределенность, неравномерную плотность населения, существенные различия в уровне развития региональных ИКТ-инфраструктур, интенсивности и результативности научно-образовательной деятельности.

Мировое экспертное сообщество ожидает к 2025 г. более чем трехкратного увеличения объемов обрабатываемых NREN научных данных. Реализация в России исследовательских проектов в рамках нацпроекта "Наука", решение задач по повышению доступности и уровня загрузки объектов научной инфраструктуры коллективного пользования ожидаемо приведет к существенному росту объемов генерируемых научных данных, потребует их передачи по сетям и распределенной обработки участниками сложившихся и новых исследовательских коллабораций.

Возрастающие в связи с этим требования со стороны научно-образовательного сообщества могут быть удовлетворены в результате комплексного решения взаимосвязанных задач опережающего развития сетевой инфраструктуры НИКС, увеличения ее пропускной способности и расширения территориального охвата, дальнейшей интенсификации взаимодействия с зарубежными NREN, развития экосистемы сервисов, необходимых для повышения эффективности и результативности научно-технической и образовательной деятельности, в интересах совершенствования телекоммуникационной связности основных участников научного,

научно-технического и инновационного взаимодействия.

Необходимо отметить, что коммерческие операторы связи ориентированы на предоставление услуг на базе стандартных технологических решений с организацией доступа к научным ресурсам и проектам через Интернет (только при наличии такой возможности). НИКС способна обеспечить пользователям существенные преимущества, включая удовлетворение особых требований к характеристикам QoS, уникальную связность с международным научно-образовательным ИКТ-пространством, специализированные сервисы и адресную техническую поддержку.

В состав ключевых мероприятий по развитию НИКС на плановый период (до 2024 г.) включено ускоренное развитие магистральной инфраструктуры сети внутри страны и за рубежом, развитие инфраструктуры доступа, создание региональных центров управления сетью, внедрение и развитие сервисов ИТ коллективного пользования, сервисов защиты инфраструктуры сети и пользователей от распределенных сетевых атак.

В отношении направлений развития магистральной инфраструктуры и инфраструктуры доступа НИКС запланированы:

- ввод в эксплуатацию (модернизация) внутрироссийских магистральных и региональных узлов связи разного типа и состава оборудования;
- ввод в эксплуатацию (модернизация) узлов доступа НИКС внутри страны;
- увеличение пропускной способности и повышение отказоустойчивости каналов связи на направлении Москва — Санкт-Петербург;
- создание отказоустойчивых кольцевых сегментов на территории Сибирского и Дальневосточного федеральных округов, а также в отдельных южных и центральных регионах (с пропускной способностью 10 Гбит/с);
- создание (модернизация) каналов связи на отдельных направлениях внутри страны в целях территориального развития и увеличения пропускной способности магистральной инфраструктуры НИКС, повышения уровня ее региональной доступности для подключения новых пользователей (с пропускной способностью от 10 до 100 Гбит/с);
- ввод в эксплуатацию (модернизация) зарубежных узлов и каналов связи в целях повышения эффективности взаимодействия с международными исследовательскими центрами и участия в глобальных проектах;
- расширение прямого межсетевого взаимодействия с зарубежными NREN и сетевыми консорциумами;
- модернизация узлов уровня ядра в Москве;
- инсталляция (аренда) городских ВОЛС в Москве, Санкт-Петербурге и ряде других крупных городов страны (от узлов доступа НИКС до пользователей).

В качестве пользователей развиваемой инфраструктурно-сервисной платформы НИКС рассматриваются:

- Минобрнауки России и иные заинтересованные ведомства;
- РАН и отраслевые академии наук;
- научные организации, организации высшего образования, в том числе базовые организации установок класса "мегасайенс", ЦКП, УНУ, СКЦ;
- организации-участницы НОЦ, НЦМУ, ЦК НТИ и других проектов, нацеленных на интенсификацию научно-образовательной и научно-технической деятельности.

Партнерами НИКС могут выступать зарубежные NREN и наднациональные сетевые консорциумы, зарубежные и международные научно-исследовательские центры и научные коллаборации, российские и международные научные фонды и институты развития, агрегаторы крупных хранилищ данных и провайдеры облачных сервисов для сферы науки и образования, производители телекоммуникационного и серверного оборудования и специализированного программного обеспечения.

Управление НИКС, как представляется авторам, должно осуществляться органом управления и координации деятельности (Минобрнауки России), представительным экспертно-координационным советом и администратором НИКС (МСЦ РАН) во взаимодействии с операторами региональных и отраслевых научно-образовательных сетей (в рамках их зон ответственности).

В настоящее время МСЦ РАН продолжает реализацию комплекса организационных и технических мероприятий по формированию единой научно-образовательной сети федерального масштаба с централизованным управлением, включая и мероприятия по верхнеуровневой интеграции в единую инфраструктуру действующих региональных и межрегиональных научно-образовательных сетей и их фрагментов.

В планах работ — предметное взаимодействие с региональными отделениями РАН, научно-образовательными кампусами, ведущими организациями науки и высшего образования страны по вопросам подключения к НИКС, выявления потенциальных потребностей в пропускной способности каналов, локальных требований к СПД, географических и иных особенностей, заинтересованности в сервисных пакетах. Весьма важным вопросом является разработка объективных критериев присоединения организаций к проекту НИКС, а также отчетных показателей, которые позволят оценить эффективность и интенсивность использования ее инфраструктуры и сервисов.

Систематическое представление текущего состояния и планов расширения сервисной платформы НИКС, а также ее совершенствования в части аппаратно-программной основы достойно отдельной публикации. Приведем только несколько примеров эксплуатируемых и развиваемых на базе сети перспективных свободных для использования сервисов с отсылками к работам, в которых можно ознакомиться с их подробным описанием.

Сервис международного роуминга в Wi-Fi сетях (проект eduroam, https://eduroam.ru) предоставляет возможности бесплатного доступа в Интернет через зоны Wi-Fi для научно-образовательного сообщества с едиными учетными данными и аутентификацией пользователей на стороне "домашней" организации [12—14].

Сервисы на основе удостоверяющей федерации НИКС RUNNetAAI (https://runnetaai.ru) [14, 15], принимающей участие в международном проекте eduGAIN [16, 17], обеспечивают идентификацию участников научно-технического взаимодействия и повсеместный доступ к востребованным научным ресурсам на базе технологий федеративной аутентификации.

Сервис вебинаров НИКС (https://vc.runnet.ru) позволяет организовывать и проводить вебинары с предоставлением возможностей ролевого доступа участников к мероприятию в качестве зрителя или модератора, "виртуальной доски", доступа к рабочему столу компьютера, средств обратной связи, создания заметок, анкетирования, проведения опросов и др.

Сервисы сбора, анализа и визуализации статистики по уровню использования телекоммуникационной инфраструктуры сети пользователями для обмена научными данными позволяют наглядно представить объемы обмена, основные направления сетевой связности и выявить устойчивые исследовательские коллаборации, включая и их зарубежных участников [18].

Принимая во внимание профильную деятельность МСЦ РАН в области суперкомпьютерных технологий, перспективным направлением работ по совершенствованию сервисной платформы НИКС является разработка и внедрение на базе ее инфраструктуры и распределенной сети СКЦ сервисов высокопроизводительных вычислений и искусственного интеллекта [19].

Заключение

К основным, прогнозируемым к концу 2024 г. количественным показателям реализации запланированных мероприятий по развитию НИКС можно отнести:

- расширение территориального охвата, обеспечение присутствия сети во всех федеральных округах и в более 60 субъектах РФ;
- рост производительности сетевой инфраструктуры НИКС (суммарной производительности маршрутизации) — в 3,5...4 раза;
- увеличение пропускной способности магистральной инфраструктуры НИКС (суммарной емкости каналов связи) — в 2,5...3 раза;
- увеличение числа организаций науки и образования, участвующих в обмене данными с использованием инфраструктуры НИКС при выполнении научных исследований и разработок — в 5 раз;
- рост объемов передаваемых научных данных в рамках сетевого взаимодействия российских организаций — в 2,5...3 раза.

В отношении социально-экономических и иных существенных результатов можно рассчитывать, в том числе, на достижение следующих:

- содействие укреплению позиций российской науки в приоритетных научных областях с использованием ресурсов НИКС при реализации на ее базе совместных проектов в рамках внутрироссийского и международного сотрудничества научных и научно-технологических коллабораций;
- поддержка интенсификации и осуществление перехода к передовым цифровым, интеллектуальным производственным технологиям, включая создание систем обработки данных больших объемов, машинного обучения и искусственного интеллекта, которые позволят получить значимые результаты в приоритетах НТР;

- внесение вклада в достижение целей и показателей нацпроекта "Наука" в части задач развития передовой инфраструктуры научных исследований и разработок, развития научной инфраструктуры коллективного пользования, создания НОЦ, НЦМУ, ЦК НТИ путем обеспечения производительной и надежной сетевой связности и предоставления уникальных сервисов;
- обеспечение высокоскоростной связности региональных научных центров для целей эффективного обмена научными данными с оптимизацией затрат на эксплуатацию посредством предоставления ресурсов НИКС в интересах создаваемой географически распределенной передовой инфраструктуры исследований и разработок, решения задач пространственного развития страны и опережающего развития приоритетных территорий;
- предоставление возможностей для прагматически обоснованной интеграции участников научно-технического процесса в международное научное ИКТ-пространство, в инфраструктуру мировых NREN, обеспечивающую доступ к новым компетенциям, идеям и технологиям, вносящую вклад в формирование интернациональной научной среды, устойчивой кооперации с мировым сообществом с учетом национальных интересов;
- разработка на базе НИКС единых политик, регламентов и организационно-технических решений, нацеленных на обеспечение информбезопасности и устойчивого функционирования ИТКИ сферы науки и образования, способных внести вклад в противодействие новым вызовам и угрозам, препятствующим решению задачи построения цифровой экономики.

Список литературы

1. **Allocchio C., Balint L., Berkhout V., Bersee J., Izhvanov Y. et al.** A History of international research networking: the people who made it happen. N. Y.: Wiley-VCH, 2010. 317 p.

2. **Lehtisalo K.** The History of NORDUnet: Twenty-five years of networking cooperation in the Nordic countries (2005). URL: http://www.nordu.net/history/book.html.

3. **GÉANT: The Case for NRENs.** A Repository of Resources to Support Funding, Advocacy and the Advancement of National and Regional R&E Networks. URL: https://www.casefornrens.org.

4. **Абрамов А. Г., Евсеев А. В.** RUNNet как национальная научно-образовательная сеть России: цели, основные задачи, телекоммуникационная инфраструктура и сервисы // Информатизация образования и науки. 2018. № 4. С. 3—15.

5. **Абрамов А. Г., Евсеев А. В.** Концептуальные аспекты создания в Российской Федерации национальной исследовательской компьютерной сети нового поколения // Информационные технологии. 2019. № 12. С. 724—733.

6. **Абрамов А. Г.** Панъевропейский научно-образовательный сетевой консорциум GÉANT: особенности инфраструктуры, ключевые проекты и сервисы // Информационные технологии. 2018. № 8. С. 546—553.

7. **GÉANT Compendium** of National Research and Education Networks in Europe — 2019 Edition. URL: https://compendium.geant.org (дата обращения: 13.11.2020).

8. **Васенин В. А.** Российские академические сети и Internet (Состояние, проблемы, решения). М.: РЭФИА, 1997. 173 с.

9. **Иванников А., Кривошеев А., Куракин Д.** Развитие сети телекоммуникаций в системе высшего образования Российской Федерации // Высшее образование в России. 1995. № 2. С. 87.

10. **Ижванов Ю. Л.** Научно-образовательные компьютерные сети. Прошлое, настоящее и тенденции развития // Образовательные ресурсы и технологии. 2017. № 2. С. 17—25.

11. **Абрамов А. Г., Евсеев А. В.** Сеть RUNNet: навстречу современным вызовам сферы телекоммуникаций в науке и образовании // Информатизация образования и науки. 2017. № 1. С. 100—115.

12. **Wierenga K., Florio L.** Eduroam: Past, present and future // Computational Methods in Science and Technology. 2005. Vol. 11(2). P. 169—173.

13. **Абрамов А. Г., Васильев И. В., Морин Ю. Н., Овсянников А. П., Порхачев В. А.** Вопросы совершенствования российского сегмента сервиса роуминга в беспроводных сетях eduroam в условиях интеграции научно-образовательных сетей RUNNet и RASNet // Труды научно-исследовательского института системных исследований РАН. 2019. № 6. С. 67—76.

14. **Абрамов А. Г., Васильев И. В., Порхачев В. А.** Развитие инфраструктуры аутентификации и авторизации для удостоверяющей федерации в рамках проектов eduGAIN и eduroam на базе сети RUNNet // ИТНОУ: Информационные технологии в науке, образовании и управлении. 2017. № 4. С. 56—64.

15. **Абрамов А. Г., Васильев И. В., Порхачев В. А.** Принципы функционирования и управления удостоверяющей федерацией RUNNetAAI в рамках интерфедеративного взаимодействия с проектом eduGAIN // Информатизация образования и науки. 2019. № 2. С. 40—47.

16. **Официальный сайт проекта eduGAIN**. URL: https://edugain.org (дата обращения: 13.11.2020).

17. **Hämmerle L., Sabatino R., Lenggenhager T. et al.** GN4-1 White Paper: Comparison of Authentication and Authorisation Infrastructures for Research. URL: https://www.geant.org/Resources/Documents/Comparison-of-AAIs-for-Research_White-Paper_v1.0.pdf (дата обращения: 13.11.2020).

18. **Абрамов А. Г., Евсеев А. В.** Мониторинг активности пользователей научно-образовательной сети России RUNNet в межсетевом взаимодействии: методики, инструментарий, результаты // Информатизация образования и науки. 2018. № 3. С. 34—49.

19. **Савин Г. И., Шабанов Б. М., Баранов А. В., Гончар А. А., Овсянников А. П.** Об использовании федеральной научной телекоммуникационной инфраструктуры для суперкомпьютерных вычислений // Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика. 2020. № 1. С. 20—35.

A. G. Abramov[1], Ph. D., Leading Researcher, e-mail: abramov@runnet.ru,
A. A. Gonchar[2], Deputy Director, e-mail: andrey.gonchar@jscc.ru,
A. V. Evseev[1], Director, e-mail: evseev@runnet.ru,
B. M. Shabanov[2], Dr. Tech. Sci., Director, e-mail: shabanov@jscc.ru,
[1] St. Petersburg Department of Joint Supercomputer Center of the RAS,
[2] Joint Supercomputer Center of the RAS Branch of FSI "Scientific Research Institute for System Analysis of the RAS"

The New Generation National Research Computer Network: Current Status and Concept for the Development

The paper systematized up-to-date information on the status of work to ensure the functioning and development plans of the new generation National Research Computer Network (NIKS), created in 2019 according to the results of integration of RUNNet and RASNet — telecommunication networks in the fields of higher education and science of Russia. The normative grounds and prerequisites for the creation of a unified research and education network in the country are presented. The key characteristics of leading foreign national research and education networks are given. The parameters of the telecommunications infrastructure and network connectivity of NIKS as of the current moment, examples of services developed in the interests of the Russian R&E community are indicated. Special attention is paid to the presentation of the main directions of the accelerated development of NIKS for 2021-2024 and expected results.

Keywords: *new generation national research computer network, NIKS, national research and education network, NREN, telecommunication infrastructure, network connectivity, network services, concept for the development*

Acknowledgements: The work was carried out at the MSC RAS within the framework of state assignment No. 0580-2021-0014.

DOI: 10.17587/it.27.115-124

References

1. **Allocchio C., Balint L., Berkhout V., Bersee J., Izhvanov Y. et al.** A History of international research networking: the people who made it happen, N. Y., Wiley-VCH, 2010, 317 p.

2. **Lehtisalo K.** The History of NORDUnet: Twenty-five years of networking cooperation in the Nordic countries, available at: http://www.nordu.net/history/book.html.

3. **GÉANT: The Case for NRENs.** A Repository of Resources to Support Funding, Advocacy and the Advancement of National and Regional R&E Networks, available at: https://www.casefornrens.org.

4. **Abramov A. G., Evseev A. V.** RUNNet as a national research and education network of Russia: goals, main tasks, telecommunication infrastructure and services, *Informatizatsiya Obrazovaniya i Nauki*, 2018, vol. 4, pp. 3—15 (in Russian).

5. **Abramov A. G., Evseev A. V.** Conceptual aspects of creating a new generation national research computer network in the Russian Federation, *Informatsionnyie Tehnologii*, 2019, vol. 2, pp. 724—733 (in Russian).

6. **Abramov A. G.** Pan-European research and education network consortium GÉANT: infrastructure features, key projects and services, *Informatsionnyie Tehnologii*, 2018, vol. 8, pp. 546—553 (in Russian).

7. **GÉANT** Compendium of National Research and Education Networks in Europe 2019 Edition, available at: https://compendium.geant.org, available at: (date of access: 13.11.20)

8. **Vasenin V. A.** *Russian academic networks and Internet (Status, problems, solutions).* Moscow, REFIA, 1997, 173 p. (in Russian).

9. **Ivannikov A., Krivosheev A., Kurakin D.** Development of the telecommunications network in the higher education system of the Russian Federation, *Vysshee Obrazovanie v Rossii*, 1995, no. 2, pp. 87 (in Russian).

10. **Izhvanov Yu. L.** Research and education computer networks. Past, present and development trends, *Obrazovatel'nye resursy i tekhnologii*, 2017, vol. 2, pp. 17—25 (in Russian).

11. **Abramov A. G., Evseev A. V.** Network RUNNet: towards the state-of-the-art challenges in the field of telecommunications in science and education, *Informatizatsiya Obrazovaniya i Nauki*, 2017, vol. 1, pp. 100—115 (in Russian).

12. **Wierenga K., Florio L.** Eduroam: Past, present and future, *Computational Methods in Science and Technology*, 2005, vol. 11(2), pp. 169—173.

13. **Abramov A. G., Vasilyev I. V., Morin Yu. N., Ovsyannikov A. P., Porhachev V. A.** Issues of improving the Russian segment of the roaming service in wireless eduroam networks in the context of the integration of RUNNet and RASNet research and education networks, *Proceedings of the Scientific Research Institute for System Analyses of the Russian Academy of Sciences*, 2019, no. 6, pp. 67—76 (in Russian).

14. **Abramov A. G., Vasilyev I. V., Porhachev V. A.** Development of the authentication and authorization infrastructure for the identity federation within the eduGAIN and eduroam projects based on the RUNNet network, *ITNOU: Informatsionnye Tekhnologii v Nauke, Obrazovanii i Upravlenii*, 2017, vol. 4, pp. 56—64 (in Russian).

15. **Abramov A. G., Vasilyev I. V., Porhachev V. A.** Principles of functioning and management of the identity federation RUNNetAAI in the framework of interfederal interaction with the eduGAIN project, *Informatizatsiya Obrazovaniya i Nauki,* 2019, no. 2, pp. 40—47 (in Russian).

16. **Official site of the project eduGAIN**, available at: https://edugain.org (date of access: 13.11.20)

17. **Hämmerle L., Sabatino R., Lenggenhager T. et al.** GN4-1 White Paper: Comparison of Authentication and Authorisation Infrastructures for Research., available at: https://www.geant.org/Resources/Documents/Comparison-of-AAIs-for-Research_White-Paper_v1.0.pdf (date of access: 13.11.20)

18. **Abramov A. G., Evseev A. V.** Monitoring of user activity of the Russian research and education network RUNNet in internetwork interaction: methods, tools and result, *Informatizatsiya Obrazovaniya i Nauki*, 2018, vol. 3, pp. 34—49 (in Russian).

19. **Savin G. I., Shabanov B. M., Baranov A. V., Gonchar A. A., Ovsyannikov A. P.** On the use of the federal research telecommunications infrastructure for supercomputer computing, *Bulletin of the South Ural State University. Series: Computational Mathematics and Informatics*, 2020, no. 1, pp. 20—35 (in Russian).

УДК 004.93’12 DOI: 10.17587/it.27.125-131

В. В. Сапунов, мл. науч. сотр., e-mail: wallowind@gmail.com,
С. А. Ботман, мл. науч. сотр., e-mail: stepan.botman@gmail.com,
Г. В. Камышов, инженер, e-mail: gv.kamyshov@mail.ru,
Н. Н. Шушарина, канд. пед. наук, руководитель службы организации НИД,
e-mail:nnshusharina@gmail.com,
Балтийский федеральный университет имени Иммануила Канта, г. Калининград

Применение свертки с периодическим граничным условием для обработки данных от цилиндрических массивов электродов[1]

Предлагается модификация искусственной нейронной сети сверточного типа для работы с электромиографическими данными, полученными от цилиндрических массивов электродов. Для учета пространственной симметрии массива операция свертки переопределяется с использованием периодических граничных условий, что позволяет построить сеть, инвариантную к вращениям массива электродов вокруг своей оси. Для проверки применимости предложенного подхода нейронная сеть, содержащая сверточный слой нового типа, была обучена на данных открытого датасета UC2018 DualMyo. При этом решалась задача классификации жестов по сигналам от одного миобраслета. Сеть, основанная на новом типе свертки, показала лучшие результаты при тестировании на данных без аугментации по сравнению со стандартными свертками, что позволило сделать вывод об инвариантности такой сети к циклическим сдвигам во входных данных.

Ключевые слова: *искусственная нейронная сеть, сверточная нейронная сеть, свертка, эквивариантность, инвариантность, циклический сдвиг, электромиография, массив электродов*

Введение

Искусственные нейронные сети сверточного типа показывают превосходные результаты при работе с изображениями и другими сигналами, которые обладают свойством локальной связности. Операция свертки (в практической реализации основных библиотек машинного обучения — кросскорреляции) позволяет эффективно выделять информацию из окрестностей каждой точки сигнала, а иерархическая структура сети, усиленная слоями субдискретизации, обеспечивает возможность глубокого обучения, при котором сеть самостоятельно определяет признаки на всех уровнях [1]. Следствием особенностей архитектуры таких сетей является наличие инвариантности по отношению к пространственным сдвигам и масштабированию обрабатываемых данных [2]. Сверточные нейронные сети успешно применяются для обработки данных различной размерности, включая изображения разных масштабов: от спутниковых снимков до цифровых микрофотографий, томографических сканов, данных спектрометрии и т.д. [3—5]. Другой перспективной областью применения является работа с временными рядами, включая распознавание и синтез звуков и речи, анализ данных с сетей датчиков, а также обработка электрофизиологических сигналов [6, 7].

Одним из перспективных приложений для сверточных нейронных сетей является обработка миографических сигналов для применения в спорте, медицине, системах управления и виртуальной реальности [8]. В последнее время с развитием электроники и компонентной базы число каналов систем регистрации электромиограммы возрастает [9, 10]. Это приводит к тому, что на каждой мышце удается разместить вместо одного электрода массив электродов, что, в свою очередь, позволяет использовать эти дополнительные данные совместно с информацией о геометрии массива для более детального анализа функционирования ин-

[1]Работа выполнена в рамках государственного задания № FZWM-2020-0013.

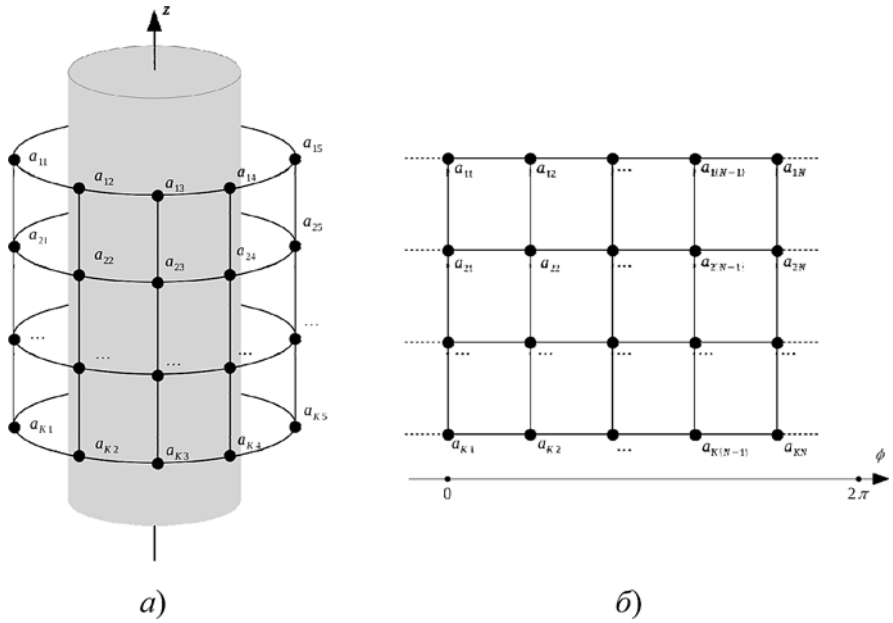
тересующей мышцы или мышечной группы. Одним из распространенных вариантов конфигурации электродов для систем распознавания жестов и движений рук является сетка электродов, свернутая в цилиндр вокруг предплечья, таким образом, что электроды равномерно распределены по цилиндру. Сверточные сети при работе способны принимать на вход сырые данные напрямую от таких массивов, что позволяет реализовывать сквозные (end-to-end) модели машинного обучения. Вместе с тем, при применении этого подхода напрямую к массивам электродов цилиндрической геометрии будут возникать проблемы, связанные с представлением данных, включая краевые эффекты [11]. Еще одной проблемой является то, что при эксплуатации неизбежно возникают сдвиги электродов от исходных позиций, а кроме того, выставление электродов в четко заданные позиции трудозатратно и поэтому неоправданно в прикладных применениях. Для устранения таких сдвигов используются дополнительные операции, такие как, например, явное вычисление угла поворота [12, 13].

В рамках данной работы предлагается компенсировать сдвиги непосредственно нейронной сетью путем наделения указанной сети инвариантностью по отношению к пространственным вращениям массивов электродов. Для этого стандартная операция свертки модифицируется с учетом периодических граничных условий. Для проверки эффективности нового типа свертки использовались открытые данные, полученные с использованием миобраслета потребительского класса с восемью парами электродов, объединенных в кольцо.

Материалы и методы

Будем рассматривать цилиндрический массив электродов, представленный на рис. 1, *а*. Такой массив может быть получен из прямоугольного массива a_{ij} размером $K \times N$ путем сшивания двух его краев (столбцов $j = 1$ и $j = N$) путем наложения периодического граничного условия: $a_{ij+N} = a_{ij}$ для $i = 1...K$, $j = 1...N$ (см. рис. 1, *б*). Электроды, расположенные на поверхности полученного цилиндра, направлены внутрь, в сторону исследуемого объекта (например предплечья человека), который расположен вдоль оси Z цилиндра. Несмотря на то что геометрия массива электродов обладает пространственной симметрией (группа вращений C_N), сам сигнал, в силу биологической природы, симметрией не обладает. Также на практике необходимость наличия надежного контакта между электродом и поверхностью кожи будет приводить к отклонениям от идеальной цилиндрической симметрии.

При работе с такими данными полезно учитывать информацию о локальной связности данных, обусловленной физической близостью соответствующих электродов. В таком случае регистрируемые за фиксированный промежуток времени данные удобно представлять в виде трехмерных массивов a_{ijt}, где $t = \overline{1,T}$ нумерует временные отсчеты. Поскольку данные локально связаны вдоль всех трех направлений массива, их можно эффективно обрабатывать с использованием нейронных сетей сверточного типа. Одним из преимуществ такого подхода является наличие эквивариантности по отношению к операциям трансляции входных данных [14], что в рассматриваемом случае будет соответствовать сдвигам по времени, сдвигам вдоль оси Z и вращениям вокруг оси Z. Проблемой при этом является тот факт, что обычная свертка не будет эквивариантна по отношению к циклическому сдвигу (т. е. вра-



а) *б)*

Рис. 1. Геометрия рассматриваемого массива электродов:

а — условное расположение электродов в пространстве относительно руки (ось руки совпадает с осью цилиндра, в котором расположены электроды); *б* — структура регистрируемых данных

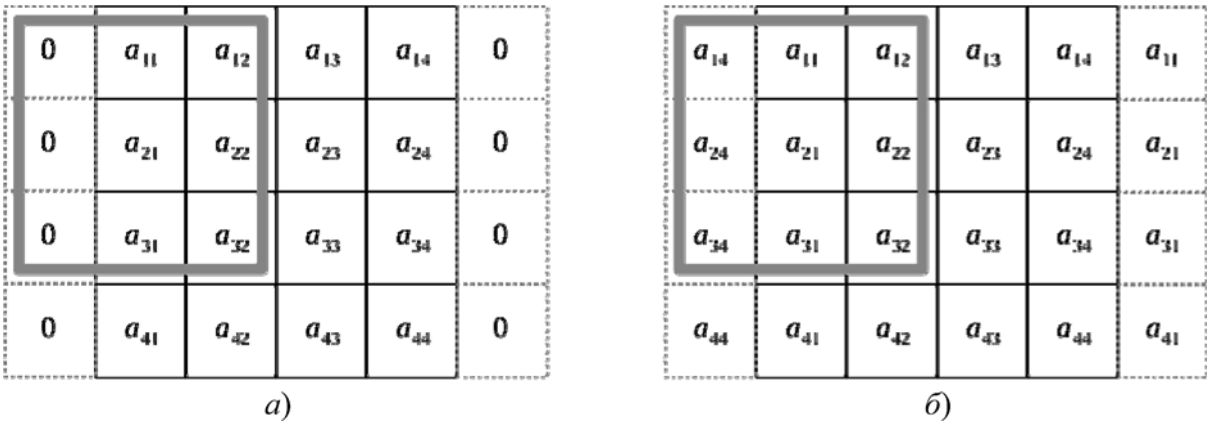


Рис. 2. Сравнение обычной свертки с паддингом (*а*) и свертки с периодическим граничным условием (*б*) на примере массива данных размером 4 × 4 и ядром 3 × 3.

щению электродных массивов вокруг оси Z), что будет приводить к появлению различного рода краевых эффектов и негативно скажется на эффективности работы сети, построенной на стандартной операции свертки.

Предлагаемым в данной работе решением указанной проблемы является введение свертки с периодическим граничным условием. Основным отличием такой свертки от обычной является работа с крайними ячейками, для которых недостающие данные добираются с противоположной стороны массива данных, как показано на рис. 2. При этом обычные свертки можно напрямую заменять свертками нового типа, тем самым адаптируя проверенные архитектуры искусственных нейронных сетей для работы с данными, имеющими циклические связи.

Основным преимуществом такого подхода является получаемая естественным образом эквивариантность отдельных сверточных слоев и, потенциально, инвариантность для всей сети. Это важное свойство, учитывая что сдвиги являются одним из основных видов артефактов для электромиографических сигналов. Альтернативно для достижения того же эффекта достаточно большую нейронную сеть можно обучить с использованием аугментации тренировочных данных. В частности, в качестве операций аугментации можно использовать вращение в пространстве на углы $2\pi/n$, где $n = \overline{1, N}$, кратные ячейке электродной сетки. В этом случае для произвольного $n < N$, где N — полное число точек массива данных вдоль угловой компоненты, новая матрица генерируется по следующему правилу: $a'_{ij} = a_{ij'}$, где $j' = (j + n) \bmod N$. Другим возможным вариантом является применение аугментации на базе вращения на произвольный угол, в которой новые данные задаются как линейная комбинация данных для соседних электродов. Несмотря на то что в теории аугментация позволяет добиться схожих результатов, тренировка на аугментированных данных повышает затраты вычислительных ресурсов в разы.

Для апробации предложенного подхода упрощенная реализация свертки с периодическим граничным условием была реализована с использованием библиотеки Tensorflow и языка программирования Python. Реализация базируется на последовательном применении операций конкатенации и обычной свертки с необходимыми параметрами. Для оценки сравнения с обычной сверткой были созданы две эквивалентные по форме нейронные сети, отличающиеся друг от друга используемым типом свертки. В качестве данных для обучения использовался открытый датасет UC2018 DualMyo Hand Gesture Dataset [15], содержащий данные ЭМГ, полученные с пары устройств Myo Armband во время последовательного выполнения человеком восьми жестов. Всего в этом датасете содержится 110 повторений по две секунды для каждого жеста. Данные снимались на частоте 200 Гц, число каналов ЭМГ — по восемь для каждого устройства, из которых для апробации использовались восемь (только первое устройство).

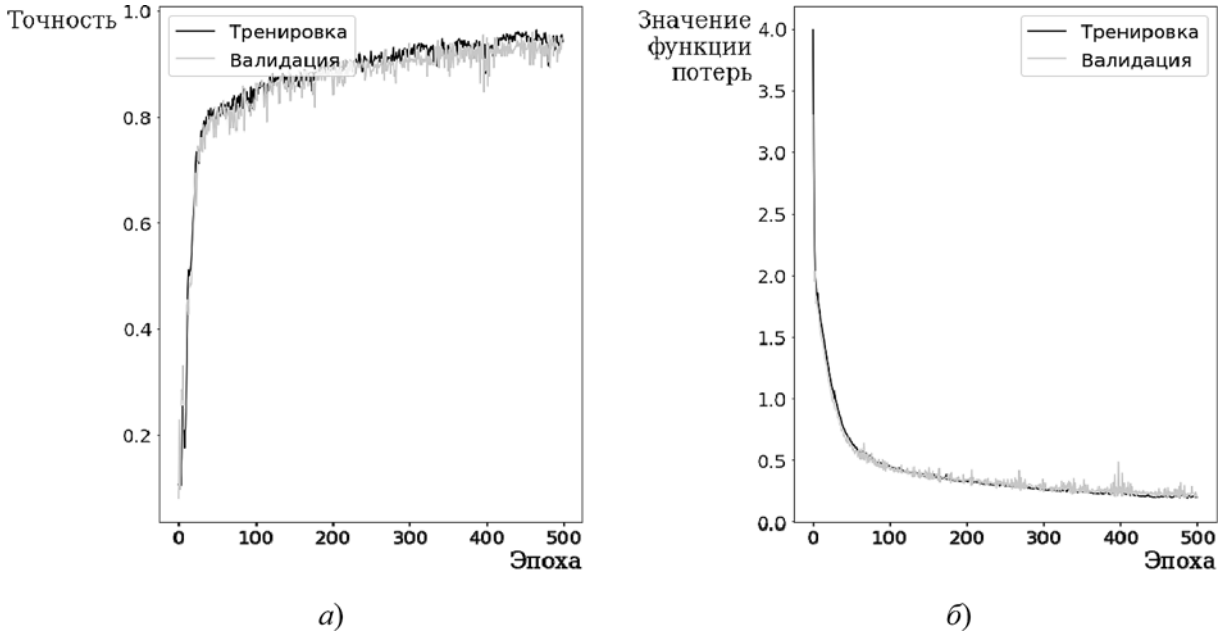
Обе нейронные сети состояли из одного сверточного слоя, глобального усредняющего слоя субдискретизации и одного полносвязного слоя. На сверточный слой с восемью ядрами размером 10 × 3 и активацией Relu подаются входные данные размером 400 × 8 × 1. Слой субдискретизации усредняет выход сверточного слоя до размеров 200 × 8 и передает эти данные на полносвязный слой из восьми нейронов с активацией Softmax. В качестве оптимизатора использовался Adam со скоростью обучения

0,01, в качестве функцией потерь была выбрана категориальная кросс-энтропия.

Для разделения данных на обучающую, проверочную и тестовую выборки в соотношении 6:2:2 использовался скрипт, поставляемый вместе с датасетом. Чтобы проверить, насколько хорошо созданная сеть на основе нового типа сверточного слоя справляется с устранением граничного разрыва между каналами ЭМГ, было проведено два испытания — с использованием аугментации исключительно тестовых данных, а также с аугментацией всех данных. Для аугментации использовалась операция вращения на кратный угол.

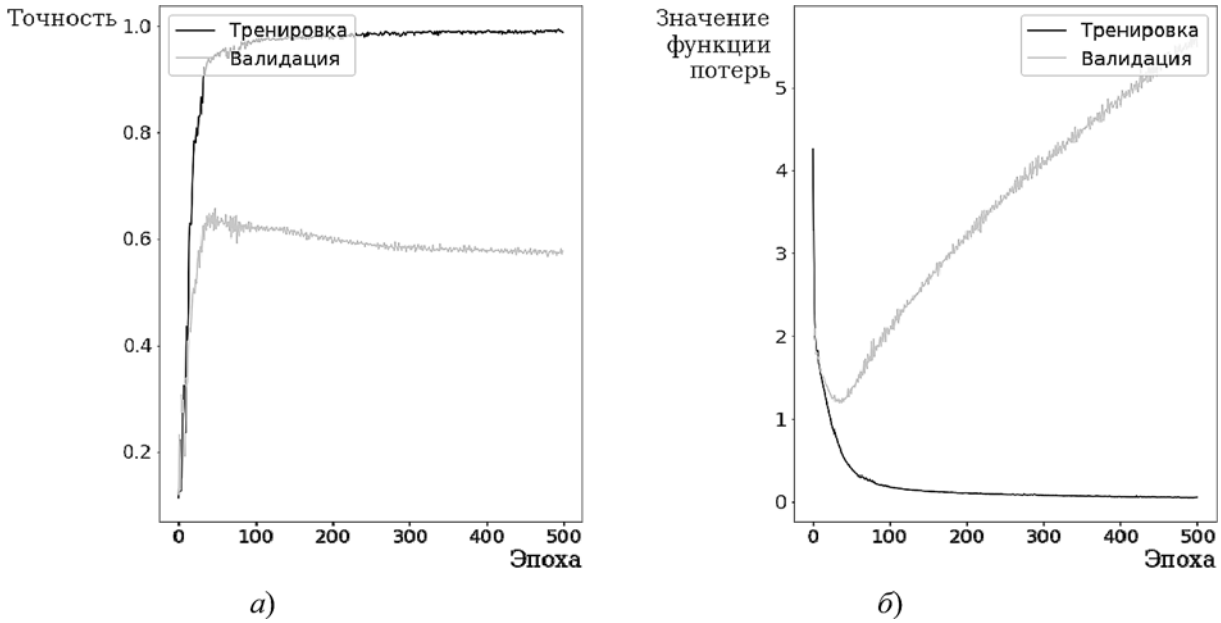
Результаты и выводы

После обучения на протяжении 200 эпох для сети с модифицированной сверткой было получено значение f1-метрики 0,96, для обычной сети — 0,65. При последующем обучении



а) *б*)

Рис. 3. Графики изменения точности (*а*) и функции потерь (*б*) для модифицированной сверточной сети в процессе обучения по данным без аугментации

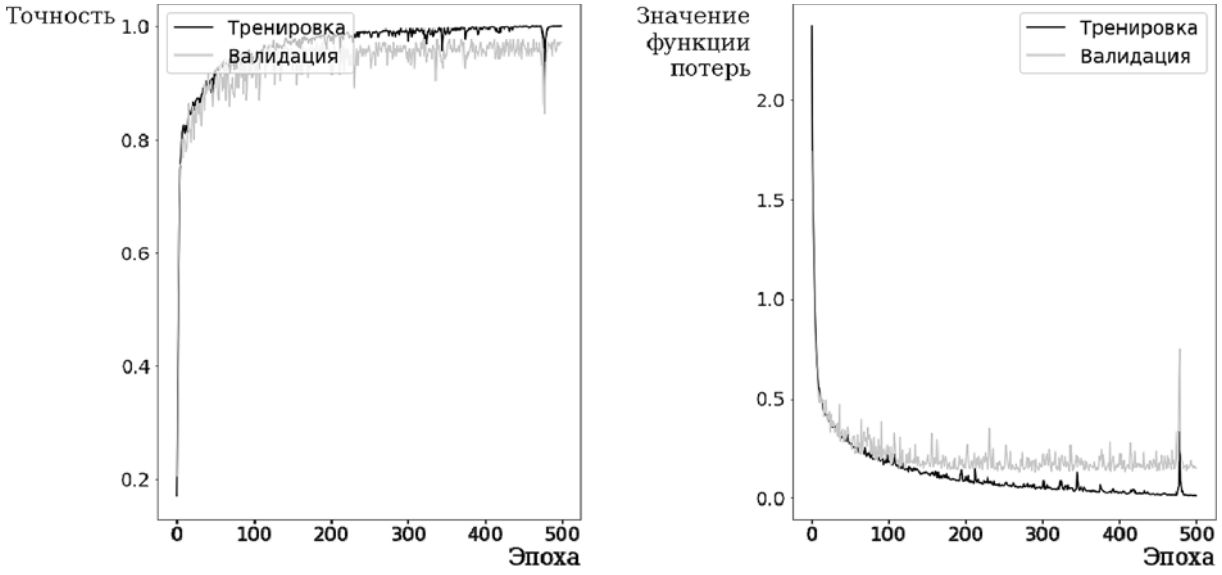


а) *б*)

Рис. 4. Графики изменения точности (*а*) и функции потерь (*б*) для обычной сверточной сети в процессе обучения по данным без аугментации

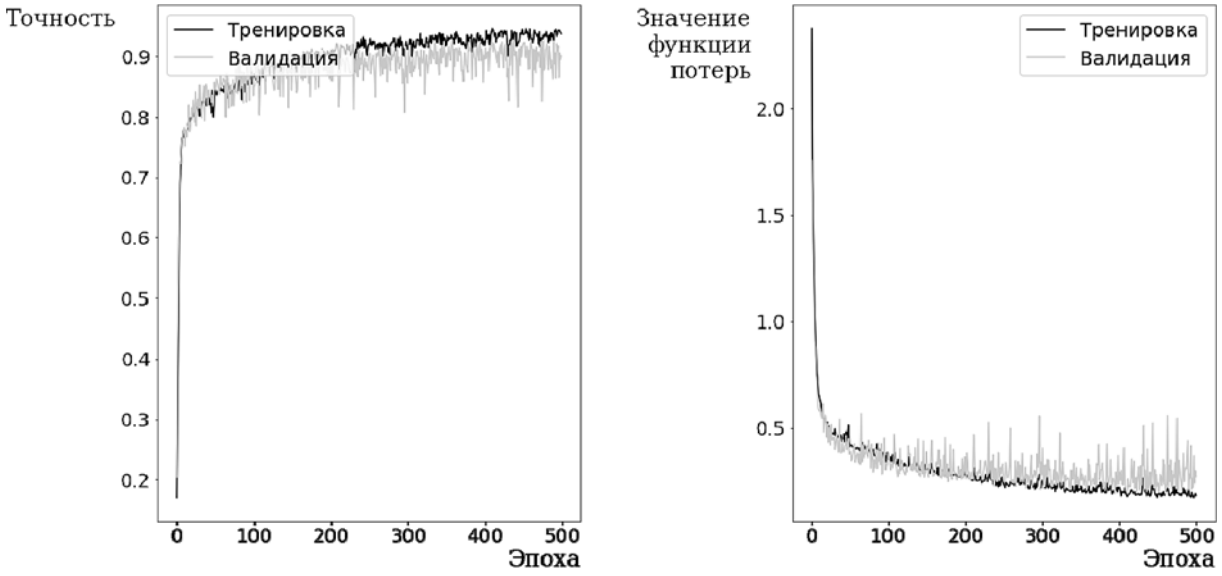
на отметке в 500 эпох было получено значение f1-метрики 0,98, в то время как результат обычной сети снизился до 0,62. Графики для метрик, рассчитанных на основе валидационных данных, представлены на рис. 3 и 4 для модифицированной и обычной сети соответственно.

Как видно из рис. 3, нейросеть на основе свертки с периодическим граничным условием успешно классифицирует аугментированные данные, несмотря на то, что тренировочные данные аугментированы не были — точность и значение функции потерь для тренировочных и проверочных данных практически идентичны. В то же время для обычной сверточной сети аугментация тестовых данных эквивалентна добавлению новых. Как видно из рис. 4, точность и значение функции потерь на тренировочных данных ведут себя так же, как у сети с модифицированной сверткой, но результаты для проверочных данных свидетельствуют о том, что вращение данных не просто затрудняет верное предсказание, но практиче-



а) *б*)

Рис. 5. Графики изменения точности (*а*) и функции потерь (*б*) для модифицированной сверточной сети в процессе обучения по аугментированным данным



а) *б*)

Рис. 6. Графики изменения точности (*а*) и функции потерь (*б*) для обычной сверточной сети в процессе обучения по аугментированным данным

ски делает его невозможным, ведь чем дольше идет обучение, тем выше значение функции потерь и ниже точность. Таким образом, вне зависимости от длительности обучения сеть со стандартной сверткой не обучается распознавать данные ЭМГ, подвергшиеся вращению.

В случае обучения на аугментированных данных сеть со стандартной сверткой достигает значения f1-метрики на уровне 0,94 после 500 эпох обучения. Сеть с модифицированной сверткой в аналогичных условиях достигала значения 0,98, что практически совпадает с результатами для этой сети при обучении без аугментации. Это служит индикатором того, что сеть на основе свертки с периодическим граничным условием способна достичь инвариантности по отношению к циклическим сдвигам (вращениям вокруг оси z) даже при обучении без использования аугментации обучающих данных. При этом следует отметить, что обучение на аргументированных данных длилось в два раза дольше, что объясняется фактическим увеличением числа сэмплов в восемь раз. Графики обучения для обеих сетей приведены на рис. 5 и 6.

Заключение

Таким образом, показано, что предложенный подход на основе сверток с периодическим граничным условием предоставляет преимущества в скорости обучения искусственной нейронной сети для решения задач анализа электромиографических сигналов с цилиндрических массивов электродов. Более точный учет геометрии и локальной связности данных позволяет получать искусственные нейронные сети, инвариантные к вращениям массива электродов. При этом можно ожидать, что для массивов электродов большего размера выигрыш за счет экономии вычислительных ресурсов может быть значительным. Вклад от таких изменений будет тем больше, чем выше соотношение числа электродов на виртуальной линии разрыва к полному числу электродов в массиве. В перспективе предложенную свертку можно применять для анализа любых данных, имеющих циклические связи такого же типа, при этом можно адаптировать уже проверенные архитектурные решения для сетей или даже целые сети, заменяя в них обычную свертку модифицированной. Используемая в данной работе примитивная реализация свертки с периодическим граничным условием не оптимальна, однако при необходимости можно создать более эффективную программную реализацию, чтобы минимизировать привносимые накладные вычислительные расходы.

Список литературы

1. **Dumoulin V., Visin F.** A guide to convolution arithmetic for deep learning // arXiv preprint arXiv:1603.07285. 2016.
2. **Tensmeyer C., Martinez T.** Improving invariance and equivariance properties of convolutional neural networks // ICLR. 2016.
3. **Pedrycz W., Chen S. M.** (ed.). Deep Learning: Algorithms and Applications. Springer, 2020.
4. **Balas V. E.** et al. (ed.). Handbook of deep learning applications. New York: Springer, 2019. T. 136.
5. **Khan A.** et al. A survey of the recent architectures of deep convolutional neural networks //Artificial Intelligence Review. 2019. С. 1—62.
6. **Wang S., Cao J., Yu P. S.** Deep learning for spatio-temporal data mining: A survey //arXiv preprint arXiv:1906.04928. 2019.
7. **Faust O. et al.** Deep learning for healthcare applications based on physiological signals: A review // Computer methods and programs in biomedicine. 2018. Vol. 161. С. 1-13.
8. **Phinyomark A., Scheme E.** EMG pattern recognition in the era of big data and deep learning //Big Data and Cognitive Computing. 2018. Vol. 2, No. 3. С. 21.
9. **Moin A. et al.** An EMG gesture recognition system with flexible high-density sensors and brain-inspired high-dimensional classifier // 2018 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2018. P. 1—5.
10. **Lara J., Paskaranandavadivel N., Cheng L. K.** Effect of Segmentation Parameters on Classification Accuracy of High-Density EMG recordings // 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2019. P. 6229—6232.
11. **Afsharipour B., Soedirdjo S., Merletti R.** Two-dimensional surface EMG: The effects of electrode size, interelectrode distance and image truncation // Biomedical Signal Processing and Control. 2019. Vol. 49. P. 298—307.
12. **Xu Z.** et al. Advanced Hand Gesture Prediction Robust to Electrode Shift with an Arbitrary Angle // Sensors. 2020. Vol. 20. N. 4. P. 11—13.
13. **Kim M., Chung W. K.** Muscle activation source model-based SEMG signal decomposition and recognition of interface rotation // 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018. P. 2780—2786.
14. **Goodfellow I., Bengio Y., Courville A.** Deep learning. MIT press, 2016.
15. **Simão M. A., Neto P., Gibaru O.** UC2018 DualMyo Hand Gesture Dataset (Version 1.0-alpha) [Data set]. 2018. URL: http://doi.org/10.5281/zenodo.1320922

V. V. Sapunov, Junior Researcher, e-mail: wallowind@gmail.com,
S. A. Botman, Junior Researcher, e-mail: stepan.botman@gmail.com,
G. V. Kamyshov, Engineer, e-mail: gv.kamyshov@mail.ru,
N. N. Shusharina, Ph.D., Head of Research Organization Department, e-mail: nnshusharina@gmail.com,
Immanuel Kant Baltic Federal University, Kaliningrad, 236016, Russian Federation

Application of Convolution with Periodic Boundary Condition for Processing Data from Cylindrical Electrode Arrays

In this paper, modification of convolutional neural networks for purposes of processing electromyographic data obtained from cylindrical arrays of electrodes was proposed. Taking into account the spatial symmetry of the array, convolution operation was redefined using periodic boundary conditions, which allowed to construct a neural network that is invariant to rotations of electrodes array around its axis. Applicability of the proposed approach was evaluated by constructing a neural network containing a new type of convolutional layer and training it on the open UC2018 DualMyo dataset in order to classify gestures basing on data from a single myobracelet. The network based on the new type of convolution performed better compared to common convolutions when trained on data without augmentation, which indicates that such a network is invariable to cyclic shifts in the input data. Neural networks with modified convolutional layers and common convolutional layers achieved f-1 scores of 0.96 and 0.65 respectively with no augmentation for input data and f-1 scores of 0.98 and 0.96 in case when train-time augmentation was applied. Test data was augmented in both cases. Potentially, proposed convolution can be applied in processing any data with the same connectivity in such a way that allows to adapt time-tested architectural solutions for networks by replacing common convolutions with modified ones.

Keywords: *artificial neural network, convolutional neural network, convolution, equivariance, invariance, cyclic shift, electromyography, electrodes array*

Acknowledgements: The work was performed within the framework of the state assignment No. FZWM-2020-0013.

DOI: 10.17587/it.27.125-131

References

1. **Dumoulin V., Visin F. A** guide to convolution arithmetic for deep learning, arXiv preprint arXiv:1603.07285, 2016.

2. **Tensmeyer C., Martinez T.** Improving invariance and equivariance properties of convolutional neural networks, *International Conference on Learning Representations, Toulon*, 2016.

3. **Pedrycz W., Chen S. M.** Deep Learning: Algorithms and Applications, Springer, 2020.

4. **Balas V. E., Roy S. S., Sharma D., Samui P.** Handbook of deep learning applications (Vol. 136), New York, Springer, 2019.

5. **Khan A., Sohail A., Zahoora U., Qureshi A. S.** A survey of the recent architectures of deep convolutional neural networks, *Artificial Intelligence Review*, 2019, pp. 1—62.

6. **Wang S., Cao J., Yu P. S.** Deep learning for spatio-temporal data mining: A survey, arXiv preprint arXiv:1906.04928, 2019.

7. **Faust O., Hagiwara Y., Hong T. J., Lih O. S.** Deep learning for healthcare applications based on physiological signals: A review, *Computer Methods and Programs in Biomedicine*, 2018, Vol. 161, pp. 1—13.

8. **Phinyomark A., Scheme E.** EMG pattern recognition in the era of big data and deep learning, *Big Data and Cognitive Computing,* 2018, vol. 2, no. 3, p. 21.

9. **Moin A., Zhou A., Rahimi A., Benatti S., Menon A., Tamakloe S., Ting J., Yamamoto N., Khan Y., Burghardt F., Benini L.** An EMG gesture recognition system with flexible high-density sensors and brain-inspired high-dimensional classifier, *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, IEEE, 2018, pp. 1—5.

10. **Lara J., Paskaranandavadivel N., Cheng L. K.** Effect of Segmentation Parameters on Classification Accuracy of High-Density EMG recordings, *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2019, pp. 6229—6232.

11. **Afsharipour B., Soedirdjo S., Merletti R.** Two-dimensional surface EMG: The effects of electrode size, interelectrode distance and image truncation, *Biomedical Signal Processing and Control*, 2019, Vol. 49, pp. 298—307.

12. **Xu Z., Shen L., Qian J., Zhang Z.** Advanced Hand Gesture Prediction Robust to Electrode Shift with an Arbitrary Angle, *Sensors*, 2020, Vol. 20, No. 4, pp. 11—13.

13. **Kim M., Chung W. K.** Muscle activation source model-based SEMG signal decomposition and recognition of interface rotation, *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),* IEEE, 2018, pp. 2780—2786.

14. **Goodfellow I., Bengio Y., Courville A.** Deep learning, MIT press, 2016.

15. **Simão M. A., Neto P., Gibaru O.** UC2018 DualMyo Hand Gesture Dataset (Version 1.0-alpha) [Data set], 2018.

УДК 621.391 DOI: 10.17587/it.27.132-137

С. В. Дворников, д-р техн. наук, проф., e-mail: practicdsv@yandex.ru,
А. В. Пшеничников, д-р техн. наук, доц., e-mail: siracooz77@mail.ru,
С. С. Манаенко, канд. техн. наук, e-mail: manaenkoss@mail.ru,
И. Н. Глухих, адъюнкт, e-mail: gluxix.86@bk.ru,
Военная академия связи имени Маршала Советского Союза С. М. Буденного, г. Санкт-Петербург

Формирование сигнальных конструкций сложных структур с высоким уровнем неопределенности их параметров

Целью статьи является формализация принципов и подходов к формированию и оценке эффективности сигнальных конструкций сложных структур. Авторами поставлена и решена задача формализованного представления полученных решений, проведен критический анализ результатов, полученных разными методами.

В ходе исследования определены свойства сигналов сложных структур. Сформулировано и доказано утверждение, позволяющее обосновать повышение структурной скрытности формируемых радиосигналов. Новизной подхода является использование теории случайных графов, что обеспечило уменьшение сложности доказательной базы, этапов имитационного моделирования.

Ключевые слова: *сигнальная конструкция сложной структуры, помехозащищенность, структурная скрытность, случайный граф*

Введение

Формирование глобального информационного пространства приводит к доступности трафика со стороны, в том числе, и нелегитимных пользователей. В таких условиях для ограничения доступа к информационным ресурсам широко используют элементы теории кодирования и шифрования.

Однако методы криптозащиты, обеспечивающие надежную защиту информации, предполагают наличие сертификационных документов, что не всегда удобно в обычной повседневной деятельности [1, 2].

Гораздо более интересным видится применение методов, снижающих возможности идентификации параметров систем передачи информации на физическом уровне [3—5], что исключает автоматический доступ нелегитимных пользователей, использующих для мониторинга общедоступные средства радиоприема.

С учетом указанных обстоятельств в настоящей статье представлен оригинальный подход к формированию сигнальных конструкций сложной структуры с высоким уровнем неопределенности их параметров для нелегитимного пользователя.

Свойства сигналов сложной структуры

В настоящее время существуют различные подходы к формированию сигналов с неопределенностью их параметров [6—10]. Наиболее рациональные из них базируются на совмещении принципов кодирования и формирования радиосигналов [7—9], что обеспечивает неопределенность их параметров уже на физическом уровне [11, 12]. Сущность такого подхода основана на расширении базы формируемых сигналов. Но при этом сам принцип расширения спектра существенно отличается от общепринятой методологии формирования широкополосных сигналов [3—5, 10].

Теоретической основой предлагаемого подхода является синтез сигнальных конструкций сложных структур (СКСС), использующих более широкую полосу частот, чем требуется для передачи вложенной в них информации [13, 14]. Применение таких СКСС затрудняет их обра-

ботку на приеме. Это объясняется неоднозначностью принятия решения о значениях их параметров при нелегитимном доступе ввиду значительной вариативности доступных комбинаций доступного вида модуляции и используемого кода. Поскольку данное направление получило активное развитие относительно недавно, то общие теоретические принципы синтеза СККС еще полностью не разработаны.

Действительно, в наиболее близких по своей сути к указанному подходу работах [11, 15—17] представлены способы определения параметров СККС на основе энтропийных оценок. Однако используемые при этом принципы, модели и методы синтеза СККС не определены в полной мере. Более того, не выявлены даже отличительные признаки СККС, что и обусловливает сложность их разработки на физическом уровне [1, 10, 13].

С учетом сделанных замечаний рассмотрим теоретические подходы к построению обобщенной модели СККС.

В качестве исходных данных определим непрерывный канал с пространством финитных сигналов $S(t)_{ij}$, $0 \leqslant t \leqslant T$, конечной энергии

$$\int_{-\infty}^{\infty} |S(t)|^2 \, dt = E < \infty, \tag{1}$$

и метрикой, определяющей их различия,

$$d[S(t)_{ij}] = \sqrt{\int_{-\infty}^{\infty} \left|S(t)_i - S(t)_j\right|^2 dt}. \tag{2}$$

На первом этапе модель СККС представим в виде матрицы координат ее сигнальных точек с позиций теории плотной упаковки заданного объема V шарами одинакового радиуса [13]. Отметим, что метрические характеристики СККС в формализованной задаче могут отличаться от оптимальных, представленных в работах [10, 13, 15], что и определяет новизну разрабатываемой модели.

В соответствии с теоремой Шеннона—Хартли о пропускной способности канала [13] будем полагать, что информационная последовательность кодируется на физическом уровне сигналами с алфавитом M, определяемом их энтропийными характеристиками. Представим множество введенных многомерных сигналов M в виде функций евклидового пространства, полученных на основе разложения в конечномерной модификации обобщенного ряда Фурье [10].

Для этого в соответствии с процедурой ортогонализации Грама—Шмидта [13] определим тригонометрические функции ряда Фурье ортонормированным базисом Ψ_i, $i = \{1, N\}$. Далее в выбранном базисе определим сигнал $S_r(t)$, $r = \{1, M\}$, $0 \leqslant t \leqslant T$, коэффициентами разложения s_{ri} (проекциями на оси координат):

$$S_r(t) = \sum_{i=1}^{N} s_{r,i} \Psi_i(t). \tag{3}$$

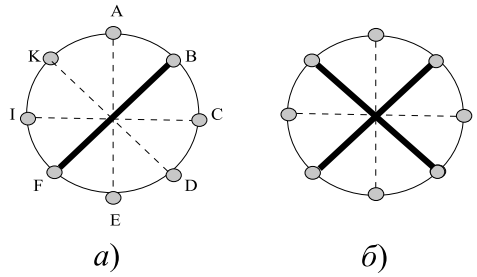
Разложение сигнала в конечную сумму ряда (3) определим в качестве основы для синтеза СККС, выбрав при этом следующие показатели:

— энергетической эффективности [18] $\sum_{i=1}^{N} s_{r,i}$ (геометрическое расположение сигнальных точек в N-мерном евклидовом пространстве такое, что при сохранении требуемой средней вероятности символьной ошибки получаем сигнал известной структуры, а формируемый сигнал будет определять соответствующий выбор сигнальных точек);

— спектральной эффективности [18] $\sum_{i=1}^{N} \Psi_i(t)$ (разработка новых ортонормированных систем базисных функций, определяющих частотно-временное распределение энергии сигнала в занимаемой полосе частот, т.е. формирование сигналов под известный спектр).

В качестве примера рассмотрим сигнальную конструкцию (СК) ФМ-8 в фазовой плоскости. Данная СК может быть получена путем объединения четырех сигналов ФМ-2 (*BPSK*), либо двух СК ФМ-4 (*QPSK*) (рис. 1) [8].

Заметим, что результирующая конструкция получена при условии ортогонального расположения формирующих ее сигналов, т. е. при условии, что сигнал *BPSK* определяется парами точек сигнального созвездия (A, E); (B, F); (C, I); (D, K) (рис. 1, *а*).



а) *б*)

Рис. 1. Возможное расположение сигнальных точек в СК при n = 8:
а — BPSK; б — QPSK

Вместе с тем, сигнал ФМ-8 можно представить как совокупность сигналов *BPSK* следующей структуры: (B, C); (D, E); (F, I); (K, A). Заметим, что такое представление не является единственно возможным. Данное обстоятельство позволяет сформулировать гипотезу, определяющую свойство СККС, связанное с множественностью формирующего базиса, что является отличительным свойством таких сигналов.

Гипотеза: для ортогональных сигналов число возможных комбинаций формируемых сигнальных конструкций L из n точек на фазовой плоскости определим в виде

$$L = \sum_{a=1}^{n/2} C^1_{n/2^{a-1}} = \sum_{a=1}^{n/2} \frac{n/2}{2^{a-1}}, \qquad (4)$$

где n — общее число точек созвездия ($n/2$ — учитывает симметрию сигнальной конструкции); C_n^m — число сочетаний из n по m.

Очевидно, что сложность оценки таких сигналов существенно возрастает с увеличением позиционности формирующей фазовой плоскости, определяемой значением C_n^m, поскольку увеличивается число комбинаций формирующих сигнальных точек.

Далее представим синтезируемые СККС ребрами графа между сигнальными точками формирующей фазовой плоскости. Тогда для сигнала *BPSK* текущий параметр *"a"* представляет собой одно ребро графа между любыми двумя ортогональными точками (рис. 1, *а*), для *QPSK* — таких ребер уже будет два (рис. 1, *б*), а для ФМ-8 их число будет равно 4. Таким образом, число возможных комбинаций из восьми точек формирующего сигнального созвездия для СК *BPSK* составляет: $\frac{n/2}{2^{a-1}} = \frac{8/2}{2^{1-1}} = 4$. Для *QPSK* таких комбинаций уже будет две.

Общее суммарное число комбинаций различных сигнальных конструкций из восьми точек созвездия равно $L = \sum_{a=1}^{n/2} \frac{n/2}{2^{a-1}} = 7$.

Поскольку для нелегитимного пользователя синтезируемая СК будет случайной, то это позволяет определить вероятность проявления каждого из возможных ребер. Таким образом, присвоив каждому из ребер значение вероятности p, получаем граф, характеризуемый свойством случайности его проявления [19]. Следовательно, применение теории графов обеспечивает уменьшение сложности формализации модели СККС за счет укрупнения формирующих базисных характеристик элементарных сигналов.

Теоретическое обоснование структурной скрытности СККС

В общем случае вероятность случайного события определения СК *BPSK* из совокупности сигнальных точек формирующего фазового созвездия ФМ-8 — это вероятность такого события, при котором из восьми вершин определяется (физически используется) только одно ребро — p^1, а остальные $n/2 - 1$ вершины не задействуются для синтеза искомого сигнала. Соответственно, вероятность такого события составляет $(1 - p)^3$.

Так как ребра графа формируются независимо, то результирующая вероятность правильного обнаружения такой сигнальной конструкции может быть представлена распределением вида $P_{BPSK} = p^1(1 - p)^3$. Рассуждая аналогичным образом, получаем, что вероятность обнаружения СК *QPSK* со случайной конструкцией точек формирующей фазовой плоскости в СККС из восьми точек составит $P_{QPSK} = p^2(1 - p)^2$.

Таким образом, формализуется следующее ***утверждение***: *повышение структурной скрытности сигнала достигается на основе увеличения базиса его формирования за счет увеличения размерности формирующей фазовой плоскости.*

Докажем сформулированное утверждение.

Обозначим $V = \{1, ..., n\}$ — множество вершин графа $G(n, p)$, где $p = \{p_1, ..., p_N\}$ — вероятности проявления ребер, $N = C_n^2$, причем $p_1 + p_2 + ... + p_N = 1$.

Осуществим выбор ребра из множества N. Будем полагать данный выбор независимым от результатов предшествующих исходов. Тогда для описания такой схемы воспользуемся аналитическим разделом теории вероятности на основе биноминального распределения. Согласно схеме Бернулли получаем случайный граф $G(V, E)$ где E — множество ребер [19]. Другими словами, определяется вероятностное пространство $G(n, p)$, в котором

$$P(G(n, p)) = p^E(1 - p)^{N-E}. \qquad (5)$$

Согласно работе [19] вероятность p характеризуется функцией, зависящей от числа вершин n, образующих сигнальные точки формирующей фазовой плоскости, т.е. $p = X(n)$. В свою очередь, указанная функция является параметром функции графа $X(G)$, математическое ожидание которой определяется следующим образом:

$$MX = \sum_G X(G)P(G). \qquad (6)$$

Здесь функция $X(G)$ определяет совокупность изолированных вершин (без текущих связей), а $P(G)$ — соответствующие вероятности перехода (вероятности ребер). В таком представлении величина MX характеризует степень сложности вскрытия структуры формируемого сигнала с позиций вероятности проявления сигнальных точек формируемой фазовой плоскости.

С учетом свойств линейности математического ожидания определим параметр c связности графа как [19]

$$c = \frac{n}{M}, \qquad (7)$$

где n — общее число вершин графа; M — число точек формирующей СК.

С учетом результатов, полученных в работе [20], выражение (6) представим в виде

$$MX = n\mathbf{e}^{-\frac{c\ln(n)(n-1)}{n}}. \qquad (8)$$

Из (8) (рис. 2) следует, что повышение структурной скрытности достигается при увеличении числа вершин графа, а также за счет увеличения потенциально возможного числа формирующих сигналов. Таким образом, утверждение доказано.

Разработанные теоретические результаты являются обобщением аналитического описания СКСС. При этом новизной представленного подхода является переход к рассмотрению формирующей фазовой плоскости в трехмерном пространстве, которая используется для синтеза СКСС с заданными свойствами. Заметим, что переход к теории графов позволяет уменьшить степень сложности решаемых задач по расчету вероятности проявления СКСС

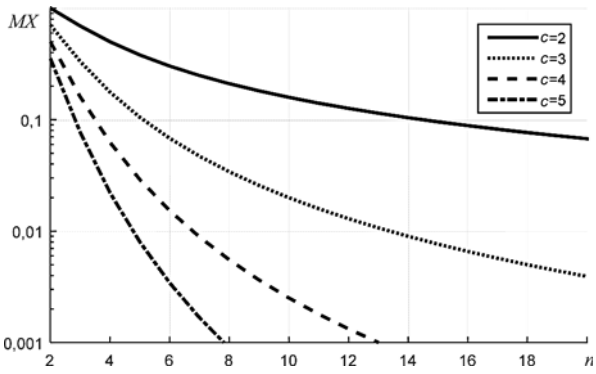


Рис. 2. Графики зависимости величины *MX* от числа вершин графа при различных значениях параметра связности *c*

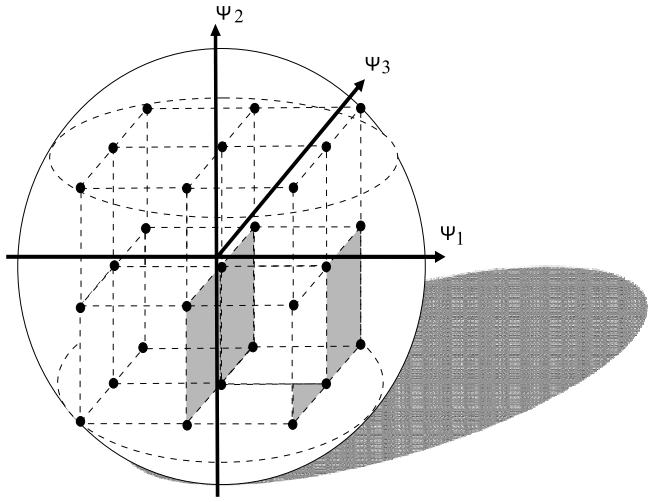


Рис. 3. Формирующая фазовая сфера

в зависимости от размерности формирующей фазовой плоскости.

С учетом объемного представления сигнальных точек СКСС в многомерном евклидовом пространстве в качестве формирующей будет выступать фазовая сфера (рис. 3). Поэтому далее целесообразно перейти к понятию формирующей фазовой сферы (ФФС).

Объемная модель в евклидовом пространстве СКСС характеризуются набором координат d_{ij}, φ_{ij}, E_m, что позволяет независимо от их формы оперировать понятием позиционности M.

Следует отметить, что с учетом формализованного геометрического представления показатель структурной сложности сигнала, введенный в работах [12, 13], преобразуется к виду

$$S_{[\text{ДИЗ}]} = \log_2\left(\prod_{i=1}^{r} k_i\varphi_i\right), \qquad (9)$$

где r — число уровней амплитуды сигнала; k — возможное число сигнальных точек на одном из уровней; φ — возможное число фазовых сдвигов в пределах одного уровня решетки ФФС.

Заключение

В целях проверки адекватности полученных решений проведено имитационное моделирование по выявлению различных структур СК в структуре формирующей фазовой плоскости СК ФМ-8 в соответствии с разработанным подходом. Результаты моделирования представлены в таблице.

Результаты моделирования выявления структур сигналов

СК	*r*	*k*	φ	*S*
BPSK	1	8	4	5
QPSK	1	8	2	4
ФМ-8	1	8	1	3

Анализ полученных результатов наглядно показывает соответствие разработанных теоретических выводов и результатов моделирования. В частности (см. рис. 2), для графа с числом вершин $n = 8$ математическое ожидание изолированных вершин MX стремится к нулю при увеличении параметра связности графа c, что равносильно стремлению к нулю вероятности проявления сигнальных точек при уменьшении позиционности формирующей СК.

Кроме того, результаты моделирования, представленные в таблице, также подтверждают, что при уменьшении позиционности синтезируемой СК на основе формирующей фазовой плоскости СК ФМ-8 структурная сложность S возрастает.

Направление дальнейших исследований связано с разработкой научно-методического аппарата оценки эффективности СККС.

Список литературы

1. **Самойленко Д. В., Финько О. А.** Имитоустойчивая передача данных в защищенных системах однонаправленной связи на основе полиномиальных классов вычетов // Нелинейный мир. 2013. Т. 11, № 9. С. 647—658.

2. **Вишневский А. К., Финько О. А.** Реализация типовых функций гибридных криптосистем арифметико-логическими полиномами // Теория и техника радиосвязи. 2011, № 1. С. 32—36.

3. **Борисов В. И., Зинчук В. М., Лимарев А. Е., Шестопалов В. И.** Помехозащищенность систем радиосвязи с расширением спектра прямой модуляцией псевдослучайной последовательностью. М.: Радиософт, 2011. 550 с.

4. **Борисов В. И., Зинчук В. М.** Помехозащищенность систем радиосвязи. Вероятностно-временной подход. М.: РадиоСофт, 2008. 260 с.

5. **Борисов В. И., Зинчук В. М., Лимарев А. Е.** Помехозащищенность систем радиосвязи с расширением спектра сигналов модуляцией несущей псевдослучайной последовательностью. М.: Радио и связь, 2003. 640 с.

6. **Дворников С. В., Пшеничников А. В.** Формирование спектрально-эффективных сигнальных конструкций в радиоканалах передачи данных контрольно-измерительных комплексов // Известия высших учебных заведений. Приборостроение. 2017. Т. 60, № 3. С. 221—228.

7. **Дворников С. В., Пшеничников А. В., Манаенко С. С.** Модель фазоманипулированного широкополосного сигнала с программной перестройкой рабочей частоты // Телекоммуникации. 2017. № 9. С. 8—12.

8. **Дворников С. В., Пшеничников А. В., Манаенко С. С., Дворников С. С.** Метод формирования многопозиционных помехозащищенных сигнальных конструкций // Информационные технологии. 2017. Т. 23, № 9. С. 669—676.

9. **Дворников С. В., Пшеничников А. В.** Модель многопозиционной помехозащищенной сигнальной конструкции на основе частотно-временных матриц // Телекоммуникации. 2017. № 6. С. 22—27.

10. **Еремеев И. Ю., Старицин С. С., Поддубных Е. В.** Модели сигналов с "быстрой" псевдослучайной перестройкой рабочей частоты // Телекоммуникации. 2015. № 8. С. 20—25.

11. **Каневский З. М., Литвиненко В. П.** Потенциальная скрытность многоканальных симплексных и дуплексных радиолиний. // Вестник Воронежского государственного технического университета. 2002. № 4-2. С. 9—10.

12. **Литвиненко В. П.** Энергетическая скрытность сигналов и защищенность радиолиний: учебное пособие. Воронеж: ВГТУ, 2009. 166 с.

13. **Скляр Б.** Цифровая связь: теоретические основы и практическое применение / Пер. с англ. Е. Е. Грозы и др. М.: Вильямс, 2016. 1099 с.

14. **Xiao Z., Su L., Jin D., Zeng L.** Performance comparison of rake receivers in sc-UWB systems and ds-UWB systems // IEICE Transactions on Communications. 2010. Vol. E93-B, № 4. P. 1041—1044.

15. **Yen-Ming Chen, Yeong-Luh Ueng.** Noncoherent Amplitude/ Phase Modulated Transmission Schemes for Releigh Block Fading Channels // IEEE Trans. Com. 2013. Vol. 61, N. 1. P. 217—227.

16. **Цепков Г. В., Яковенко И. Н.** Анализ нестационарных сигналов в адаптивном секвентном базисе // Контроль. Диагностика. № 4. 2010. С. 26—31.

17. **Raphaeli D.** Noncoherent Coded Modulation // IEEE Transaction on communication. 1996. Vol. 44, № 2. Febrary. P. 172—183.

18. **Дворников С. В., Пшеничников А. В., Аванесов М. Ю.** Модель деструктивного воздействия когнитивного характера // Информация и космос. 2018. № 2. С. 22—29.

19. **Колчин В. Ф.** Случайные графы. М.: Физмат. 2004. 256 с.

20. **Kolchin A. V., Kolchin V. F.** On transition of distributions of sums of independet identically distributed random variables from one lattice to another in the generalised allocation sheme // Discrete Mathematics and Applications. 2006. Vol. 16, N. @. C. 527—540.

S. V. Dvornikov, Professor, e-mail: practicdsv@yandex.ru,
A. V. Pshenichnicov, Assistant Professor, e-mail: siracooz77@mail.ru,
S. S. Manaenko, Assistant Professor, e-mail: manaenkoss@mail.ru,
I. N. Glukhikh, e-mail: gluxix.86@bk.ru,
Military Communications Academy, St. Petersburg, 190000, Russian Federation

The Formation of Signal Construct of Complex Structures with a High Level of Uncertainty in Their Parameters

Radio communication systems with enhanced resistance to illegitimate influences are of great interest in the development of modern information technologies. The modern methods of implementing such radio systems are the technologies of cryptographic information protection, methods of direct spectrum expansion. Despite the achievements of the selected subject area, the practical implementation of radio systems of this level is hampered by models of radio communication channels. Known results in their totality do not take into account the potential capabilities of implemented modulation formats, which limits their practical application and indicates the relevance of the study. The aim of the scientific work is to formalize theoretical principles and approaches to the formation and evaluation of the effectiveness of signals of complex structure. The authors posed and solved the problem of a formalized presentation of the obtained solutions, conducted a critical analysis of the results obtained by different methods. In this case, methods of statistical radio engineering, harmonic analysis of discrete signals, random graph theories, and decision making were used. The authors conducted a simulation of the processes of manifestation of various structures of the generated signals. The coincidence of theoretical and practical results is substantiated. During the study, the properties of signals of complex structures were determined. A theoretical statement is formulated and proved that justifies the increase in the structural secrecy properties of the generated radio signals. The novelty of the approach is the use of random graph theory, which has reduced the complexity of the evidence base, the stages of simulation. The theoretical results obtained are critically evaluated in relation to the final modeling data, which made it possible to substantiate the adequacy of the results. The obtained solutions substantiate theoretical approaches to the formation of signal construct of complex structure. In contrast to the known studies, the study reveals the potential stability boundaries of radio systems based on signals with uncertain parameters. The theoretical significance of the presented results is the expansion of the practical implementation of the methods of the general theory of communication in the field of building noise-immune systems. Practical significance is based on the results of analytical and simulation modeling presented in the work.

Keywords: *signal construct of complex structure, noise immunity, structural complexity, random graph*

DOI: 10.17587/it.27.132-137

References

1. **Samoilenko D. V., Finko O. A.** Imitation-robust data transmission in secure unidirectional communication systems based on polynomial residue classes, *Nelineynyy Mir*, 2013, vol. 11, no. 9, pp. 647—658 (in Russian).

2. **Vishnevsky A. K., Finko O. A.** Implementation of typical functions of hybrid cryptosystems by arithmetic-logical polynomials, *Teoriya i Tekhnika Radiosvyazi*, 2011, no. 1, pp. 32—36 (in Russian).

3. **Borisov V. I., Zinchuk V. M., Limarev A. E., Shestopalov V. I.** Interference immunity of radio communication systems with spreading of the spectrum by direct modulation by a pseudo-random sequence, Moscow, Radiosoft, 2011, 550 p. (in Russian).

4. **Borisov V. I., Zinchuk V. M.** Interference immunity of radio communication systems. Probabilistic-time approach, Moscow, RadioSoft, 2008, 260 p. (in Russian).

5. **Borisov V. I., Zinchuk V. M., Limarev A. E.** Interference immunity of radio communication systems with the expansion of the spectrum of signals by modulation of the carrier pseudo-random sequence, Moscow, Radio and communications, 2003, 640 p. (in Russian).

6. **Dvornikov S. V., Pshenichnikov A. V.** The formation of spectrally-efficient signal structures in radio channels for data transmission of control and measuring systems, *Izvestiya Vysshih Uchebnyh Zavedenij. Priborostroenie*, 2017, vol. 60, no. 3, pp. 221—228 (in Russian).

7. **Dvornikov S. V., Pshenichnikov A. V., Manaenko S. S.** A model of a phase-shifted broadband signal with software tuning of the operating frequency, *Telekommunikacii*, 2017, no. 9, pp. 8—12 (in Russian).

8. **Dvornikov S. V., Pshenichnikov A. V., Manaenko S. S., Dvornikov S. S.** Method for the formation of multi-position noise-protected signal structures, *Informacionnye Tehnologii*, 2017, vol. 23, no. 9, pp. 669—676 (in Russian).

9. **Dvornikov S. V., Pshenichnikov A. V.** A model of a multi-position noise-protected signal structure based on time-frequency matrices, *Telekommunikacii*, 2017, no. 6, pp. 22—27 (in Russian).

10. **Eremeev I. Yu., Staritsin S. S., Poddubnykh E. V.** Signal models with "fast" pseudo-random tuning of the operating frequency, *Telekommunikacii*, 2015, no. 8, pp. 20—25 (in Russian).

11. **Kanevsky Z. M., Litvinenko V. P.** Potential stealth of multi-channel simplex and duplex radio links, *Vestnik Voronezhskogo Gosudarstvennogo Tekhnicheskogo Universiteta*, 2002, no. 4-2, pp. 9—10 (in Russian).

12. **Litvinenko V. P.** Energy signal stealth and radio link security: tutorial. Voronezh, VSTU, 2009, 166 p. (in Russian).

13. **Sklyar B.** Digital communication: theoretical foundations and practical application, Moscow, Williams, 2016, 1099 p. (in Russian).

14. **Xiao Z., Su L., Jin D., Zeng L.** Performance comparison of rake receivers in sc-UWB systems and ds-UWB systems, *IEICE Transactions on Communications*, 2010, vol. E93-B, no. 4, pp. 1041—1044.

15. **Yen-Ming Chen, Yeong-Luh Ueng.** Noncoherent Amplitude/ Phase Modulated Transmission Schemes for Releigh Block Fading Channels, *IEEE Trans. Com.*, 2013, vol. 61, no. 1, pp. 217—227.

16. **Tsepkov G. V., Yakovenko I. N.** Analysis of non-stationary signals in an adaptive sequential basis, *Control. Diagnostics*, 2010, no. 4, pp.26—31 (in Russian).

17. **Raphaeli D.** Noncoherent Coded Modulation, *IEEE Transaction on Communication*, Febrary 1996, vol. 44, no. 2, pp. 172—183.

18. **Dvornikov S. V., Pshenichnikov A. V., Avanesov M. Yu.** Cognitive character destructive model, *Informaciya i Kosmos*, 2018, no. 2, pp. 22—29 (in Russian).

19. **Kolchin V. F.** Random graphs, Moscow, Fizmat, 2004, 256 p. (in Russian).

20. **Kolchin A. V., Kolchin V. F.** On transition of distributions of sums of independet identically distributed random variables from one lattice to another in the generalised allocation sheme, *Discrete Mathematics and Applications*, 2006, vol. 16, no. 6, pp. 527—540.

УДК 004.021, 004.822 DOI: 10.17587/it.27.138-146

А. О. Корней, аспирант, e-mail: korney.alena@yandex.ru,
Е. Н. Крючкова, канд. физ.-мат. наук, доц., e-mail: kruchkova_elena@mail.ru,
Алтайский государственный технический университет им. И. И. Ползунова, Барнаул

Категоризация текстов на основе сконденсированного графа

Предлагается комбинированный семантико-статистический алгоритм аспектного анализа текстов большого объема, основанный на использовании семантического графа. Метод выделения аспектов содержит фазы выделения множества значимых слов, вычисления весов вершин семантического графа методом релаксации, фильтрации аспектов на основе градиентного метода. Рассматривается алгоритм построения таких домензависимых множеств наиболее значимых слов, которые характеризуются одинаковыми статистическими характеристиками для разных доменов и учитывают структуру и лексическое разнообразие текстов.

Ключевые слова: *семантический граф, категоризация текстов, семантико-статистический алгоритм, извлечение знаний, извлечение аспектных терминов*

Введение

Резонансные мировые события последних лет привели к увеличению количества информации в сети Интернет, в том числе криминальной, недостоверной, заказных негативных отзывов. Борьба с такого рода информацией сделала особенно актуальной задачу автоматического определения тематической направленности и краткого содержания текста, его эмоциональной окраски. Ложная негативная информация может распространяться очень быстро, поэтому разработка эффективных алгоритмов выявления такого рода информации является теоретической базой для практической реализации автоматизированных систем своевременного реагирования. Для большинства обычных информационных ресурсов с оценочными характеристиками пользователей о тех или иных событиях или явлениях, сервисных центрах или частных поликлиниках, фильмах или телепередачах эффективная автоматическая обработка позволяет быстро получать актуальное краткое содержание множества документов на одну тему или об одном объекте.

В данной работе рассматривается инструмент для категоризации текстов, обладающий относительно невысокой вычислительной сложностью. В основу инструмента положен семантический граф русского языка, содержащий обобщенные знания о мире.

Задача категоризации текстов

Задача категоризации текстов может рассматриваться как частный случай классификации и обычно включает четыре этапа: предобработка и индексация документов, уменьшение размерности пространства признаков, построение и обучение классификатора, оценка качества классификации. Первые два этапа, как правило, предполагают стандартный набор действий. Предобработка текста строится на основе токенизации, лемматизации, удаления стоп-слов и т.д. [1]; индексация — построение числовой модели документа — может быть основана на методах Bag of Words [2], Word2vec [3], TF-IDF [4], учете *n*-грамм [2] и т.д.

Вычислительная сложность алгоритмов классификации напрямую зависит от размерности пространства признаков. Поэтому разумной мерой повышения эффективности является взвешивание и уменьшение числа признаков. Для этого применяют, например, латентный семантический анализ (LSA) [3, 5], поточечную взаимную информацию (PMI) [6], линейный дискриминантный анализ (LDA) [1]. Кроме того, в работе [1] рассматриваются и другие методы: стохастическое вложение соседей с *t*-распределением (*t*-SNE), метод случайных проекций и т.д.

Наиболее важным шагом является непосредственно этап классификации. Подходы,

применяемые для построения классификаторов, очень разнообразны. Наиболее известны такие решения, как наивный байесовский классификатор (NBC) [7], классификатор на основе *k*-ближайших соседей (KNN) [8], а также метод опорных векторов (SVM) [9]. Более сложные современные решения связаны с методами машинного обучения, использованием нейросетей, LSTM [10] и т.д.

В связи с тем, что производительность является одним из критических аспектов при категоризации текстов, современные системы строятся по одному из двух принципов: без понижения размерности, но с использованием "быстрого" классификатора; с понижением размерности, но с более качественным классификатором. Второй вариант более предпочтителен, поскольку область его применения включает и те задачи, где "быстрые" классификаторы работают плохо.

В рамках данной работы рассматривается метод построения пространства признаков, основанный на анализе семантических графов. Предполагается, что для каждой категории может быть определен набор семантических подграфов (кластеров), включающих в себя данные о лексико-семантических и статистических характеристиках категории. За счет построения семантических кластеров вокруг ключевых понятий можно достичь понижения размерности, а внутренние данные подграфа (веса связей и вершин) могут использоваться в качестве весов отдельных признаков.

Семантический граф в системах информационного поиска

Современные информационно-поисковые системы работают с текстами без ограничения смыслового диапазона, поэтому используют в минимальной степени знания о мире и о языке, базируясь на статистических методах анализа. И причина одна: использование семантических знаний может существенно усложнить обработку текста. Тем не менее, стремление повысить качество информационных систем приводит к появлению современных исследований в области использования знаний при обработке документов. Например, на основе онтологий в работе [11] предлагается модель семантического поиска в больших коллекциях. Таким образом, использование систем управления знаниями не теряет актуальности, однако требует эффективных решений.

В рамках данной работы предлагается комбинированный семантико-статистический подход, когда в качестве знаний о мире используется семантический граф русского языка, автоматически построенный авторами на базе лингвистических словарей (толкового словаря и словаря синонимов). Вершинами графа являются канонические формы слов русского языка, связанные тремя типами нечетких отношений — синонимии, ассоциации и определения. Структура графа и анализ содержимого, извлеченного из словарей, приведены в работе [12]. Данные, извлеченные из общелингвистических словарей, не зависят от предметной области и при этом достаточно полны и потому могут быть использованы в системах информационного поиска в качестве источника общих знаний о мире, а информация о связях между отдельными понятиями может интерпретироваться различным образом в зависимости от конкретной задачи. Коллектив разработчиков успешно использовал этот граф для решения нескольких проблем, в том числе для семантического поиска в больших текстовых коллекциях[13], для классификации сложных изображений [14], в системах сентимент-анализа [15]. Семантические связи между словами в графе используются для расширения знаний системы, для извлечения неявной информации об отдельных лексических единицах. Такой подход позволяет скомпенсировать недостаточность статистической информации в системах, основанных только на статистике.

Состав и структура обучающих данных

При тестировании предлагаемого алгоритма были использованы два набора обучающих данных, соответствующие двум различным доменам — "Фильмы" и "Рестораны". По домену "Фильмы" использовался набор отзывов, автоматически извлеченных с сайта "КиноПоиск" [16], куда вошли положительные и отрицательные отзывы о фильмах различных жанров, эпох и с различным рейтингом. Для домена "Рестораны" использовался полный набор отзывов, опубликованный в рамках SemEval-2016 (Task 5, *Aspect based sentiment analysis*) [17]. На рис. 1 приведены численные характеристики, позволяющие оценить объем и лексическое разнообразие выбранных наборов данных.

Введем следующие обозначения: l — общее число слов в тексте; m — число уникальных канонических форм; $d = l/m$ — коэффициент постоянства корпуса слов. Величина d пока-

Рис. 1. Численные характеристики для оценки лексического разнообразия текстовых наборов

зывает, сколько слов в среднем приходится на одну каноническую форму, и чем выше это значение, тем менее лексически разнообразны тексты. Для домена "Фильмы" d = 12,825, для домена "Рестораны" d = 19,605, что позволяет оценить "Фильмы" как существенно более разнообразный домен. Структурные, количественные и семантические различия текстов из разных областей знаний требуют высокой универсальности алгоритмов обработки и извлечения информации.

Алгоритм построения конденсированного семантического графа на основе обучающей выборки

Для построения домен-специфичных подграфов в данной работе применяется метод конденсации исходного графа на основе данных, извлеченных из обучающей выборки. Алгоритм включает несколько этапов:

- фильтрация обучающих данных;
- релаксация на базе домен-специфичного каркаса и последующее отсечение;
- расчет центральностей и градиентов вершин;
- выбор ключевых терминов домена.

Каждый из перечисленных этапов строится на основе известных алгоритмов и концепций, а двукратное отсечение незначимых данных снижает вычислительную нагрузку.

Фильтрация обучающих данных

Вершины исходного семантического графа — это канонические формы слов, а дуги представляют связи между соответствующими словами и принадлежат одному из четырех типов: определение, ассоциация, синонимия, контекстная зависимость. Первые три типа дуг принадлежат семантическому графу изначально, а четвертый тип связи представляет собой контекстное усиление ассоциации, его будем достраивать в процессе обработки текста. Граф взвешенный, на первом этапе для построения каркаса в качестве веса вершин была выбрана частотность отдельных канонических форм, встречающихся в наборе документов в рамках домена. В качестве дополнительного источника информации при модификации графа использовалась частотность биграмм, сформированных из канонических форм слов. Выбор частотных характеристик биграмм обусловлен структурой исходного семантического графа: с точки зрения семантики наличие высокочастотных биграмм с большой вероятностью свидетельствует о том, что между словами биграммы в пределах выбранного домена существует ассоциативное отношение.

Очевидно, что для построения домен-специфичных подграфов следует выбирать слова и биграммы, несущие существенную смысловую нагрузку в пределах выбранной предметной области. Необходим механизм, позволяющий выбрать наиболее значимые слова с учетом разницы в структуре и составе текстов, характеризующих предметную область. Иными словами, необходимо ввести пороги отсечения по частотности для униграмм и биграмм, учитывая статистические характеристики доменов.

В квантитативной лингвистике для текстов ограниченного объема применяется эмпирический закон Хипса [18], связывающий число уникальных слов в тексте с его длиной:

$$V_R(l) = Kl^{\beta}, \qquad (1)$$

где $V_R(l)$ — число разных слов в тексте длины l; $10 \leqslant K \leqslant 100$ и $0{,}4 \leqslant \beta \leqslant 0{,}6$ — свободные параметры. Для выявления оптимального соотношения порогов частотности для доменов разной структуры переформулируем соотношение (1). Очевидно, что коэффициент K существенно варьируется от текста к тексту именно по причине разнообразия лексикона этого текста. Поэтому параметр l заменим в выражении (1) на d, где d — введенный нами коэффициент постоянства корпуса слов, тогда число значимых для домена слов, а значит, и ранг r (порядковый номер в упорядоченном по убыванию частотности списке) последнего, включаемого в рассмотрение слова, удовлетворяет условию

$r \sim 1/d^{\beta}$. В соответствии с известным соотношением [19, 20] между нормированным коэффициентом частотности c и его рангом r выполняется соотношение $c = 1/r$ или $r = 1/c$. Тогда $c \sim d^{\beta}$. Поскольку лексическое разнообразие домена в выражении (1) учтено в параметре d, мы можем считать, что при обработке разных текстов в целях получения результатов с одинаковыми статистическими характеристиками достаточно выбрать коэффициент пропорциональности в сотношении $c \sim d^{\beta}$ не зависящим от текста домена данной группы текстов (отзывы, научные статьи, ленты новостей и т.п.), тогда для двух доменов справедливо равенство

$$\frac{c_1}{c_2} = \frac{r_2}{r_1} = \frac{d_1^{\beta}}{d_2^{\beta}}. \qquad (2)$$

Выражение (2) связывает относительные частотности порогов отсечения слов, выбираемых для подграфов. Для перехода к абсолютным значениям необходимо умножить коэффициенты c_1 и c_2 на максимальные частотности f_1 и f_2 соответствующих доменов (для "Фильмов" $f_1 = 3069$, для "Ресторанов" $f_2 = 1303$). Табл. 1 демонстрирует связь абсолютных значений пороговых частотностей для разных β.

Согласно табл. 1 порог частотности отдельных канонических форм для домена "Фильмы" должен не более чем вдвое превышать порог для домена "Рестораны".

Более сложную структуру, чем униграммы, имеют n-граммы. Вероятность появления слов в одной n-грамме зависит от вероятности появления отдельных слов. Чем больше n, тем меньше вероятность совместного появления данного упорядоченного набора. В рамках данной работы авторы предлагают устанавливать соотношение порогов для n-грамм исходя из их длины и вычисленного соотношения порогов для униграмм. Пусть p_1 — соотношение пороговых значений для униграмм выбранной пары доменов, тогда для n-грамм соотношение порогов следует выбирать приблизительно равным $\sqrt[n]{p_1}$. В случае доменов "Фильмы" и "Рестораны" соотношение порогов выбирали близким к $\sqrt{1,99} \approx 1,41$. При этом в рассмотрение включали биграммы, в которых оба слова подходили под порог частотности, выбранный для униграмм.

Фактические частотности униграмм и биграмм, подсчитанные по данным обучающих выборок, действительно подчиняются закону обратной зависимости ранга слова и частотности. Для проведения экспериментов в рамках данной работы были выбраны пороги, представленные в табл. 2.

Таблица 1

Зависимость соотношения порогов от β

Параметр	Домен		Отношение
	Фильмы	Рестораны	
d	12,825	19,605	
f_{max}	3069	1303	
$f_{max} d^{0,4}$	8515,69	4284,41	**1,99**
$f_{max} d^{0,5}$	10990,71	5769,36	**1,91**
$f_{max} d^{0,6}$	14185,06	7768,99	**1,83**

Таблица 2

Экспериментальные и теоретические пороговые значения частотности для построения домен-специфических подграфов

Домен	Порог для униграмм	Порог для биграмм
Фильмы	50	16
Рестораны	25	11
Расчет	$\frac{50}{1,99} = 25,12$	$\frac{16}{\sqrt{1,99}} = 11,34$

Графики зависимости между рангом r и частотностью c униграмм, а также выбранные пороги представлены на рис. 2 и 3. Схожесть графиков для двух столь различных доменов позволяет оценить теоретические расчеты как достаточно верные и адекватные. Для доменов пороги экспериментально подобраны таким

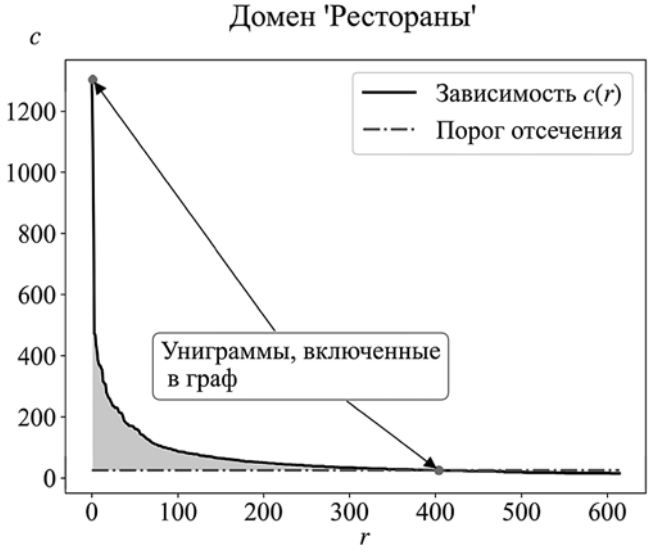


Рис. 2. Зависимость частотности униграмм от их рангов и порог отсечения для домена "Рестораны"

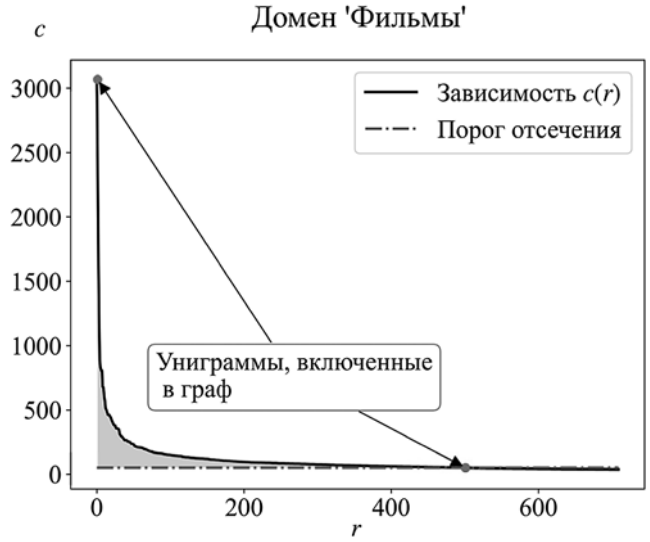


Рис. 3. Зависимость частотности униграмм от их рангов и порог отсечения для домена "Фильмы"

образом, чтобы низкочастотные и слабые по значимости биграммы и униграммы оказались за пределами рассмотрения и не включались в итоговые домен-специфические подграфы.

Релаксация на базе домен-специфичного каркаса

Пусть G — семантический граф, построенный на основе словарей. Для построения каркаса G_D домен-специфических подграфов используются слова и биграммы, отфильтрованные на основании значений из табл. 2. Пусть D — множество слов, удовлетворяющих выбранному порогу для униграмм, а B — множество биграмм, удовлетворяющих критериям выбора биграмм. Каркасы строятся на основании следующих правил:

- каждое слово $v \in D$ порождает вершину v графа G_D, которой присваивается вес $w_v = f_v$, где f_v — частота появления слова в домене;
- каждая биграмма $b \in B$ формирует двунаправленную *контекстную* ассоциативную связь в G_D между формирующими ее словами. Вес такой связи в проведенных экспериментах выбирался равным 0,8.

Полученный каркас G_D затем достраивается с использованием семантического графа G. Для этого используется процесс релаксации, в основе которого лежит классический алгоритм BFS — обход графа в ширину, модифицированный следующим образом:

- в качестве начальной вершины всегда берется очередное слово $v \in D$, которому присваивается текущий вес $\omega = \sqrt{w_v}$;
- поиск ведется параллельно по обоим графам G_D и G, G_D в ходе поиска динамически расширяется за счет включения домен-независимых вершин и связей из G. Веса w_u новых вершин u сначала принимаются равными нулю: $w_u = 0$;
- запускаем BFS от вершины v с весом ω. При посещении еще не рассмотренной очередной вершины v' ей передается релаксационный вес $\omega = r(u, v')$, где u — непосредственный предок вершины v'. Ребро (u, v'), принадлежит графу $G_D \cup G$, а $r(u, v')$ — функция релаксации;
- критерий окончания BFS: вершина не включается в очередь, если она уже была рассмотрена ранее, а также, если релаксационный вес, передаваемый ей, близок к 0 (не превышает некоторого ε).

Релаксация запускается для всех вершин из D независимо. При каждом запуске BFS соседям передаются только веса, полученные от текущей начальной вершины, а ранее накопленные веса не учитываются. Все релаксационные веса в конечном счете суммируются в итоговые значения Ω_u для всех $u = G_D$. Функция релаксации в данной работе имеет вид $r(u, v') = \alpha\omega_u f(u, v')$, где α — коэффициент затухания, не зависящий от домена; ω_u — текущий релаксационный вес вершины u и $f(u, v')$ — вес связи между u и v', определяемый как максимум из весов отношений синонимии, определения, ассоциации. В рамках данной работы эксперименты проводились для $\alpha = 0{,}5$. Веса отношений выбирались равными 0,3; 0,45 и 0,6 для определения, синонимии и ассоциации соответственно. При наличии доменной связи вес ассоциации увеличивался до 0,8. Выбор соотношения весов основан на предположении, что активное включение гипонимов и синонимов в доменный граф может понижать его специфичность, поскольку синонимия и гипонимия на общеязыковом уровне не всегда согласуется с внутридоменными отношениями.

После окончания релаксации проводится повторное отсечение малозначимых слов. В качестве метрики выбирается $p_u = \frac{\sqrt{w_u} + \Omega_u}{2}$. Порог для p_u устанавливается таким образом, чтобы после отсечения осталось не более 15 % от всех слов в графе. Фрагмент графа, полученного после релаксации и отсечения для домена "Рестораны", приведен на рис. 4.

В ходе экспериментов для домена "Фильмы" полный граф включал 4221 вершину, после отсечения — 521 (12,34 %). Для домена "Ресто-

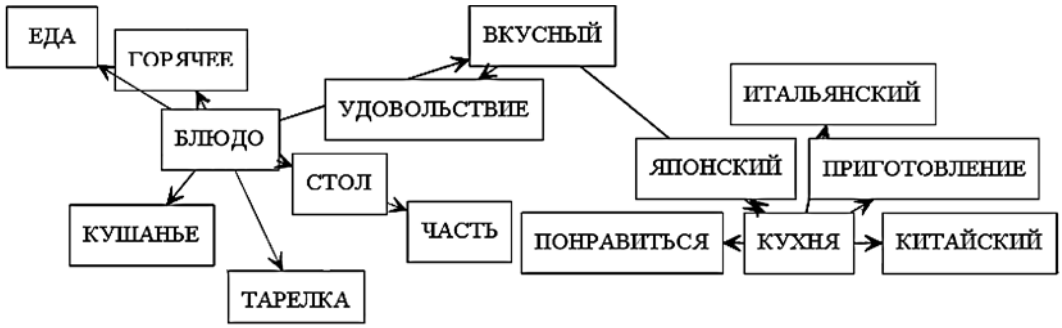


Рис. 4. Фрагмент семантического графа домена "Рестораны", образованный термином "блюдо" и его ближайшими соседями

раны" — соответственно 3142 и 253 вершины (8,05 %).

Расчет центральностей и градиентов вершин

Структура предложенного графа такова, что после релаксации на основе домен-специфической информации должны возникнуть области сгущения семантических данных за счет весов, присваиваемых вершинам. Для выявления центров сгущения авторы работы предлагают использовать две величины — центральность и градиент вершины.

Пусть $\sigma \in \{0, 1, 2...\}$ — ширина захвата контекста, т. е. при $\sigma = 0$ контекст вершины не учитывается вообще, при $\sigma = 1$ в контекст включаются вершины, непосредственно связанные с данной через исходящее ребро, — соседи первого порядка, при $\sigma = 2$ — соседи второго порядка и т. д. Использование большой ширины захвата контекста в данной задаче нецелесообразно, так как приводит к размыванию основных показателей значимости и снижению смысловой связности.

Введем величину центральности вершины. Пусть Ω_v — вес вершины, вычисленный в ходе релаксации; $G_D = (V, H)$ — домен-специфический граф с вершинами V и ребрами H. Тогда центральность вершины будем определять по формуле

$$C(v) = \begin{cases} \Omega_v, \sigma = 0; \\ \Omega_v \dfrac{\sum_{(v,u)\in H} \Omega_u}{\sqrt{\sum_{(v,u)\in H} \Omega_u^2}}, \sigma = 1; \\ \Omega_v \dfrac{\sum_{(v,u)\in H} C(u)}{\sqrt{\sum_{(v,u)\in H} C^2(u)}}, \sigma > 1. \end{cases} \quad (3)$$

Рассмотрим вектор градиента для функции $C(v)$:

$$g(v) = (C(v) - C(u_1), C(v) - C(u_2), \ldots, \ldots, C(v) - C(u_n)),$$

где $\{u_1, u_2, ..., u_n\}$ — соседи первого порядка для вершины v. Значение градиента определяется формулой

$$M_{g(v)} = \sqrt{(C(v) - C(u_1))^2 + (C(v) - C(u_2))^2 + \ldots + (C(v) - C(u_n))^2}. \quad (4)$$

В данной работе для графов, прошедших через процесс релаксации и отсечения незначимых вершин, были рассчитаны значения центральности для $\sigma = 1$ и значения градиентов. Обе вычисляемые характеристики позволяют ранжировать слова по значимости для того или иного домена, однако результаты ранжирования отличаются. В множество слов, выбранных по центральности, попадают слова, имеющие собственный высокий вес и находящиеся в центре большого "сгустка" вершин с достаточно высокими весами. Отсечение по градиенту позволяет выбрать множество вершин, окруженных большим числом соседей, но при этом обладающих собственным существенно более высоким весом по сравнению с ними. Пересечение этих множеств позволяет выбрать вершины, удовлетворяющие обеим характеристикам.

Выбор ключевых терминов домена

Для оценки согласованности ранжирования по центральности и градиенту, а также для выбора основной характеристики, позволяющей выделять ключевые термины домена, авторы работы провели ряд экспериментов.

Пусть $C(v)$ и $M_{g(v)}$ — значения функций центральности и градиента соответственно, а V_C и V_g — множества слов, отсортированных по убыванию $C(v)$ и $M_{g(v)}$. Пусть N — число наиболее важных терминов домена, которые должны остаться после отсечения по выбранной характеристике. Тогда можно говорить об абсолютном уровне отсечения — это значение центральности или градиента для последнего слова, включаемого в топ наиболее значимых. Обозначим абсолютные уровни отсечения $A_C(v_N)$ и $A_g(v_N)$ для V_C и V_g соответственно. Кроме этого, введем понятие относительного уровня отсечения: $a_C(v_N) = A_C(v_N) / \max_{v \in V_C} C(v)$ и $a_g(v_N) = A_g(v_N) / \max_{v \in V_g} M_{g(v)}$.

В ходе экспериментов для каждого из двух выбранных доменов выбирали различные значения N, а затем вычисляли число k слов, которые попали одновременно и в топ-N по центральности, и в топ-N по градиенту. Результаты экспериментов представлены в табл. 3 и 4.

Как видно из табл. 3, 4, центральность и градиент дают достаточно разный набор максимально значимых слов. Очевидно, что с ростом N согласованность растет, поскольку N приближается к полному охвату графа. Однако выбор больших N нецелесообразен, поскольку с ростом N падает значимость отдельных слов, включаемых в рассмотрение. Возникает необходимость выбора оптимального N, а также ведущей характеристики для выбора ключевых терминов домена.

Таблица 3

Оценка согласованности ранжирования по центральности и градиенту для домена "Фильмы"

N	$A_g(v_N)$	$A_C(v_N)$	$a_g(v_N)$	$a_C(v_N)$	k/N, %	k
100	457,27	77,718	0,054	0,053	56	56
70	583,054	104	0,069	0,07	49	34
50	631,624	127,641	0,075	0,086	54	27
45	642,041	135,648	0,076	0,092	53	24
40	647,329	149,817	0,077	0,101	58	23
35	711,55	163,329	0,084	0,111	60	21
30	1284,5	177,184	0,152	0,12	60	18

Таблица 4

Оценка согласованности ранжирования по центральности и градиенту для домена "Рестораны"

N	$A_g(v_N)$	$A_C(v_N)$	$a_g(v_N)$	$a_C(v_N)$	k/N, %	k
55	204,832	61,749	0,144	0,161	64	35
50	215,645	64,583	0,152	0,168	60	30
45	224,509	70,045	0,158	0,183	58	26
40	234,126	73,381	0,164	0,191	58	23
35	238,946	81,731	0,168	0,213	57	20
30	262,065	87,102	0,184	0,227	57	17

Проведем выборочный анализ терминов, попадающих в топ-20 для домена "Рестораны". В список наиболее значимых существительных при ранжировании по градиенту вошли такие слова, как: *место, кухня, время, ресторан, интерьер, блюдо, столик, обслуживание, салат, впечатление.* При ранжировании на основе центральностей в списке существительных остаются: *кухня, место, ресторан, время, интерьер, блюдо, салат, столик, впечатление, минута, день.* Очевидно, что в списки вошли значимые для оценки ресторанов термины. Объединение или пересечение списков позволяет формировать более обширные или узкие перечни аспектов, однако информативность и значимость все равно остаются на высоком уровне. При этом при ранжировании по частотности, без релаксации и расчета характеристик захват аналогичного набора терминов происходит приблизительно на уровне топ-50 (например, важный с точки зрения ABSA термин "впечатление" по частотности имеет ранг 46 и исключается из анализа).

Для домена "Фильмы" пересечение топ-40 содержит, в частности, следующий набор существительных: *фильм, время, герой, год, действие, жизнь, конец, момент, работа.* Представленный список содержит в себе ключевые термины, используемые при описании сюжета и написании рецензий. Кроме того, исходные списки топ-40 содержат такие важные слова, как *образ* (ранг по частотности 48), *сценарий* (92), *действие* (161), которые при использовании традиционного выбора по частотности в список не попадают. Таким образом, предложенный авторами подход позволяет эффективнее отбирать значимые для домена слова по сравнению с простым подсчетом частотных характеристик по домену.

Окончательное решение по выбору основного подхода к ранжированию слов может быть принято на основании условий конкретной задачи. К примеру, для задач ABSA может быть более применимо ранжирование по градиенту — на примере домена "Рестораны" такой подход позволяет извлечь разные типы лексики — для передачи смысла и отношения.

Заключение

Предложен и реализован комбинированный семантико-статистический алгоритм аспектного анализа текстовых документов, обладающий невысокой временной сложностью и направленный на комплексное решение задачи применения знаний о языке и о мире для повышения качества автоматической обработки текстов. Показано, что предложенный алгоритм позволяет значительно улучшить качество выделения категорий. Проведенные эксперименты подтверждают применимость предложенной формализованной модели для выбора соотношения порогов отсечения при обработке разных текстов в целях получения результатов с одинаковыми статистическими характеристиками. В проведенных авторами экспериментах фаза достройки семантического словаря и релаксации занимает не более 10 % от времени работы системы. Основной и наиболее затратной частью является подсчет частотных характеристик для обучающей выборки.

Список литературы

1. **Kowsari K., Jafari Meimandi K., Heidarysafa M., Mendu S., Barnes L., Brown D.** Text classification algorithms: A survey // Information. 2019. Vol. 10, N. 4. P. 150.

2. **Zhang X., Zhao J., LeCun Y.** Character-level Convolutional Networks for Text Classification // Proc. of the Neural Information Processing Systems Conf. (NIPS 2015). Montreal, Canada, 2015. URL: https://arxiv.org/abs/1509.01626 (accessed July 18, 2016).

3. **Ju R.** An Efficient Method for Document Categorization Based on Word2vec and Latent Semantic Analysis // 2015 IEEE Int. Conf. on Computer and Information Technology. 2015. P. 2276—2283.

4. **Pontiki M., Galanis D., Pavlopoulos J., Papageorgiou H., Androutsopoulos I., Manandhar S.** SemEval-2014 Task 4: Aspect based sentiment analysis // The 8th Intern. Workshop on Semantic Evaluation (SemEval 2014). Dublin, Ireland. 2014. P. 27—35.

5. **Medhat W., Hassan A., Korashy H.** Sentiment analysis algorithms and applications: a survey // Ain Shams Eng. Jour. 2014. N. 5. P. 1093—1113.

6. **Xu Y., Jones G. J., Li J., Wang B., Sun C.** A Study on Mutual Information-based Feature Selection for Text Categorization // Journal of Computational Information Systems. 2007. N. 3. P. 1007—1012.

7. **Dai W., Xue G. R., Yang Q., Yu Y.** Transferring naive bayes classifiers for text classification // In AAAI. 2007. Vol. 7. P. 540—545.

8. **Guo G., Wang H., Bell D., Bi Y., Greer K.** Using kNN model for automatic text categorization // Soft Computing. 2006. Vol. 10, N. 5. P. 423—430.

9. **Joachims T.** Text categorization with Support Vector Machines: Learning with many relevant features // NÉdellec C., Rouveirol C. (eds) Machine Learning: ECML-98. ECML 1998. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence). Vol 1398. Berlin, Heidelberg: Springer. URL: https://doi.org/10.1007/BFb0026683

10. **Luan Y., Lin S.** Research on Text Classification Based on CNN and LSTM // 2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Dalian, China. 2019. P. 352-355. doi: 10.1109/ICAICA.2019.8873454.

11. **Курейчик В. В., Бова В. В., Лещанов Д. В.** Модель семантического поиска в системах управления знаниями на основе генетических процедур // Информационные технологии. 2017. Т. 23, № 12. С. 876—883.

12. **Крайванова В. А., Кротова А. О., Крючкова Е. Н.** Математическая модель естественного языка в задачах нечеткого ассоциативного поиска // XIV Международная конференция "Речь и компьютер" (SPECOM'2011). С. 402—406.

13. **Савченко В.** Алгоритм семантического поиска в больших текстовых коллекциях" // Supplementary Proceedings of the 3rd International Conference on Analysis of Images, Social Networks and Texts (AIST'2014). P. 161—166.

14. **Казаков М. Г.** Классификация сложных изображений на основе семантического графа понятий // Прикладная информатика. 2014. Т. 54, № 6. С. 79—89.

15. **Корней А. О., Крючкова Е. Н.** Анализ тональности коротких текстов на основе семантического графа // Робототехника и искусственный интеллект. Матер. X Всеросс. науч.-техн. конф. с международным участием. 2018. С. 168—174.

16. **КиноПоиск.** Все фильмы планеты. URL: https://www.kinopoisk.ru/.

17. **Pontiki M.** et al. SemEval-2016 Task 5: Aspect Based Sentiment Analysis // Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016). 2016. P. 19-30. San Diego, California: The Association for Computational Linguistics. URL: https://doi.org/10.18653/v1/S16-1002.

18. **Heaps, Harold Stanley**. Information Retrieval: Computational and Theoretical Aspects. Inc.6277 Sea Harbor Drive Orlando, FL, United States, Academic Press, 1978.

19. **Zipf G. K.** Human behavior and the principle of least effort. Cambridge (Mass.): Addison—Wesley, 1949, pp. 573.

20. **Li W.** Random texts exhibit Zipf's-law-like word frequency distribution // IEEE Transactions on information theory. 1992. Vol. 38, N. 6. P. 1842—1845.

A. O. Korney, Postgraduate Student, korney.alena@yandex.ru,
E. N. Kryuchkova, Cand. Ph.-Math. Sc., Associate Professor, kruchkova_elena@mail.ru,
Polzunov Altai State Technical University, Barnaul, 656038, Russian Federation

Text Categorization Based on Condensed Graph

The resonant world events of 2020 led to an increase in the amount of information on the Internet, including criminal, fake news, and fake negative reviews. False negative information can spread very quickly, and methods are needed to suppress this process. The development of effective algorithms for automatic text analysis is especially relevant today. The most important

subtasks include thematic categorization, sentiment analysis, including ABSA (aspect-based sentiment analysis). The paper proposes a combined semantic-statistical algorithm for the aspect analysis of large texts, based on the use of a semantic graph. The aspect extraction method contains the phases of selecting a set of significant words, calculating the weights of the vertices of the semantic graph by the relaxation method, filtering aspects based on the gradient method. The method proposed allows to extract domain-dependent aspect terms from training data. Different aspect term sets extracted from different domains have the same statistical features, and in the same time lexical diversity and structure are taken into account.

Keywords: *semantic graph, text categorization, semantic-statistical algorithm, knowledge extraction, aspect term extraction*

DOI: 10.17587/it.27.138-146

References

1. **Kowsari K., Jafari Meimandi K., Heidarysafa M., Mendu S., Barnes L., Brown D.** Text classification algorithms: A survey, *Information*, 2019, vol. 10, no. 4, pp. 150.

2. **Zhang X., Zhao J., LeCun Y.** Character-level Convolutional Networks for Text Classification, *Proc. of the Neural Information Processing Systems Conf. (NIPS 2015),* Montreal, Canada, 2015, available at: https://arxiv.org/abs/1509.01626 (accessed July 18, 2016).

3. **Ju R.** An Efficient Method for Document Categorization Based on Word2vec and Latent Semantic Analysis, *2015 IEEE Int. Conf. on Computer and Information Technology*, 2015, pp. 2276—2283.

4. **Pontiki M., Galanis D., Pavlopoulos J., Papageorgiou H., Androutsopoulos I., Manandhar S.** SemEval-2014 Task 4: Aspect based sentiment analysis, *The 8th Intern. Workshop on Semantic Evaluation (SemEval 2014*), Dublin, Ireland, 2014, pp. 27—35.

5. **Medhat W., Hassan A., Korashy H.** Sentiment analysis algorithms and applications: a survey, *Ain Shams Eng. Jour.*, 2014, no. 5, pp. 1093—1113.

6. **Xu Y., Jones G. J., Li J., Wang B., Sun C.** A Study on Mutual Information-based Feature Selection for Text Categorization, *Journal of Computational Information Systems*, 2007, no. 3, pp. 1007—1012.

7. **Dai W., Xue G. R., Yang Q., Yu Y.,** Transferring naive bayes classifiers for text classification, *In AAAI*, 2007, vol. 7, pp. 540—545.

8. **Guo G., Wang H., Bell D., Bi Y., Greer K.** Using kNN model for automatic text categorization. Soft Computing, 2006, vol. 10, no. 5, pp. 423—430**.**

9. **Joachims T.** Text categorization with Support Vector Machines: Learning with many relevant features, NÉdellec C., Rouveirol C. (eds) *Machine Learning: ECML-98. ECML 1998. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence),* 1998, vol 1398, Springer, Berlin, Heidelberg, available at: https://doi.org/10.1007/BFb0026683.

10. **Luan Y., Lin S.** Research on Text Classification Based on CNN and LSTM, *2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA),* Dalian, China, 2019, pp. 352—355, doi: 10.1109/ICAICA.2019.8873454.

11. **Kureichik V. V., Bova V. V., Leshchanov D. V.** Semantic search model for knowledge management systems based on genetic procedures, *Information Technology*, 2017, vol. 23, no. 12, pp. 876—883.

12. **Krotova A, Krayvanova V., Kryuchova E**. Mathematical model of natural language for fuzzy associative search tasks, *SPECOM 2011 Proceedings*, Kazan, 2011, pp. 402—406 (in Russian).

13. **Savchenko V.** Semantic Search Algorithms in Large Text Collections, *in Supplementary Proceedings of AIST* 2014, pp. 161—166 (in Russian).

14. **Kazakov M., Kruchkova E.** Classication of complex images based on semantic graph, *Journal of Applied informatics*, vol. 6, no. 54, pp. 79—89 (in Russian).

15. **Korney A., Kryuchkova E0,** Short text sentiment analysis based on semantic graph, *ROBOTICS AND ARTIFICIAL INTELLIGENCE X Proceedings*, Zheleznogorsk, 2018, pp. 168—174 (in Russian).

16. **КиноПоиск.** Все фильмы планеты, available at: https://www.kinopoisk.ru/.

17. **Pontiki M.** et al. SemEval-2016 Task 5: Aspect Based Sentiment Analysis // Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), 2016, pp. 19-30. San Diego, California: The Association for Computational Linguistics, available at: https://doi.org/10.18653/v1/S16-1002.

18. **Heaps Harold Stanley.** Information Retrieval: Computational and Theoretical Aspects. Academic Press, Inc.6277 Sea Harbor Drive Orlando, FL, United States, 1978.

19. **Zipf G. K.** Human behavior and the principle of least effort. Cambridge, Mass., Addison—Wesley, 1949, pp. 573.

20. **Li W**. Random texts exhibit Zipf's-law-like word frequency distribution, *IEEE Transactions on information theory,* 1992, vol. 38, no. 6, pp. 1842—1845.

УДК 004.056

П. А. Шиловских, советник, e-mail: pashilovskikh@hse.ru,
Национальный исследовательский университет "Высшая школа экономики", г. Пермь, Россия

Протокол цифровой подписи на основе криптокодовых конструкций

Развитие вычислительных ресурсов в постквантовый период ставит под сомнение обеспечение требуемого уровня стойкости алгоритмов симметричной и несимметричной криптографии. Появление полномасштабного квантового компьютера на основе алгоритмов Шора и Гровера значительно увеличивает возможности киберпреступников и снижает стойкость используемых криптосистем в протоколах обеспечения основных услуг безопасности. В статье анализируются основные требования к стойкости алгоритмов постквантовой криптографии. В таких условиях необходимо использование модифицированных криптосистем, обеспечивающих требуемый уровень стойкости и оперативности криптопреобразований. Одним из таких механизмов являются криптокодовые конструкции Р. Мак-Элиса и Х. Нидеррайтера, обеспечивающие требуемые показатели стойкости, оперативности и достоверности. В работе проводится анализ их построения на эллиптических кодах, модифицированных эллиптических кодах и ущербных кодах, дается оценка их стойкости. Предлагается усовершенствованный протокол формирования цифровой подписи с использованием криптокодовых конструкций.

Ключевые слова: *постквантовый период, цифровая подпись, криптокодовые конструкции Мак-Элиса, эллиптические коды*

Введение

Вступление человечества в эру высоких технологий позволило выделить безопасность в киберпространстве (абстрактное понятие на основе компьютерных сетей и Интернет-технологий) в отдельную составляющую информационной безопасности. Для обеспечения информационной безопасности в киберпространстве, как правило, используются симметричные криптосистемы с временной стойкостью, но в 3...5 раз более быстрые по сравнению с несимметричными криптосистемами, обеспечивающими доказуемый уровень стойкости (стойкость основана на NP-полных задачах), что позволяет их использовать при передаче ключевых данных и формировать протоколы цифровой подписи (ЦП), обеспечивающие услугу аутентичности (подлинности источника сообщения). Бурное развитие вычислительных средств обеспечивает рост вычислительных возможностей в два раза каждые 18 месяцев, что позволяет существенно увеличить сферу распространения услуг в киберпространстве. Однако анализ специалистами Национального института стандартов и технологий (НИСТ) США алгоритмов традиционной криптографии [1—3] и алгоритмов несимметричной криптографии, протоколов цифровой подписи (включая алгоритмы с использованием эллиптических кривых) показал, что вычислительные возможности в постквантовый период — использование полномасштабных квантовых компьютеров и алгоритмов взлома Гровера и Шора [4] — позволяют за полиномиальное время взламывать используемые в компьютерных системах и сетях киберпространства данные криптосистемы, что ставит под сомнение качество обеспечения основных услуг безопасности: конфиденциальности, целостности и аутентичности. В работах [4—7] указывается, что с ростом вычислительных возможностей происходит не только расширение ИТ-услуг практически во всех сферах деятельности человечества, но и значительное увеличение гибридных, обеспечивающих синергетический эффект атак с элементами социальной инженерии. Таким образом, возникает научно-техническая задача реализации основных услуг информационной безопасности на основе альтернативных подходов, обеспечивающих, в первую очередь, криптостойкость применяемых алгоритмов.

Исследование требований к алгоритмам постквантовой криптографии

При реализации полномасштабного квантового компьютера алгоритм Шора позволяет разложить на сомножители число N за время $O(\lg^3 N)$, используя $O(\lg N)$-битовый регистр, что существенно быстрее любого классического метода факторизации. Преимуществом использования квантовых регистров является существенная экономия памяти (N квантовых битов могут содержать 2^N битов информации), кроме того, взаимодействие между кубитами дает возможность за одну операцию воздействовать на весь регистр (квантовый параллелизм).

Таким образом, алгоритм Шора поставил под вопрос само существование несимметричной криптографии, поскольку на его основе возможно эффективное решение задач о дискретном логарифмировании и других задач, на сложности которых базируются криптографические алгоритмы. Этот вывод был подтвержден в марте 2018 г. в отчете специалистов НИСТ США (Report on Post-Quantum Cryptography) [1, 2], в котором отмечается, что появление полномасштабных квантовых компьютеров ставит под сомнение криптостойкость алгоритмов несимметричной криптографии. В феврале 2019 г. специалисты НИСТ США при открытии конкурса на алгоритмы постквантовой криптографии заявили, что под сомнение ставятся даже алгоритмы на эллиптических кривых. Таким образом, человечество входит в так называемый постквантовый период — промежуток времени в будущем, когда будут существенно усовершенствованы классические методы и созданы квантовые компьютеры с необходимыми для успешного криптоанализа длинами регистров (в кубитах) и необходимое для их реализации математическое и программное обеспечение. К основным задачам, которые могут быть решены на квантовом компьютере, необходимо отнести следующие:

1) квантовый алгоритм факторизации Шора;

2) квантовый алгоритм Гровера поиска элемента в несортированной базе;

3) квантовый алгоритм Шора для решения дискретного логарифма в конечном поле;

4) квантовый алгоритм решения дискретного логарифма в группе эллиптических кодов (*EC*) Шора;

5) квантовые алгоритмы криптоанализа для преобразований в фактор-кольце;

6) квантовый алгоритм криптоанализа Ксионга и Ванга и его совершенствование и другие.

В табл. 1 приведены результаты сравнительного анализа сложности факторизации для классического и квантового алгоритмов, в табл. 2 — сложность реализации метода Шора дискретного логарифмирования в группе точек *EC*.

Представленные в табл. 1, 2 результаты сравнений указывают на существенное сокращение энергетических затрат на реализацию взлома криптоалгоритмов несимметричной криптографии, к которым относятся и алгоритмы ЦП при использовании квантового компьютера, что в значительной степени снижает уровень "доверия" к алгоритмам и протоколам обеспечения основных услуг безопасности: конфиденциальности, целостности и аутентичности.

В условиях постквантовой криптографии специалисты НИСТ предлагают рассматривать атаки специального вида (SIDE-CHANEL ATTACKS). Реализация этих атак направлена на поиск уязвимостей в криптосистемах.

Предложена следующая классификация специальных атак по следующим признакам:

- контроль над вычислительным процессом;
- способ доступа к системе или средства аутентификации;
- метод непосредственного осуществления атаки и тому подобное.

Таблица 1

Сравнительный анализ сложности факторизации для классического и квантового алгоритмов

Размер модуля N, бит	Число необходимых кубитов $2n$	Сложность квантового алгоритма $4n^3$	Сложность классического алгоритма
512	1024	$0{,}54\cdot10^{9}$	$1{,}6\cdot10^{19}$
3072	6144	$12\cdot10^{10}$	$5\cdot10^{41}$
15360	30720	$1{,}5\cdot10^{13}$	$9{,}2\cdot10^{80}$

Таблица 2

Сложность реализации метода Шора дискретного логарифмирования в группе точек EC

Размер порядка базовой точки, бит	Число необходимых кубитов $f(n) = 7n + 4\log_2 n + 10$	Сложность квантового алгоритма $360n^3$	Сложность классического алгоритма
163	1210	$1{,}6\cdot10^{9}$	$3{,}4\cdot10^{24}$
256	1834	$6\cdot10^{9}$	$3{,}4\cdot10^{38}$
571	4016	$6{,}7\cdot10^{10}$	$8{,}8\cdot10^{85}$
1024	7218	$3{,}8\cdot10^{11}$	$1{,}3\cdot10^{154}$

В основу защиты от атак специального вида могут быть положены следующие особенности:
- фиксированное число обращений к хеш-функции, рандомизация данных;
- независимость ключей от значений и тому подобное.

Основными требованиями НИСТ по безопасности в условиях постквантового периода являются:

△ *требования по безопасности:*
- замена стандарта ЭП FIPS 186;
- замена стандартов распределения ключей SP 800-56A, SP 800-56B;
- использование нового стандарта в протоколах: TLS, SSH, IPSec и тому подобное;
- модель безопасности для шифрования и распределения — схема "семантически безопасного шифрования". Модель безопасности — IND-CCA2;

△ *условия безопасности*:
- доступ злоумышленника менее чем до 2^{64} избранных пар шифртекст—ключ;

△ *требования к устойчивости:*

1) 128 битов классической безопасности / 64 бита квантовой защищенности (запас устойчивости AES-128);

2) 128 битов классической безопасности / 80 битов квантовой защищенности (запас устойчивости SHA-256 / SHA3-256 / SHA-384 / SHA3-384);

3) 256 битов классической безопасности / 128 битов квантовой защищенности (запас устойчивости AES-256).

Таким образом, НИСТ США предлагает рассматривать следующие модели:
- для алгоритмов симметричной криптографии — IND-CCA2 (Indistinguishability Adaptive Ciphertext Attack), что определяет устойчивость к адаптивной атаке на основе выбранного шифр-текста;
- для электронной цифровой подписи — EUF-CMA (existentially unforgeable under adaptive chosen message attacks);
- для протокола инкапсуляции ключей — Canetti-Krawczyk (CK-безопасность).

В качестве предварительного критерия НИСТ предлагает подход, при котором квантовые атаки ограничены фиксированным временем работы, или "глубиной" схемы. Такой параметр назван MAXDEPTH.

Возможные значения для диапазона MAXDEPTH:

— 2^{40} логических вентилей, которые будут последовательно выполняться в год;

— 2^{64} логических вентилей, которые могут выполняться последовательно за десять лет;

— не более чем 2^{96} логических вентилей, которые могут выполняться за тысячелетия.

Таким образом, проведенный анализ показал, что использование ЭЦП на основе несимметричных криптоалгоритмов в постквантовый период не может обеспечить гарантированный уровень криптостойкости, и, соответственно, криптосистема может быть подвержена атаке специального вида на основе полномасштабного квантового компьютера.

Исследование криптокодовых конструкций

Особое место среди симметричных и несимметричных криптосистем занимают несимметричные криптосистемы на основе криптокодовых конструкций (KKK) Мак-Элиса и Нидеррайтера, которые являются участниками конкурса НИСТ на постквантовый алгоритм и интегрированно обеспечивают не только требуемый уровень криптостойкости (при их реализации в поле Галуа для кодирования информации, обозначаемого $GF(2^{10}—2^{13})$), но и достоверность передаваемой информации на основе помехоустойчивых кодов (реализуется способ передачи с прямым исправлением ошибок). Однако существенным недостатком являются трудности их практической реализации в алфавите GF $(2^{10}—2^{13})$, а также значительные энергетические затраты. Кроме этого в работе В. М. Сидельникова [10] предложен практический алгоритм взлома данных конструкций при использовании циклических помехоустойчивых кодов, суть которого сводится к нахождению элементов порождающей матрицы и снятию действия матриц маскировки. Ортогональность матриц — порождающей и проверочной — позволяет рассматривать эффективность атаки и на схему Нидеррайтера. В качестве направления устранения выявленных закономерностей Сидельников предлагает использовать каскадные или алгеброгеометрические коды, построенные на основе алгебры теории помехоустойчивого кодирования и геометрических параметров кривой, в частности эллиптических кривых.

Общая классификация KKK приведена в работе [4] (рис. 1).

Проведенный в работах [4, 8, 9] анализ показал, что данные криптосистемы позволяют обеспечивать доказуемый (математически) уровень безопасности (стойкость основывается на NP-полной задаче декодирования случайного кода), обеспечивают оперативность криптопреобразований на уровне скорости шифрования алго-

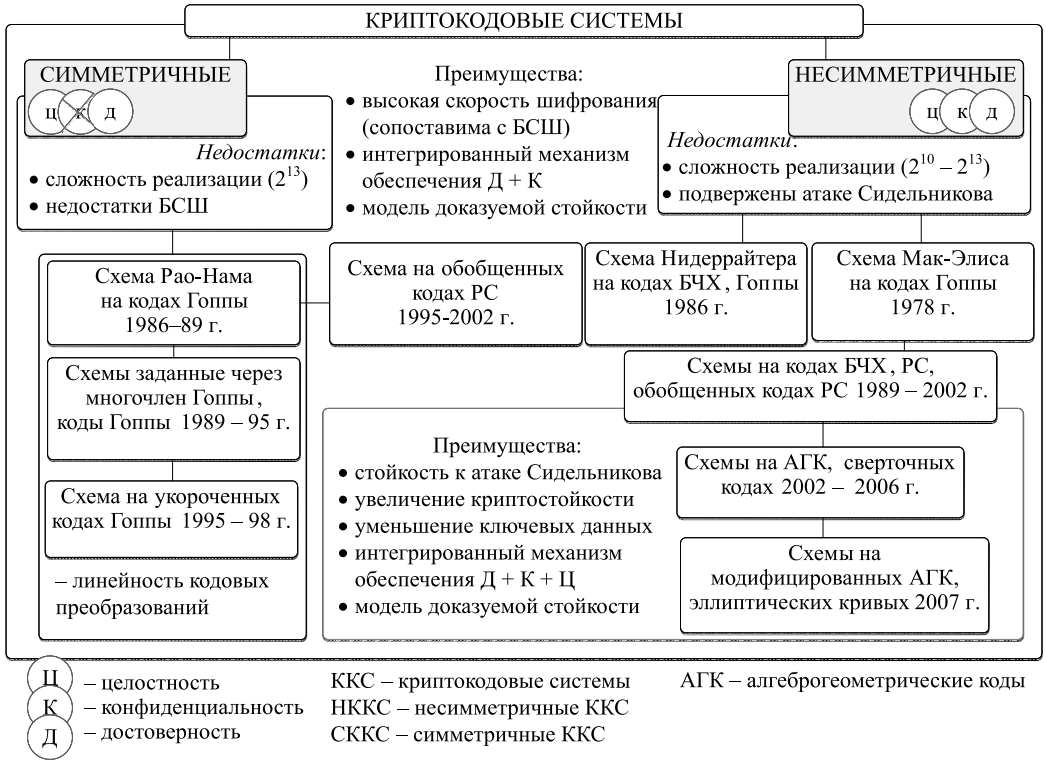


Рис. 1. Общая классификация ККК

ритмами традиционной криптографии и достоверность за счет применения помехоустойчивых кодов. Кроме этого в отчете специалистов НИСТ [1, 2] отмечено, что именно ККК позволяют обеспечить требуемый уровень криптостойкости в постквантовой криптографии.

В качестве закрытого (личного ключа) используется порождающая матрица G линейного (n, k, d) кода над $GF(q)$ и матрицы маскировки: невырожденная $(k \times k)$-матрица X над $GF(q)$, диагональная $(n \times n)$-матрица D, перестановочная $(n \times n)$-матрица P. Перестановочная матрица реализует перестановку координат вектора в виде матричного умножения.

Открытым (публичным) ключом является матрица $G_X = XGPD$.

Шифрование:

$$c_X = iG_X + e,$$

где вектор $c_X = iG_X$ принадлежит (n, k, d)-коду с порождающей матрицей G_X; i — k-разрядный информационный вектор; вектор e — вектор ошибок веса $\leqslant t$, служит дополнительным секретным параметром (сеансовым ключом).

На приемной стороне получатель, зная публичный ключ и используя алгоритм раскодирования (полиномиальной сложности) Берлекемпа—Месси[1], получает исходный текст.

Для устранения недостатка — возможности реализации атаки Сидельникова — предлагается использовать алгеброгеометрические коды (АГК) — коды, построенные на кривых (как пример, на эллиптических кривых).

Для формирования АГК (EC) используются сингулярные (суперсингулярные) кривые 3-го рода.

АГК по кривой X над $GF(q)$ — это линейный код длины $n \leqslant N$, кодовые слова $C(c_1, c_2, ..., c_n)$ которого задаются равенством

$$\sum_{j=0}^{k-1} i_j F_j(P_i) = c_i,$$

где i — k-разрядный информационный вектор; $P_i(X_i, Y_i, Z_i)$ — проективные точки кривой X, т.е. (X_i, Y_i, Z_i) — решения однородного алгебраического уравнения, задающие кривую X, $i = \overline{1, n}$; $F_j(P_i)$ — значения генераторных функций в точках кривой.

[1]Алгоритм Берлекэмпа–Мэсси — алгоритм поиска кратчайшего регистра сдвига с линейной обратной связью для поданной на вход алгоритма требуемой генерируемой последовательности. Алгоритм был открыт Э. Берлекэмпом (англ.) в 1968 г. Применение алгоритма к линейным кодам было найдено Дж. Мэсси в следующем году.

Это определение равносильно матричному представлению АГК [4]:

$$G(i_0, i_1, \ldots, i_{k-1})^T = (c_0, c_1, \ldots, c_{n-1}),$$

где G — порождающая матрица размерности $k \times n$, $k = \alpha - g + 1$, $\alpha = \deg X \deg F$ вида

$$G = \begin{pmatrix} F_0(P_0) & F_0(P_1) & \ldots & F_0(P_{n-1}) \\ F_1(P_0) & F_1(P_1) & \ldots & F_1(P_{n-1}) \\ \ldots & \ldots & \ldots & \ldots \\ F_{k-1}(P_0) & F_{k-1}(P_1) & \ldots & F_{k-1}(P_{n-1}) \end{pmatrix} = \left\| F_j(P_i) \right\|_{n,k}.$$

Однако построение ККК на *EC* не устраняет недостаток — значительную энергоемкость при практической реализации. Для устранения этого недостатка предлагается использовать модифицированные *EC* (*MEC*), предложенные в работах [4, 8].

Рассмотрим *криптосистему на основе ККК Нидеррайтера*, впервые предложенную в работе [12]. Закрытый (личный) ключ — проверочная матрица H линейного (n, k, d)-кода над $GF(q)$, *матрицы маскировки:* невырожденная $(r \times r)$-матрица X над $GF(q)$, диагональная $(n \times n)$-матрица D, перестановочная $(n \times n)$-матрица P. Открытый (публичный) ключ матрица $H_X = XHPD$. Шифрование по правилу

$$S_X = eH_X^T,$$

где вектор e — вектор длины n и веса $\leqslant t$, вычисляется предварительно на основе равновесного кодирования и является преобразованной входной последовательностью. На приемной стороне получатель находит из q^k решений выражения $S_X = c_X H_X^T$. Далее используется расшифрование на основе алгоритма Берлекемпа—Месси.

Для ККК Нидеррайтера используется дополнительный вектор инициализации, определяющий кодовые слова, удовлетворяющие алгоритму раскодирования.

Для дальнейшего снижения уровня энергоемкости с сохранением криптостойкости криптосистемы в работах [4, 9] предлагается использовать гибридные ККК Мак-Элиса и Нидеррайтера на ущербных кодах.

На рис. 2 приведена структурная схема одного шага универсального механизма нанесения ущерба.

Криптографическими ущербными текстами называются тексты, полученные следующими способами [13, 14]:

— *подход 1*: нанесением ущерба исходному тексту с последующим шифрованием ущербного текста и/или его ущербов (рис. 3);

— *подход 2*: нанесением ущерба шифртексту (рис. 4);

— *подход 3*: нанесением ущерба исходному тексту и шифртексту ущербного текста (рис. 5).

Для определения оптимального метода проанализируем отношение числа требуемых дополнительных операций для реализации подхода к размеру результирующих исходящих данных на примере ККК Нидеррайтера.

Зависимость групповых операций реализации несимметричной криптокодовой системы (НККС) от мощности поля приведена в табл. 3.

В табл. 4 представлена длина передаваемых данных.

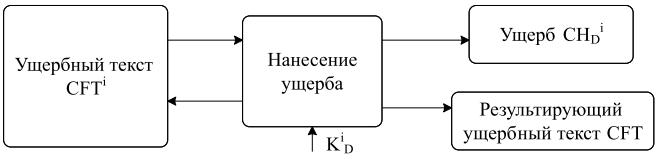


Рис. 2. Структурная схема одного шага механизма нанесения ущерба

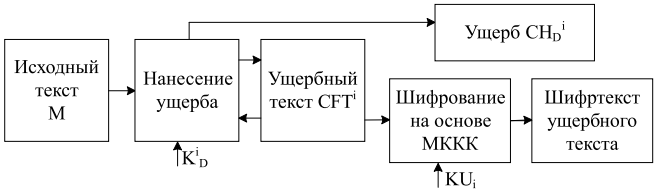


Рис. 3. Структурная схема построения гибридной криптосистемы на основе нанесения ущерба исходному тексту (подход 1)

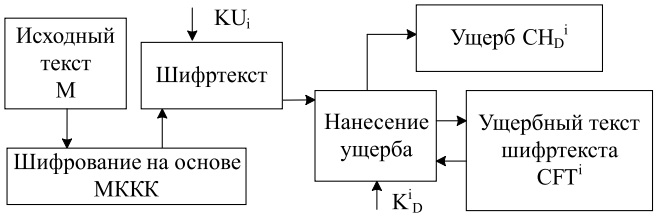


Рис. 4. Структурная схема построения гибридной криптосистемы на основе нанесения ущерба шифртексту (подход 2)

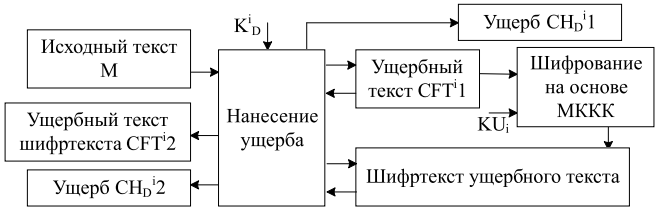


Рис. 5. Структурная схема построения гибридной криптосистемы на основе нанесения ущерба исходному тексту и шифртексту (подход 3)

Таблица 3

Зависимость программной реализации от мощности поля (число тысяч доп. операций перед шифрованием /после/ сумма)

Подход	Мощность поля			
	2^5	2^7	2^9	2^{11}
1	1002/—/1002	3285/—/3285	6322/—/6322	11078/—/8247
2	—/1501/1501	—/4289/4289	—/9296/9296	—/15908/15908
3	992/1487/2479	2952/4428/7380	5793/8690/14483	10086/15130/25216

Таблица 4

Длина передаваемых данных в байтах

Подход	Мощность поля			
	2^5	2^7	2^9	2^{11}
1	500902	902403	1642357	2374489
2	375298	667029	1072313	1652979
3	627533	1044069	1868102	2716713

Таблица 5

Число битов на дополнительную операцию

Подход	Мощность поля			
	2^5	2^7	2^9	2^{11}
1	2,5e-04	4,55e-04	4,812e-04	4,341e-04
2	4,999e-04	8,038e-04	**10,836e-04**	**12,03e-04**
3	4,938e-04	8,836e-04	9,691e-04	11,602e-04

Отношение этих значений показывает число битов пропускной способности на каждую дополнительную операцию (табл. 5).

Таким образом, использование подхода при нанесении ущерба шифртексту с модифицированной ККК на *MEC*, представленного на рис. 4 (второй подход), увеличивает пропускную способность, начиная с поля $GF(2^9)$. Этот способ является оптимальным подходом для построения гибридной ККК Нидеррайтера (Мак-Элиса) на *MEC*.

Под информационным ядром некоторого текста понимается ущербный текст *CFT*, полученный циклическим преобразованием универсального механизма нанесения ущерба C_m.

Исследование стойкости криптокодовых конструкций

Проведем оценку параметров ККК Мак-Элиса с использованием *EC* (*MEC*). Введем следующие обозначения [4]:

l_I — длина информационной последовательности (блока), поступающей на вход теоретико-кодовой схемы (в битах); l_K — длина открытого ключа (в битах); l_{K+} — длина закрытого ключа (в битах); l_s — длина кодограммы (в битах); O_K — сложность формирования кодограммы (число групповых операций); O_{SK} — сложность раскодирования кодограммы (число групповых операций); O_{K+} — сложность решения задачи анализа (число групповых операций); K_C — коэффициент сжатия остатка; K_f — коэффициент сжатия флага; s — число отрезков ущербного текста; $u(n)$, $v(r)$ — положительные числа ключа ущерба; $z(m)$ — раунды ущерба; L_0 — длина исходного текста; L_{DT} — длина ущербного текста.

Для построения графиков были использованы условные сокращения (приставки): ukh/udh — гибридные ККК с укороченными *MEC* (ГККК)/гибридные ККК с удлиненными *MEC*; uk — модифицированная ККК (МККК) с укороченными *MEC*; ud — МККК с удлиненными *MEC*.

При расчетах параметров криптосистем были использованы поля Галуа: для ККК Мак-Элиса — $GF(2^{10})$; для МККК с укороченными/удлиненными *MEC* — $GF(2^6)$; для гибридных ККК — $GF(2^4)$.

Для оценивания *длины информационной последовательности (в битах)*, поступающей на вход МККК с алгебраическим (n, k, d)-кодом над $GF(2^m)$ используем выражения:

- для ККК на *EC* $l_I = km$;
- для МККК на укороченных кодах *MEC* $l_I = (1/2k)m$;
- для МККК на удлиненных кодах *MEC* $l_I = km$.

В табл. 6 и на рис. 6 (см. вторую сторону обложки) представлены зависимости сложности формирования криптограммы от мощности поля (мощности алфавита).

Таким образом, сложность формирования криптограммы в ККК на укороченных и удлиненных кодах значительно ниже (в 5 раз и более), чем в оригинальной реализации ККК Мак-Элиса на *EC*.

Для оценивания *длины кодограммы* (в битах) используем выражения:

- для ККК на *EC* $l_S = nm$;
- для МККК на укороченных *MEC*

$$l_S = (2\sqrt{q} + q + 1 - 1/2k)m;$$

- на удлиненных *MEC*

$$l_S = (2\sqrt{q} + q + 1 - 1/2k + 1/2k)m.$$

В табл. 7 и на рис. 7 (см. вторую сторону обложки) представлены зависимости сложности раскодирования кодограммы от мощности поля.

Анализ рис. 7 и табл. 7 показывает существенный выигрыш при использовании МККК.

Оценка *длины открытого ключа* (в битах) определяется суммой элементов матрицы G_X^{EC} и задается выражениями:

- для ККК на *EC* $l_K = knm$;
- для МККК на укороченных *MEC*

$$l_K = (1/2k)(2\sqrt{q} + q + 1 - 1/2k)m;$$

- для МККК на удлиненных *MEC*

$$l_K = (1/2k)(2\sqrt{q} + q + 1 - 1/2k + 1/2k)m.$$

Оценка *длины закрытого ключа* (в битах) определяется суммой элементов матриц *X*, *P*, *D* (в битах) и задается следующими выражениями:

- для ККК на *EC* $l_{K+} = n^2k^2m$;
- для МККК на укороченных *MEC*

$$l_{K_+} = (1/2k)\left\lceil \log_2(2\sqrt{q} + q + 1)\right\rceil;$$

- для МККК на удлиненных *MEC*

$$l_{K_+} = (1/2k - 1/2k)\left\lceil \log_2(2\sqrt{q} + q + 1)\right\rceil.$$

В табл. 8 и на рис. 8 (см. вторую сторону обложки) представлены зависимости сложности взлома на основе перестановочного декодирования от мощности поля.

Анализ рис. 8 показал, что уменьшение мощности алфавита в МККК на *MEC* обеспечивает требуемый уровень криптостойкости.

Таблица 6

Зависимость сложности формирования криптограммы

$GF(2^m)$	*R*					
	0,5	0,75	0,5(ud)	0,75(ud)	0,5(uk)	0,75(uk)
3	31	87	242	603	817	968
4	76	340	760	980	2140	6282
5	335	872	2241	6121	8706	11461
6	582	2170	**6348**	**9830**	**10722**	**60760**
7	1023	6172	17092	61751	83000	210170
8	5237	10673	67016	105265	207422	605005
9	10563	50487	98765	510780	710920	1018079
10	**52704**	**103822**	497309	908243	4572881	5561379

Таблица 7

Зависимость сложности раскодирования криптограммы

$GF(2^m)$	*R*					
	0,5	0,75	0,5(ud)	0,75(ud)	0,5(uk)	0,75(uk)
1	43	57	78	81	82	96
2	67	98	456	457	457	556
3	120	640	1024	1168	1280	5127
4	680	2378	7672	8232	11028	23674
5	2092	7512	21073	42082	78634	277830
6	12397	61246	**103862**	**281472**	**760553**	**5220573**
7	127523	136495	642648	752018	4566721	19768512
8	1203984	1494284	3564898	3957812	12948312	52694229
9	10637991	12768954	54678128	67458242	92516734	102564872
10	**175645127**	**193648924**	1e + 09	1e + 09	1e + 09	1e + 09

Таблица 8

Зависимость сложности взлома в различных $GF(2^m)$

$GF(2^m)$	*R*					
	0,5	0,75	0,5(ud)	0,75(ud)	0,5(uk)	0,75(uk)
1	1,056	1,38	2,786	2,835	4,122	4,257
2	2,237	3,017	4,978	5,961	6,233	6,781
3	2,868	4,867	7,568	8,120	8,234	9,764
4	4,843	6,613	9,87	12,1	12,647	13,32
5	6,22	8,03	12,017	14,224	14,742	16,892
6	7,891	12,245	**14,983**	**17,483**	**18,767**	**19,76**
7	8,995	13,13	17,14	20,32	21,102	22,93
8	10,37	15,16	19,55	23,23	24,05	26,11
9	11,74	17,18	21,96	26,15	27,002	29,302
10	13,19	19,23	24,37	29,06	29,95	32,484

Оценка с*ложности формирования кодограммы* определяется выражениями:

- для ККК на *ЕС* при реализации систематического кодирования $O_K = (r + 1)n$; для несистематического кодирования $O_K = (k + 1)n$;
- для МККК на укороченных *МЕС* при реализации систематического кодирования

$$O_K = (r+1)(2\sqrt{q} + q + 1 - 1/2k);$$

для несистематического кодирования

$$O_K = (k+1)(2\sqrt{q} + q + 1 - 1/2k);$$

- для МККК на удлиненных *МЕС* при реализации систематического кодирования

$$O_K = (r+1)(2\sqrt{q} + q + 1 - 1/2k + 1/2k);$$

для несистематического кодирования

$$O_K = (k+1)(2\sqrt{q} + q + 1 - 1/2k + 1/2k).$$

Оценка *сложности раскодирования кодограммы* определяется выражениями:

- для ККК на *ЕС*

$$O_{SK} = 2n^2 + k^2 + 4t^2 + (t^2 + t - 2)^2/4,$$

(t — потенциальная стойкость криптограммы);

- для МККК на укороченных *МЕС*

$$O_{SK} = 2(2\sqrt{q} + q + 1 - 1/2k)^2 + \\ + 1/2k^2 + 4t^2 + (t^2 + t - 2)^2/4;$$

- для МККК на удлиненных *МЕС*

$$O_{SK} = 2(2\sqrt{q} + q + 1 - 1/2k + 1/2k) + \\ + k^2 + 4t^2 + (t^2 + t - 2)^2/4.$$

В табл. 9 представлены результаты исследований емкостной характеристики при программной реализации от мощности поля.

Результирующая табл. 9 показывает число групповых операций программной реализации ККК в зависимости от мощности поля. Видно, что если для реализации ККК Мак-Элиса в поле 2^{10} необходимо $82{,}5 \cdot 10^6$ групповых операций, то реализация МККК на укороченных/удлиненных *МЕС* в поле 2^6 требует $(17{,}7...18{,}6) \cdot 10^6$ групповых операций, т. е. в 4,5 раза меньше.

Проведем ***сравнительный анализ параметров МККК Мак-Элиса на МЕС с параметрами ГККК на ущербных кодах.*** Оценка *длины информационной последовательности* (в битах), поступающей на вход криптосистемы с *МЕС* определяется следующими выражениями:

- для ГККК на укороченных кодах

$$l_I = l_z^c + l_z^f,$$

где $l_z^c = K_c L + \frac{1}{K_f} s$ — длина ущербного текста; $l_z^f = L + us$ — длина ущерба; $s = \left\lceil \frac{L_0 - L_{DT}}{L_{DT}} \right\rceil$ — число отрезков ущербного текста; $K_C = 1 - K_f \approx 0{,}758$ — коэффициент сжатия остатка (ущербного текста) (при $u = 8$, $v = 3$, $z = 5$); $K_f = \frac{2 - 2^{v-u+1}}{u} \approx 0{,}242$ — коэффициент сжатия флага (ущерба) (при $u = 8$, $v = 3$, $z = 5$); $z = \frac{\log(uL) - 7}{\log(1/K_C)}$ — необходимое для рандомизации шифра MV2 число допустимых раундов преобразования;

- для ГККК на удлиненных *МЕС*

$$l_I = (1/2k)m + l_z^c + l_z^f.$$

В табл. 10 и на рис. 9 (см. вторую сторону обложки) приведены результаты исследований сложности формирования криптограммы в различных $GF(2^m)$.

Оценка *длины кодограммы* (в битах) определяется выражениями:

- для ГККК на укороченных *МЕС*

$$l_S = (2\sqrt{q} + q + 1 - 1/2k)m;$$

Таблица 9

Зависимость скорости программной реализации от мощности поля (число групповых операций)

Криптосистемы	Мощность поля					
	2^5	2^6	2^7	2^8	2^9	2^{10}
ККК на *ЕС*	10018042	18048068	32847145	47489784	63215578	**82467897**
МККК на укороченных *МЕС*	10007947	**17787431**	28595014	44079433	61974253	79554764
МККК на удлиненных *МЕС*	11156138	**18561228**	33210708	48297112	65171690	84051337

- для гибридных криптокодовых конструкций с ущербными кодами (ГККУК) на удлиненных *MEC*

$$l_S = (2\sqrt{q} + q + 1 - 1/2k + 1/2k)m.$$

В табл. 11 и на рис. 10 (см. вторую сторону обложки) приведены результаты исследований сложности раскодирования криптограммы в различных $GF(2^m)$.

Таким образом, анализ табл. 10, 11 и рис. 9, 10 показал, что использование ущербных кодов и дальнейшее уменьшение мощности поля Галуа приводит к значительному уменьшению сложности формирования (примерно в 12 раз) и раскодирования (примерно в 20 раз) криптограммы.

В табл. 12 и на рис. 11 (см. вторую сторону обложки) приведены результаты исследований

Таблица 10

Зависимость сложности формирования криптограммы

$GF(2^m)$	*R*							
	0,5(ud)	0,75(ud)	0,5(uk)	0,75(uk)	0,5(udh)	0,75(udh)	0,5(ukh)	0,75(ukh)
3	242	603	817	968	643	780	923	998
4	760	980	2140	6282	**905**	**1085**	**1563**	**5125**
5	2241	6121	8706	11461	1863	2450	6137	8282
6	**6348**	**9830**	**10722**	**60760**	6273	7016	9183	10341
7	17092	61751	83000	210170	16582	15985	16563	16925
8	67016	105265	207422	605005	65278	65450	66137	68282
9	98765	510780	710920	1018079	95327	96037	97134	97841

Таблица 11

Зависимость сложности раскодирования криптограммы

$GF(2^m)$	*R*							
	0,5(ud)	0,75(ud)	0,5(uk)	0,75(uk)	0,5(udh)	0,75(udh)	0,5(ukh)	0,75(ukh)
1	78	81	82	96	148	153	1568	1621
2	456	457	457	556	835	897	6112	9624
3	1024	1168	1280	5127	1240	1307	12283	14817
4	7672	8232	11028	23674	**5224**	**11937**	**34673**	**225017**
5	21073	42082	78634	277830	12348	25597	95088	1246572
6	**103862**	**281472**	**760553**	**5220573**	123548	127137	1316373	4383507

Таблица 12

Зависимость сложности взлома ГККК над $GF(2^m)$

$GF(2^m)$	*R*							
	0,5(ud)	0,75(ud)	0,5(uk)	0,75(uk)	0,5(udh)	0,75(udh)	0,5(ukh)	0,75(ukh)
1	2,786	2,835	4,122	4,257	1,089	1,864	2,391	3,46
2	4,978	5,961	6,233	6,781	2,569	3,643	4,108	4,962
3	7,568	8,120	8,234	9,764	3,57	4,131	5,382	7,623
4	9,87	12,1	12,647	13,32	**4,92**	**5,817**	**6,836**	**8,972**
5	12,017	14,224	14,742	16,892	7,591	8,617	10,13	12,005
6	**14,983**	**17,483**	**18,767**	**19,76**	10,85	12,53	14,673	14,962

сложности взлома алгоритмом перестановочного декодирования в различных $GF(2^m)$.

Сложность формирования кодограммы определяется выражениями:

- для ГККК на укороченных *MEC* при реализации систематического кодирования

$$O_K = (r+1)(2\sqrt{q}+q+1-1/2k) + O\left(\frac{1-K_C^u}{K_f}L\right);$$

- для несистематического кодирования

$$O_K = (k+1)(k+1)\times$$
$$\times(2\sqrt{q}+q+1-1/2k) + O\left(\frac{1-K_C^u}{K_f}L\right);$$

- для ГККК на удлиненных *MEC:* при реализации систематического кодирования

$$O_K = (r+1)(2\sqrt{q}+q+1-1/2k+1/2k) +$$
$$+ O\left(\frac{1-K_C^u}{K_f}L\right);$$

- для несистематического кодирования

$$O_K = (k+1)(2\sqrt{q}+q+1-1/2k+1/2k) +$$
$$+ O\left(\frac{1-K_C^u}{K_f}L\right).$$

Оценка *сложности раскодирования кодограммы* определяется выражениями:

- для ГККК на укороченных *MEC*

$$O_{SK} = 2(2\sqrt{q}+q+1-1/2k)^2 + 1/2k^2 +$$
$$+ 4t^2 + (t^2+t-2)^2/4 + O\left(\frac{\alpha - z\log k}{\left|K_z^c L\right|}\right);$$

- для ГКККУК на удлиненных *MEC*

$$O_{SK} = 2(2\sqrt{q}+q+1-1/2k+1/2k)^2 +$$
$$+ k^2 + 4t^2 + (t^2+t-2)^2/4 + O\left(\frac{\alpha - z\log k}{\left|K_z^c L\right|}\right).$$

Оценка *сложности решения задачи анализа (декодирования)* определяется выражениями:

- для ГККК на укороченных *MEC*

$$O_{K+} = N_{\text{покр}} \times$$
$$\times(2\sqrt{q}+q+1-1/2k)\times r + N_F \text{ или } (N_K),$$

где $N_{\text{покр}}$ — коэффициент положительности ключа ущерба; $N_F \approx \frac{K_C{}^z}{2^{1-K_C{}^{z+1}}} \times |F|$, $K_C = 97/128$, $|F|$ — суммарная длина выходных флагов (ущербов) (бит) — при известном злоумышленнику остатке (ущербном тексте) и заданных флагах (ущербах), при неизвестном ключе — $N_K \approx 2^{1190\times z}$, $z = 16$;

- для ГККК на удлиненных *MEC*

$$O_{K+} = N_{\text{покр}}(2\sqrt{q}+q+1-$$
$$-1/2k+1/2k)r + N_F \text{ или } (N_K).$$

В табл. 13 и на рис. 12 (см. третью сторону обложки) приведены результаты исследований сложности взлома и сложности кодирования для различных скоростей R в различных $GF(2^m)$.

В табл. 14 представлены результаты исследований емкостной характеристики при программной реализации в зависимости от мощности поля.

Как и при исследовании МККК результаты, приведенные в табл. 14 подтверждают правильность теоретических выражений, что позволяет на практике подтвердить, что существенное уменьшение открытых ключевых данных для

Таблица 13

Сложность взлома и сложности кодирования

$\lg(l_s)$	R							
	0,5(ud)	0,75(ud)	0,5(uk)	0,75(uk)	0,5(udh)	0,75(udh)	0,5(ukh)	0,75(ukh)
1	15,6	18,23	19,12	19,82	7,21	9,17	12,54	14,56
2	32,47	35,67	38,63	39,18	21,46	23,72	27,48	29,82
3	43,75	51,61	56,88	58,03	31,68	33,83	37,38	38,43
4	59,43	72,81	78,92	80,52	**41,72**	**42,27**	**47,48**	**58,23**
5	68,26	87,32	94,91	104,56	56,63	58,91	62,86	66,53
6	**101,72**	**112,46**	**120,83**	**128,79**	72,32	74,79	89,5	97,71

Таблица 14

Зависимость скорости программной реализации от мощности поля (число групповых операций)

Системы	Мощность поля						
	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}
ККК на укороченных *МЕС*	8293075	10007947	**17787431**	28595014	44079433	61974253	79554764
КККК на удлиненных *МЕС*	8506422	11156138	**18561228**	33210708	48297112	65171690	84051337
ГККК на удлиненных *МЕС*	**5612316**	7900315	14892945	25565274	42279183	58963778	76564173
ГККК на укороченных *МЕС*	**5942627**	7905257	14682411	25595014	42116327	58468143	75474764

Таблица 15

Результаты исследований статистической безопасности

Алгоритм	Число тестов, в которых тестирование прошли более 99 % последовательностей	Число тестов, в которых тестирование прошли более 96 % последовательностей	Число тестов, в которых тестирование прошли менее 96 % последовательностей
ККК McEliece *EC*	149 (78,83 %)	189 (100 %)	0 (0 %)
МККК McEliece на укороченных *МЕС*	151 (79,89 %)	189 (100 %)	0 (0 %)
МНККС McEliece на удлиненных *МЕС*	152 (80,42 %)	189 (100 %)	0 (0 %)
ГККК на удлиненных *МЕС*	153 (80,95 %)	189 (100 %)	0 (0 %)
ГККК на укороченных *МЕС*	155 (82 %)	189 (100 %)	0 (0 %)

гибридных ККК на ущербных кодах приводит к суммарному увеличению относительной скорости передачи.

Для проведения статистических исследований стойкости исследуемых криптосистем воспользуемся пакетом НИСТ STS 822 [15]. Результаты исследований представлены в табл. 15.

Результаты, представленные в табл. 15, подтверждают, что использование гибридных ККК на ущербных кодах и модифицированных ККК на *МЕС* обеспечивают требуемый уровень стойкости. Все криптосистемы прошли 100 % тестов НИСТ, причем наилучший результат показала ГККК на укороченных *МЕС*: 155 из 189 тестов пройдено на уровне 0,99, что составляет 82 % от всего числа тестов. При этом традиционная ККК на *ЕС* Мак-Элиса над $GF(2^{10})$ показала 149 тестов на уровне 0,99.

Разработка модифицированного протокола цифровой подписи на основе криптокодовых конструкций

Для обеспечения услуги аутентичности в киберпространстве используется протокол DSS (Digital Signature Standard — стандарт цифровой подписи), описывающий DSA (Digital Signature Algorithm — алгоритм цифровой подписи) на основе алгоритмов RSA и Эль-Гамаля[2]. Основным отличием несимметричных криптоалгоритмов является относительно более высокий уровень стойкости в алгоритме Эль-Гамаля и возможность использования эллиптических кривых для формирования ЦП. Однако протокол ЦП на RSA обеспечивает более быстрое формирование ЦП. Криптостойкость основывается на стойкости применяемых алгоритмов RSA (NP-полная задача факторизации числа), алгоритма Эл-Гамаля (NP-полная задача нахождения дискретного алгоритма в группе чисел или в группе точек эллиптической кривой в зависимости от использования уравнения *ЕС*). Однако в современных тенденциях развития постквантового периода данные алгоритмы могут не обеспечить требуемый уровень криптостойкости и могут быть взломаны за полиномиальное время. Поэтому предлагается модифицированный протокол DSS на основе ККК, структурная схема приведена на рис. 13.

В качестве ключевых данных отправитель использует личные данные: сеансовый ключ ККК — e (вектор ошибки) и векторы инициализации (IV_1 — символы укорочения, IV_2 — символы удлинения *МЕС*). Получатель в каче-

[1]Тахер Эль-Гамаль — американский криптограф.

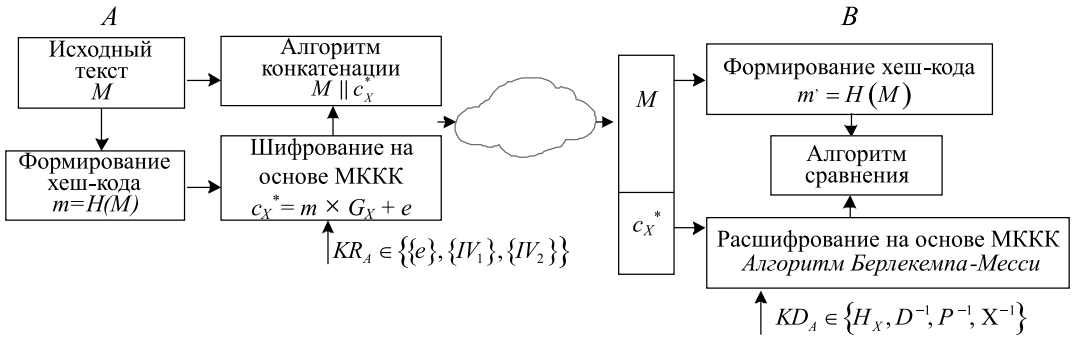


Рис. 13. Модифицированный протокол DSS

стве открытого ключа отправителя использует ортогональную проверочную матрицу *MEC* и обратные матрицы маскировки.

Таким образом, использование ККК в протоколе DSS позволит обеспечить требуемый уровень стойкости в условиях постквантового периода и синергии и/или гибридности современных атак.

Заключение

1. Проведенный анализ вычислительных ресурсов в постквантовый период ставит под сомнение использование для обеспечения услуг безопасности алгоритмов традиционной криптографии и криптографии с открытым ключом. Дальнейшее развитие и появление квантового компьютера позволят киберзлоумышленникам использовать комплексирование угроз для достижения эффекта синергизма и/или гибридности. В таких условиях необходима модификация и/или разработка принципиально новых алгоритмов, обеспечивающих требуемый уровень криптостойкости.

2. Схема модифицированного протокола DSS на основе модифицированных (гибридных) ККК обеспечивает требуемый уровень стойкости к современным угрозам постквантового периода. Проведенные исследования подтверждают, что применение *MEC* (*EC*) обеспечивает быстродействие на уровне скорости криптопреобразований симметричных криптоалгоритмов, доказуемую криптостойкость на основе теоретико-сложностной задачи декодирования случайного кода (обеспечивается 10^{30} ...10^{35} групповых операций) и достоверность на основе использования укороченного алгеброгеометрического кода (обеспечивается $P_{\text{ош}} = 10^{-9}...10^{-12}$). Для дальнейшего уменьшения мощности алфавита (поля Галуа до $GF(2^4—2^6)$) предлагается использовать системы на ущербных кодах, позволяющих одновременно формировать многоканальные криптосистемы.

Список литературы

1. **Guide** for Cybersecurity Event Recovery [Online]. URL: https://nvlpubs.nist.gov/nistpubs/.../NIST.SP.800-184.pdf (Accessed on Februray 1, 2020).

2. **Security** requirements for cryptographic modules [Online]. URL: https://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf (Accessed on Februray 1, 2020).

3. **Guide** to LTE Security [Online]. URL: https://csrc.nist.gov/publications/drafts/800-187/sp800_187_draft.pdf (Accessed on Februray 1, 2020.

4. **Hryshchuk R., Yevseiev, S. Shmatko A.** Construction methodology of information security system of banking information in automated banking systems. Vienna: Premier Publishing s. r. o., 2018. 284 p.

5. **A Comprehensive** Survey of Prominent Cryptographic Aspects for Securing Communication in Post-Quantum IoT Networks. [Online]. URL: https://www.sciencedirect.com/science/article/pii/S2542660520300159 (Accessed on Februray 1, 2020).

6. **Assessing** the quantum-resistant cryptographic agility of routing and switching IT network infrastructure in a large-size financial organization [Online]. URL: https://www.sciencedirect.com/science/article/abs/pii/S2214212618301212 (Accessed on Februray 1, 2020).

7. **Blockchain** for smart communities: Applications, challenges and opportunities [Online]. URL: https://www.sciencedirect.com/science/article/pii/S1084804519302231 (Accessed on Februray 1, 2020).

8. **Сидельников В. М.** Криптография и теория кодирования // Материалы конференции "Московский университет и развитие криптографии в России". МГУ, 2002. С. 1—22.

9. **McEliece R. J.** A Public-Key Criptosystem Based on Algebraic Theory // DGN Progres Report 42-44, Jet Propulsi on Lab. Pasadena, CA. 1978. January — February. P. 114—116.

10. **Niederreiter H.** Knapsack-Type Cryptosystems and Algebraic Coding Theory // Probl. Control and Inform. Theory. 1986. Vol. 15. P. 19—34.

11. **Мищенко В. А., Виланский Ю. В.** Ущербные тексты и многоканальная криптография. Минск: Энциклопедикс, 2007.

12. **Мищенко В. А., Виланский Ю. В., Лепин В. В.** Криптографический алгоритм MV 2. Минск: Энциклопедикс, 2006.

P. A. Shilovskikh, Advisor, e-mail: pashilovskikh@hse.ru,
Higher School of Economics, Perm, 614000, Perm Branch, Russian Federation

Digital Signature Protocol Based on Cryptocode Constructions

The development of computational resources in the post-quantum period challenges the required level of robustness of symmetric and non-symmetric cryptography algorithms. The emergence of a full-scale quantum computer based on Shore and Grover algorithms significantly increases the capabilities of cybercriminals and reduces the robustness of used cryptosystems in the protocols for basic security-free services. The article analyzes the basic requirements for the robustness of post-quantum cryptography algorithms. In such circumstances, it is necessary to use modified cryptosystems that provide an integrated required level of strength and speed of crypto-transformations. One of such mechanisms are cryptocode constructions of McAleese and Niederreiter, providing the required indicators of firmness, efficiency and reliability. In work the analysis of their construction on elliptic (EC), modified elliptic codes (MES), and damaged codes, an estimation of their firmness is spent. An improved protocol for generating a digital signature using crypto-code structures is proposed.

Keywords: *post-quantum period, digital signature, McAleese cryptocode constructions, elliptic codes*

Acknowledgements: This work was supported by the Perm Campus of the Higher School of Economics Russian Federation

References

1. **Guide** for Cybersecurity Event Recovery, [Online], available at: https://nvlpubs.nist.gov/nistpubs/.../NIST.SP.800-184.pdf. Accessed on Februray 1, 2020.

2. **Security** requirements for cryptographic modules, [Online], available at: https://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf. Accessed on Februray 1, 2020.

3. **Guide** to LTE Security, [[Online], available at: https://csrc.nist.gov/publications/drafts/800-187/sp800_187_draft.pdf. Accessed on Februray 1, 2020.

4. **Hryshchuk R., Yevseiev S. Shmatko A.** Construction methodology of information security system of banking information in automated banking systems, Vienna, Premier Publishing s. r. o., 2018, 284 p.

5. **A Comprehensive** Survey of Prominent Cryptographic Aspects for Securing Communication in Post-Quantum IoT Networks, [Online], available at: https://www.sciencedirect.com/science/article/pii/S2542660520300159. Accessed on Februray 1, 2020.

6. **Assessing** the quantum-resistant cryptographic agility of routing and switching IT network infrastructure in a large-size financial organization. [Online], available at: https://www.sciencedirect.com/science/article/abs/pii/S2214212618301212. Accessed on Februray 1, 2020.

7. **Blockchain** for smart communities: Applications, challenges and opportunities. [Online], available at: https://www.sciencedirect.com/science/article/pii/S1084804519302231. Accessed on Februray 1, 2020.

8. **Sidelnikov V. M.** Cryptography and coding theory, *Proceedings of the conference "Moscow University and the development of cryptography in Russia",* Moscow State University, 2002, pp. 1—22.

9. **McEliece R. J.** A Public-Key Criptosystem Based on Algebraic Theory, *DGN Progres Report* pp. 42-44, *Jet Propulsi on Lab. Pasadena*, CA. January — February, 1978, pp. 114—116.

10. **Niederreiter H.** Knapsack-Type Cryptosystems and Algebraic Coding Theory, *Probl. Control and Inform. Theory*, 1986, vol. 15, pp. 19—34.

11. **Mishchenko V. A.**, **Vilansky Yu. V.** Damaged Texts and Multichannel Cryptography, Minsk, Encyclopedics, 2007.

12. **Mishchenko V. A.**, **Vilansky Yu. V.**, **Lepin V. V.** Cryptographic Algorithm MV 2, Minsk, Encyclopedics, 2006.

УДК 004.652.43 DOI: 10.17587/it.27.160-166

М. А. Колокольцев, старший преподаватель, e-mail: makolokoltsev@gmail.com,
У. А. Михалёва, канд. техн. наук, e-mail: uamikhaleva@mail.ru,
ФГАОУ ВО "Северо-восточный федеральный университет им. М. К. Аммосова", г. Якутск

Автоматизированная информационная система "Распределение учебной нагрузки преподавателя"

Рассматривается разработка автоматизированной информационной системы "Распределение учебной нагрузки преподавателя" на основе реляционной модели данных. Создание данной системы способствует экономии времени заведующего кафедрой, ответственного за проверку правильности составления распределения годовой учебной нагрузки между преподавателями. Данный программный продукт является логически завершенным продуктом, готовым к применению.

Ключевые слова: *реляционная модель данных, учебная нагрузка преподавателя, автоматизированная информационная система*

Введение

В настоящее время существует много программных продуктов, автоматизирующих деятельность организаций, в том числе образовательных учреждений. В последних в силу большого числа групп обучающихся, специальностей и предметов очень сложно найти тот программный продукт, который бы автоматизировал составление и расписания, и карточек учебных поручений (КУП) с учетом внутреннего Положения о порядке планирования и учета работы профессорско-преподавательского состава и других нормативных документов, регулирующих учебный процесс.

В данный момент среди образовательных учреждений распространены следующие программные продукты: "1С: Университет ПРОФ" [1], "Программный комплекс "Планы" [2], которые позволяют разрабатывать рабочие учебные планы всех форм и уровней обучения, формировать и распределять учебную нагрузку для структурных подразделений университета и многое другое.

В работе [3] описан опыт успешного внедрения программного обеспечения "1С: Университет ПРОФ" в ФГБОУ ВО "Брянский государственный технический университет". Автор статьи делится опытом внедрения одного из компонентов данного продукта, а также планами по внедрению других компонентов для расширения функциональных возможностей.

В работе [4] рассмотрена модернизация информационной среды "Планы". Авторы разработали свою программную информационную систему, которая преобразовывает XML-формат в Excel-формат для дальнейшей работы с данными. В своей статье авторы подчеркивают многофункциональность информационной среды "Планы", но в то же время считают, что нет необходимости приобретать дополнительную функцию для формирования учебной нагрузки и составления расписания.

В работе [5] описана полностью разработанная автоматизированная система формирования учебных планов и распределения учебной нагрузки преподавателей. Данный продукт учитывает особенности образовательных стандартов и был реализован на базе двухуровневой модели технологии клиент—сервер. Данная система внедрена на кафедре информационных и вычислительных технологий Кыргызско-Российского Славянского университета (г. Бишкек).

Описанную в работе [6] систему распределения учебной нагрузки для кафедры менеджмента и информационных технологий ФГБОУ ВПО "Казанская государственная академия ветеринарной медицины и биотехнологии имени К. И. Скрябина", по словам авторов, не целесообразно применять, когда число

преподавателей меньше 5, а число дисциплино-потоков меньше 10.

Московский государственный университет имени М. В. Ломоносова использует свою автоматизированную информационную систему "Педагогическая нагрузка", которая автоматизирует сбор и анализ данных о занятости в учебном процессе профессорско-преподавательского состава [7].

Существующие в данный момент на рынке программные продукты являются коммерческими проектами отдельно взятых вузов или организаций, и большинство этих систем перегружены деталями, которые являются лишними для отдельно взятой кафедры или структурного подразделения университета. Но в то же время процесс распределения учебной нагрузки между преподавателями кафедры требует большого количества времени и труда. Это указывает на целесообразность разработки и использования автоматизированной информационной системы для формирования общей учебной нагрузки кафедры и структурного подразделения (в часах), а также распределения годовой нагрузки между преподавателями.

В данной статье рассматривается создание программного продукта для автоматизации распределения учебной нагрузки преподавателей на основе реляционной модели данных.

1. Постановка задачи

Программный продукт для автоматизации распределения учебной нагрузки преподавателей разработан для внутреннего пользования в ФГАОУ ВО "Северо-Восточный федеральный университет им. М. К. Аммосова" (далее — СВФУ). Данная автоматизированная информационная система (АИС) разрабатывалась с учетом федеральных государственных образовательных стандартов (ФГОС) и норм времени, указанных в Приложении 1 к Положению о порядке планирования и учета работы профессорского-преподавательского состава в первой половине рабочего дня (СМК-П-2.5-418-18, Версия 2.0) [8].

Исходными данными для разработанного продукта являются:

1) список рабочих учебных планов (РУП). В СВФУ для разработки рабочих учебных планов используется автоматизированная система "Планы" Лаборатории ММИС [2];

2) список существующих кафедр в университете;

3) профессорско-преподавательский состав (ППС);

4) перечень укрупненных групп специальностей (УГС).

Исходные данные обновляются ежегодно.

Необходимо разработать программный продукт, который автоматически формирует отчетные документы в точном соответствии с формами отчетной документации ФГАОУ ВО СВФУ [8].

На рис. 1 представлена модель предлагаемой АИС "Распределение учебной нагрузки преподавателя".

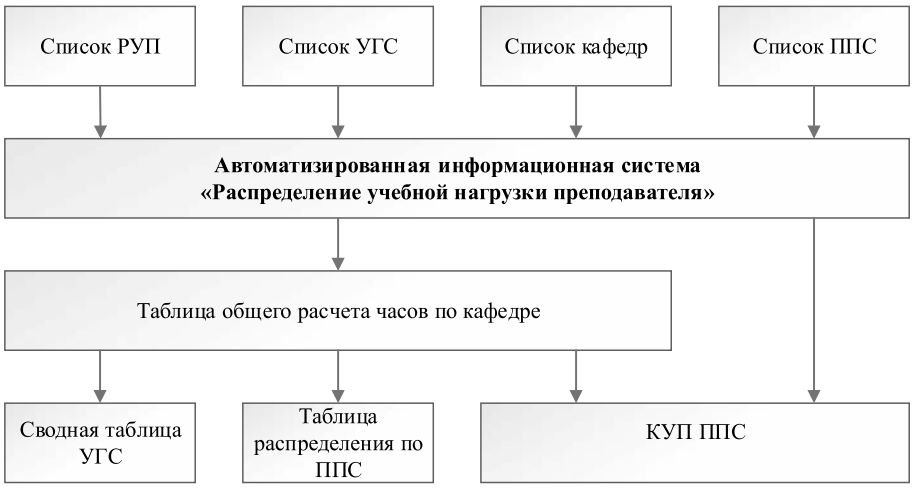


Рис. 1. Модель предлагаемой АИС "Распределение учебной нагрузки преподавателя"

2. Функциональная модель АИС "Распределение учебной нагрузки преподавателя"

Для решения данной задачи целесообразно использовать реляционную модель данных, так как такая модель позволяет представлять информацию с помощью взаимосвязанных таблиц, в которой записи являются уникальными. Данная модель предложена Е. Ф. Коддом, сотрудником компании IBM в 1970 г. [9]. В настоящее время многие системы управления базами данных (СУБД) ориентируются на данную модель. В основе реляционной модели лежит математическая теория отношений.

Перед тем как описать модель, введем некоторые определения, использующиеся в данной статье.

Отношение — это таблица, которая является объектом реляционной модели, состоящая из строк и столбцов. Столбцы называются *атрибутами*, а строки — *записями*. Для идентификации записей используется *первичный ключ*. Представление таблиц и связей между ними отображается *схемой данных*. Существуют *базовое* и *производное* отношения. *Базовое* отношение — это именованное отношение, содержащее один или несколько столбцов, характеризующих свойства объекта, а также первичный ключ. *Производное* отношение — это отношение, которое определено через другие именованные отношения [10].

Считаем, что даны:

1) базовое отношение B_1 — список существующих кафедр СВФУ. Имеет следующие атрибуты: порядковый номер и название кафедры. Номер кафедры является первичным ключом;

2) базовое отношение B_2 — список существующих специальностей СВФУ. Имеет следующие атрибуты: порядковый номер, название специальности и название УГС. Код специальности является первичным ключом;

3) базовое отношение B_3 — профессорско-преподавательский состав СВФУ. Имеет следующие атрибуты: фамилия, имя, отчество, должность, ставка, ученая степень, ученое звание. Порядковый номер является первичным ключом;

4) базовое отношение B_4 — список всех существующих рабочих учебных планов для всех форм обучения. Имеет следующие атрибуты: наименование дисциплины, номер семестра, форма контроля, число лекционных/ практических/ лабораторных часов и т.д. Рабочие учебные планы корректируются каждый новый учебный год;

5) базовое отношение B_5 — список дисциплин по выбору. Имеет следующие атрибуты: порядковый номер и название дисциплины.

Необходимо разработать программный продукт, который должен выполнять следующие функции:

1) автоматизацию детализированного общего расчета часов по кафедре на текущий учебный год (R_1);

2) автоматизацию сводной информации о суммарных часах по УГС по кафедре на текущий учебный год;

3) автоматизацию формирования КУП ППС кафедры на текущий учебный год (R_2);

4) автоматизацию распределения объема часов кафедры по преподавателям кафедры на текущий учебный год.

Решение: производное отношение (R_1), которое получается при установлении отношений B_1, B_2, B_4, B_5, является общим перечнем часов, читаемых сотрудниками кафедры:

$$R_1 \subseteq B_1 \times B_2 \times B_4 \times B_5. \quad (1)$$

Производное отношение (R_2), которое получается при установлении отношений R_1 и B_3, является КУП преподавателя кафедры:

$$R_2 \subseteq R_1 \times B_3. \quad (2)$$

Взаимосвязь базовых отношений представлена в виде схемы на рис. 2 (см. третью сторону обложки).

3. Этапы работы АИС "Распределение учебной нагрузки преподавателя"

В данный момент АИС "Распределение учебной нагрузки преподавателя" является частью комплекса АИС, поддерживающих учебный процесс СВФУ. В дальнейшем планируется создание платформы для расписания и для других отчетных документов, связанных с отчетами кафедры и упрощающих их составление. Тот факт, что данная система является отдельным модулем, позволяет использовать ее в любом в вузе. Необходимые данные подгружаются составителем отчетных данных. Все отчеты формируются в формате MS Excel и могут быть

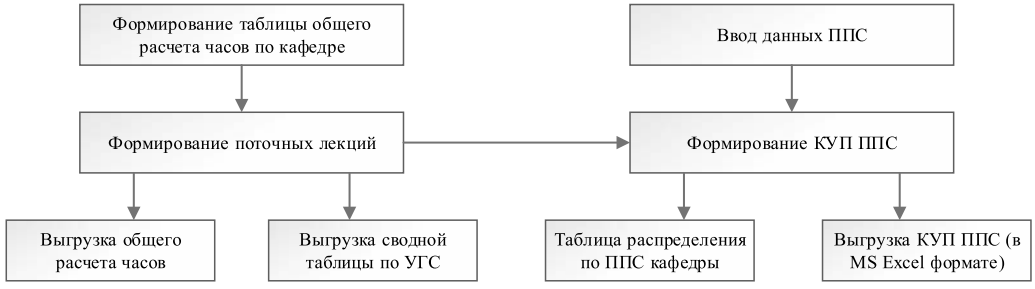


Рис. 3. Этапы работы предлагаемой АИС "Распределение учебной нагрузки преподавателя"

использованы для дальнейшего анализа и обработки.

Работу автоматизированной информационной системы по выполнению поставленных задач можно разбить на этапы, представленные на рис. 3.

На первом этапе проводится загрузка данных из рабочих учебных планов и формируется производное отношение общего расчета часов по кафедре (R_1).

Далее пользователь при необходимости формирует поточность лекций. Полученное в итоге отношение используется для выгрузки общего расчета часов и сводной таблицы по УГС.

На втором этапе выполняется ввод данных о ППС, на основе которых совместно с R_1 формируется производное отношение R_2 КУП для каждого преподавателя кафедры, которое далее используется для экспорта таблицы распределения часов по ППС кафедры и КУП.

4. Результат разработки

В настоящее время данный продукт прошел апробацию на кафедре многоканальных телекоммуникационных систем института математики и информатики СВФУ.

Перед началом работы необходимо выбрать кафедру, для которой будет проводиться расчет часов (рис. 4).

Далее выполняется загрузка информации из рабочих учебных планов (рис. 5). При этом формируются отношения R_1.

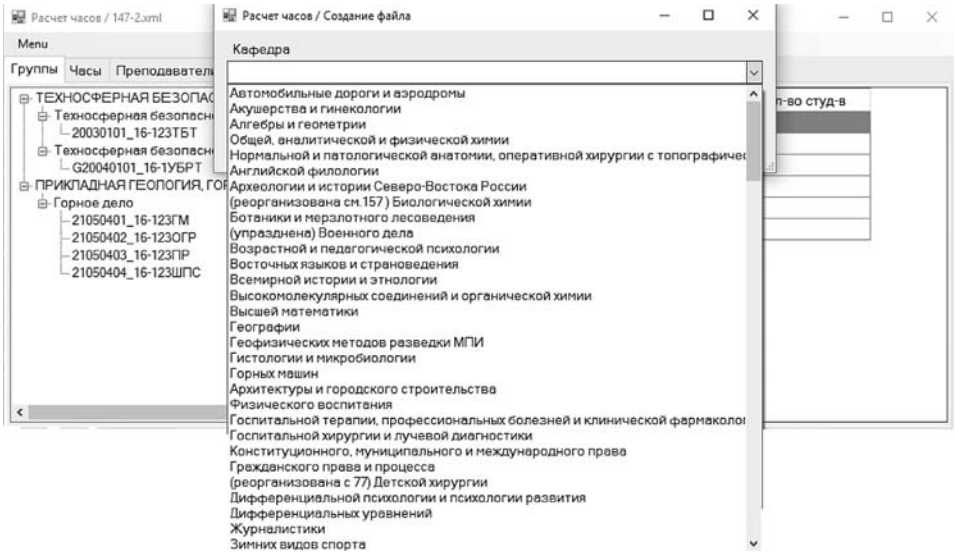


Рис. 4. Выбор кафедры

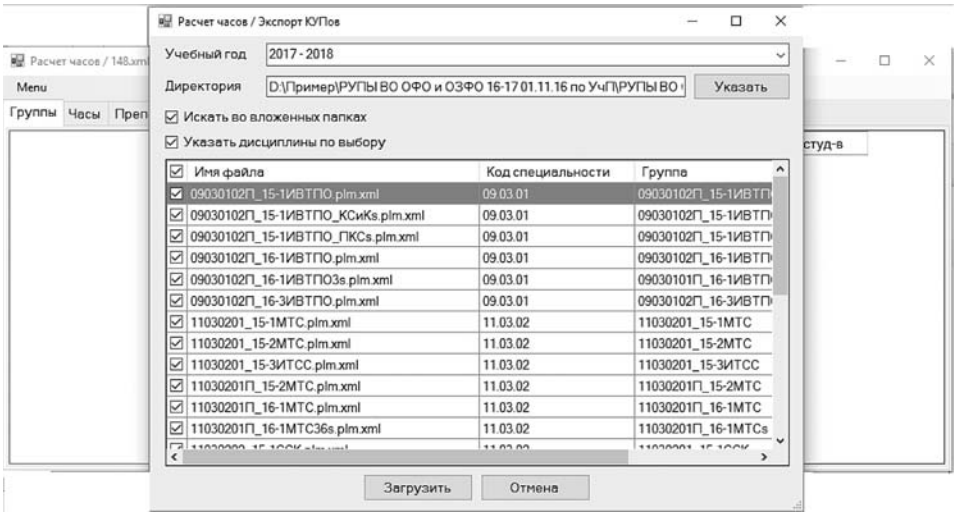


Рис. 5. Загрузка РУП

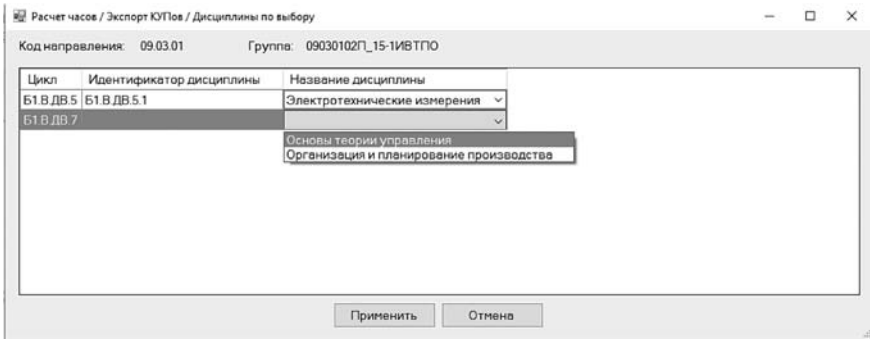


Рис. 6. Определение дисциплины по выбору

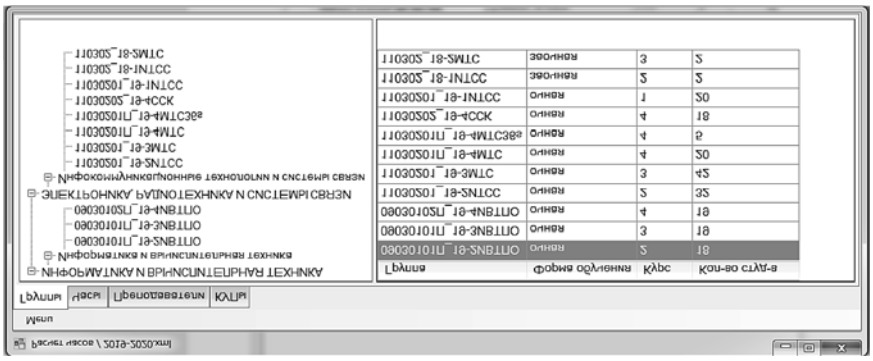


Рис. 7. Корректировка числа студентов

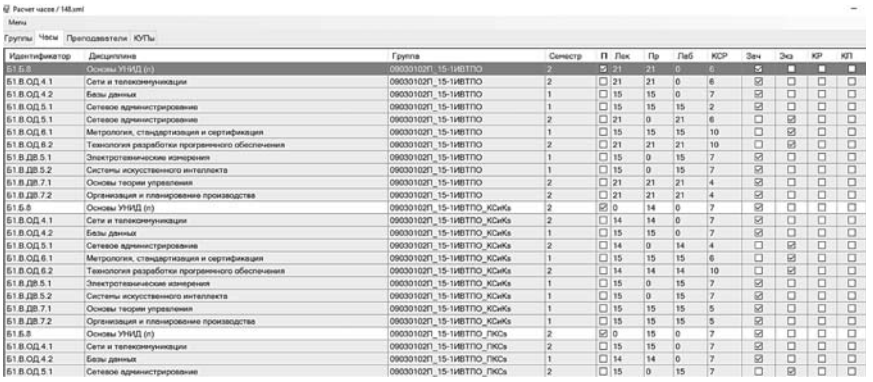


Рис. 8. Внешний вид объединения дисциплин в поток

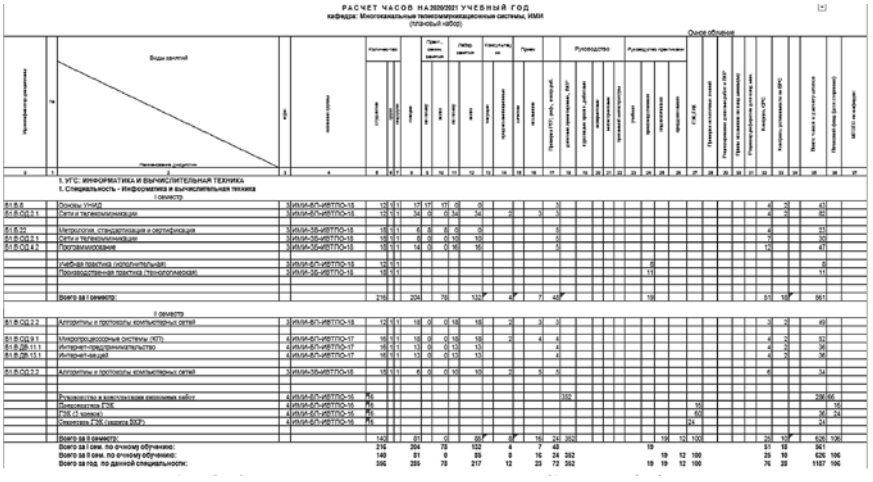


Рис. 9. Общий расчет часов в MS Excel формате

Далее формируется список дисциплин по выбору (рис. 6).

На следующем шаге при необходимости выполняется корректировка числа студентов в группах, если фактическое значение числа учащихся отличается от загруженных из РУП значений (рис. 7).

Далее формируются потоки дисциплин, при этом АИС упрощает эту процедуру, оставляя доступными только строки, соответствующие часам, которые можно объединить в поток (рис. 8).

На данном этапе возможно выгрузка общего расчета часов и сводной информации часов по УГС (рис. 9).

Для формирования КУП на каждого преподавателя необходимо ввести информацию о профессорско-преподавательском составе (рис. 10).

На следующем шаге каждому преподавателю назначаются часы. При этом для выбора АИС выводит список свободных часов, не закрепленных за другими преподавателями (рис. 11).

После того как все часы распределены среди ППС, осуществляется выгрузка КУП и распределение объема часов кафедры по преподавателям кафедры на текущий учебный год (рис. 12).

Заключение

В целом можно сделать заключение, что предложенная модель АИС "Распределение учебной нагрузки преподавателя" на основе реляционной модели данных представляет собой законченный программный продукт, который написан на языке C# и имеет следующие преимущества:

1) повышает эффективность работы заведующего кафедрой, сокращая время, затрачиваемое на отчетные документы;

2) повышает эффективность работы ответственного за пла-

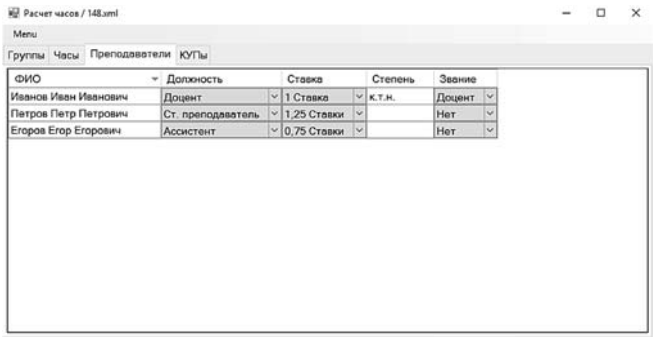


Рис. 10. Ввод данных ППС

нирование и распределение учебной нагрузки структурного подразделения;

3) позволяет экономить время на проверку расчета часов каждой кафедры и структурного подразделения в целом;

4) минимизирует вероятность появления ошибок в отчетных документах, обусловленных человеческим фактором;

5) обладает интуитивно понятным интерфейсом, несложным в освоении.

Таким образом, предложенная в статье реляционная модель данных является эффективным инструментом для создания АИС "Распределение учебной нагрузки преподавателя".

Список литературы

1. **1С: Университет** ПРОФ. URL: https://solutions.1c.ru/catalog/university-prof (дата обращения: 22.08.2020).

2. **Программный** комплекс Планы. Сайт лаборатории математического моделирования и информационных систем (ММиИС). URL: http://www.mmis.ru/ (дата обращения: 22.08.201).

3. **Максимяк И. Н.** От комплексной автоматизации управления образовательным процессом высшего учебного заведения к цифровой трансформации вуза // НОВЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ. Сб. науч. тр. 20-й междунар. науч.-практ. конф. Под общей редакцией Д. В. Чистова. М.: 1С-Паблишинг, 2020. С. 51—52.

4. **Мунтян Е. Р., Поленов М. Ю., Костюк А. И.** О подходе к модернизации программной системы поддержки управленческих решений // Известия ЮФУ. Технические науки. 2015. № 3(164). С. 54—62.

5. **Гаврилец Е. З., Медведева О. А.** Автоматизированная система формирования учебных планов и распределения учебной нагрузки преподавателей кафедры вуза // Современные наукоемкие технологии. 2007. № 2. С. 40—41.

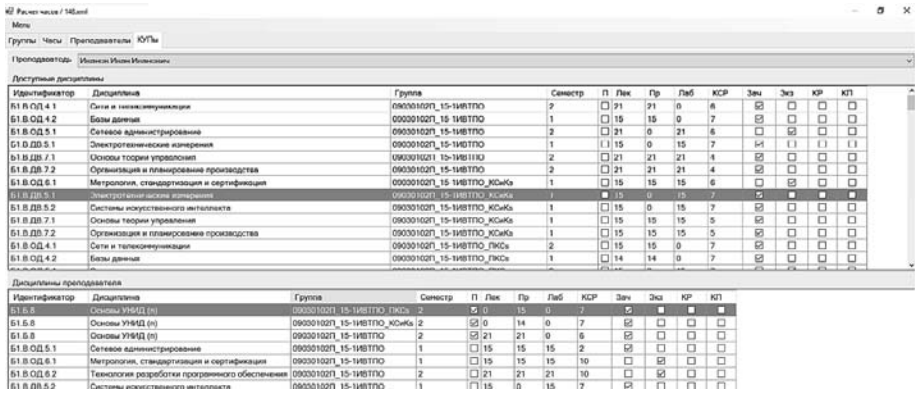


Рис. 11. Формирование КУП ППС

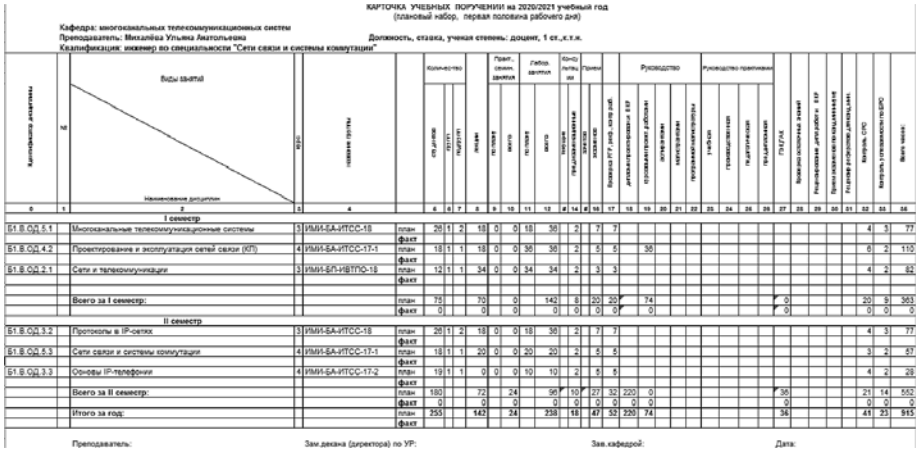


Рис. 12. КУП ППС в MS Excel формате

6. **Леонтьев А. Ю., Василевский Н. М., Акмуллин А. И.** Автоматическая система распределения учебной нагрузки с учетом квалификации преподавателей // Ученые записки Казанской государственной академии ветеринарной медицины им. Н. Э. Баумана. 2013. Т. 216. С. 192—197.

7. **Авраамова О. Д., Болотова И. Н., Владимиров А. В., Вржещ П. В., Галактионова И. А., Ермаков К. В., Зуева С. Ю., Павлов А. П., Рыбин С. И., Рыльская Т. В., Садовникова Е. В., Эрлих Л. И.** Автоматизированная информационная система "Педагогическая нагрузка" / Под.ред. проф. А. В. Тихонравова. М.: Изд-во МГУ, 2011. 46 с.

8. **Положение** о порядке планирования и учета работы профессорского-преподавательского состава в первой половине рабочего дня (СМК-П-2.5-418-18, Версия 2.0).

9. **Ревунков Г. И.** Структуры баз данных: учебное пособие по курсу "Банки данных". М.: Изд.-во МГТУ им. Н. Э. Баумана, 2009. 16 с. URL: http://www.iprbookshop.ru/31569.html (дата обращения: 11.09.2020).

10. **Кузнецов С. Д.** Введение в реляционные базы данных. М.: Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. 247 с. URL: http://www.iprbookshop.ru/73671.html (дата обращения: 11.09.2020).

M. A. Kolokoltsev, Senior Lecturer, e-mail: makolokoltsev@gmail.com,
U. A. Mikhalyova, Assistant Professor, e-mail: uamikhaleva@mail.ru,
Federal State Autonomous Educational Institution of Higher Education "M. K. Ammosov North-Eastern Federal University"

Automated Information System "Distribution of the Teaching Load of the Teacher"

This article discusses the development of an automated information system "Distribution of teaching load of the teacher" based on a relational data model. Existing modern software products for planning the teaching load of a teacher are a commercial product of universities or private organizations. They have many additional functions that are not entirely necessary for a particular university. Therefore, there is a need to create its own automated system for distributing the teaching load of a teacher, which will be easy to use and automatically generates reporting documents in exact accordance with the forms of the university's reporting documentation. The proposed automated information system "Distribution of the teaching load of the teacher" uses a relational data model, since such a model allows information to be represented using interrelated tables in which records are unique. The need to create this system is to save the time of the head of the department, who are responsible for checking the correctness of the distribution of the annual teaching load between teachers. This software product is a logically complete product and ready to use. In the future, it is planned to create a platform for the schedule also for other reporting documents related to the reports of the department and simplifying their preparation. The fact that this system is a separate module allows you to use it in any university. The required data is loaded by the reporting data compiler. All reports are generated in MS Excel format and can be used for further analysis and processing.

Keywords: *a relational data model, a teaching load, an automated information system*

DOI: 10.17587/it.27.160-166

References

1. **1C:** University PROF, available at: https://solutions.1c.ru/catalog/university-prof (date of access: 22.08.2020).

2. **The software** package Planny. Site of the laboratory of mathematical modeling and information systems (MMiIS), available at: http://www.mmis.ru/ (date of access: 22.08.201) (in Russian).

3. **Maksimyak I. N.** From integrated automation of management of the educational process of a higher educational institution to digital transformation of a university, *NEW INFORMATION TECHNOLOGIES IN EDUCATION. Collection of scientific papers of the 20th international scientific and practical conference. Under the general editorship of D. V. Chistova*, Moscow, Limited Liability Company "1C-Publishing", 2020, pp. 51—52 (in Russian).

4. **Muntyan E. R., Polenov M. Yu., Kostyuk A. I.** On the approach to modernization of the software system for supporting managerial decisions, *Izvestia SFU. Technical Science,* 2015, no. 3(164), pp. 54—62 (in Russian).

5. **Gavrilets E. Z., Medvedeva O. A.** Automated system for the formation of curricula and distribution of the teaching load of the teachers of the department of the university, *Modern Scientific Technologies*, 2007, no. 2, pp. 40—41 (in Russian).

6. **Leontiev A. Yu., Vasilevsky N. M., Akmullin A. I.** Automatic system of distribution of the teaching load, taking into account the qualifications of teachers, *Scientific Notes Of The Kazan State Academy Of Veterinary Medicine N. E. Bauman*, 2013, vol. 216, pp. 192—197 (in Russian).

7. **Avraamova O. D., Bolotova I. N., Vladimirov A. V., Vrzhesh P. V., Galaktionova I. A., Ermakov K. V., Zueva S. Yu., Pavlov A. P., Rybin S. I., Rylskaya T. V., Sadovnikova E. V., Ehrlich L. I.** Automated information system "Pedagogical load", Moscow, Publishing house of MSU, 2011, pp. 46 (in Russian).

8. **Regulations** on the procedure for planning and accounting for the work of the teaching staff in the first half of the working day (SMK-P-2.5-418-18, Version 2.0) (in Russian).

9. **Revunkov G I.** Database structures: a tutorial on the course "Data banks", Moscow, Publishing house of MSTU named after N. E. Bauman, 2009, 16 p., available at: http://www.iprbookshop.ru/31569.html (date accessed: 09/11/2020).

10. **Kuznetsov S D.** Introduction to relational databases, Moscow, Internet University of Information Technologies (INTUIT), 2016. 247 p. available at: http://www.iprbookshop.ru/73671.html (date accessed: 09/11/2020).

XIV-я Всероссийская Мультиконференция по проблемам управления (МКПУ-2021)

27 сентября – 2 октября 2021 г.

с. Дивноморское, Геленджик, Краснодарский край, Россия

https://niimvus.org.ru/

ОРГАНИЗАТОРЫ И СПОНСОРЫ

- Российская академия наук
- Министерство науки и высшего образования РФ
- Южный научный центр РАН
- Научный совет по мехатронике и робототехнике РАН
- Академия навигации и управления движением
- ГНЦ РФ ОАО «Концерн «ЦНИИ «Электроприбор»
- Южный федеральный университет
- НИИ многопроцессорных вычислительных систем им. А.В. Каляева ЮФУ
- Институт проблем управления им. В.А. Трапезникова РАН
- Институт проблем механики им. А.Ю. Ишлинского РАН
- Санкт-Петербургский институт информатики и автоматизации РАН
- НИЦ суперЭВМ и нейрокомпьютеров
- Журнал «Известия РАН. Теория и системы управления»
- Журнал «Автоматика и телемеханика»
- Журнал «Мехатроника, автоматизация, управление»
- Журнал «Труды СПИИ РАН»
- Журнал «Известия ЮФУ. Технические науки»

ЦЕЛЬ МУЛЬТИКОНФЕРЕНЦИИ

Обсуждение результатов фундаментальных и прикладных исследований в области процессов управления и их практического применения в различных сферах человеческой деятельности.

СОСТАВ МУЛЬТИКОНФЕРЕНЦИИ

Мультиконференция включает **четыре локальные научно-технические конференции:**

- **Робототехника и мехатроника (РиМ-2021),** председатель – академик Ф.Л. Черноусько
- **Управление в распределенных и сетевых системах (УРСС-2021)**, председатель – академик И.А. Каляев, сопредседатель – академик член-корр. РАН Д.А. Новиков
- **Управление транспортными системами (УТС-2021)**, председатель – член-корр. РАН В.М. Приходько

РОБОТОТЕХНИКА И МЕХАТРОНИКА (РиМ-2021)

Программный комитет

Ф.Л. Черноусько	академик (ИПМех РАН, Москва) – *председатель*
И.Л. Ермолов	проф. РАН (ИПМех РАН, Москва) – *учёный секретарь*
Е.С. Брискин	д.т.н. (ВолгГТУ, Волгоград)
Ю.В. Визильтер	проф. РАН (ГосНИИАС, Москва)
В.А. Глазунов	д.т.н. (ИМАШ РАН, Москва)
З.А. Годжаев	д.т.н. (ВИМ, Москва)
О.В. Даринцев	д.т.н. (ИМех УНЦ РАН, Уфа)
О.Г. Котиев	д.т.н. (МГТУ им. Н.Э. Баумана, Москва)
А.В. Лопота	д.т.н. (ЦНИИ РТК, С.-Петербург)
В.М. Лохин	д.т.н. (МИРЭА, Москва)
А.Л. Ронжин	проф. РАН (СПИИ РАН, С.-Петербург)
И.В. Рубцов	к.т.н. (МГТУ им. Н.Э. Баумана, Москва)
В.В. Серебренный	к.т.н. (МГТУ им. Н.Э. Баумана, Москва)
М.В. Сильников	член-корр. РАН (НПО «СМ», С.-Петербург)
В.Ф. Филаретов	д.т.н. (ИАПУ ДВО РАН, Владивосток)
Н.Б. Филимонов	д.т.н. (МГУ, Москва)
С.П. Хрипунов	д.т.н. (ФПИ, Москва)
С.Г. Цариченко	д.т.н. (НИИ Геодезия, Красноармейск)
А.Ф. Щербатюк	член-корр. РАН (ИПМТ ДВО РАН, Владивосток)
А.И. Якушенко	академик (С.-Петербург)
С.Ф. Яцун	д.т.н. (ЮЗГУ, Курск)

Направления работы

- Кинематика и динамика роботов и мехатронных систем
- Средства очувствления и навигации роботов
- Алгоритмы и системы управления роботов и мехатронных систем
- Планирование поведения роботов в недетерминированных средах
- Групповое управление роботов
- Безэкипажные машины и автономные наземные роботы
- Беспилотные летательные аппараты
- Роботы для ликвидации ЧС
- Роботизация в топливо-энергетическом комплексе
- Роботизация в сельскохозяйственном производстве
- Прикладные аспекты проектирования и применения роботов и мехатронных систем

Секретариат конференции РиМ-2021

119526, Москва, пр. Вернадского, 101, корп. 1, ИПМех РАН

Тел.: +7(495) 434-3547, факс: +7(499) 739-9531

E-mail: **ermolov@ipmnet.ru**

Ермолов Иван Леонидович, д.т.н., профессор РАН

Подробная информация о Мультиконференции размещена на сайте https://niimvus.org.ru/

31 мая – 4 июня 2021 г., ДГТУ, г. Ростов-на-Дону, Россия

Международная научная мультиконференция

Кибер-физические системы: проектирование и моделирование» «Cyber-physical systems design and modelling» (CyberPhy-2021) (Scopus, Springer)

Секции

1. Cyber-Physical Systems: digital technologies and applications (Кибер-физические системы: цифровые технологии и приложения)
2. Cyber-physical systems: design and application for Industry 4.0 (Кибер-физические системы: проектирование и применение для Индустрии 4.0)
3. Cyber-Physical Systems: Modelling and Intelligent Control (Кибер-физические системы: моделирование и интеллектуальное управление)
4. Society 5.0: Cyberspace for advanced human-centered society (Общество 5.0: киберпространство для развитого общества, ориентированного на человека)

XXXIV Международная научная конференция

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ТЕХНИКЕ И ТЕХНОЛОГИЯХ — ММТТ-34 (РИНЦ, DOI)

Секции

1. Качественные и численные методы исследования дифференциальных и интегральных уравнений
2. Оптимизация, автоматизация и оптимальное управление технологическими процессами
3. Математическое моделирование технологических и социальных процессов
4. Математическое моделирование и оптимизация в задачах САПР, аддитивных технологий, цифрового производства
5. Математические методы в задачах радиотехники, радиоэлектроники и телекоммуникаций, геоинформатики, авионики и космонавтики
6. Математические методы и интеллектуальные системы в робототехнике и мехатронике
7. Математические методы в медицине, биотехнологии и экологии
8. Математические методы в экономике и гуманитарных науках
9. Информационные и интеллектуальные технологии в технике и образовании
10. Математические и инструментальные методы технологий Индустрии 4.0
11. Обсуждение квалификационных работ

Подача заявок на участие с 15 декабря 2020 г.

Подробная информация о конференции и условиях участия в ней размещается на сайте http://mmtt.sstu.ru/

Адрес редакции:

107076, Москва, Стромынский пер., 4

Телефон редакции журнала **(499) 269-5510**
E-mail: it@novtex.ru

Технический редактор *Е. В. Конова*.
Корректор *М. Ю. Безменова*.
Сдано в набор 11.01.2021. Подписано в печать 26.02.2021. Формат 60×88 1/8. Бумага офсетная.
Усл. печ. л. 8,86. Заказ IT321. Цена договорная.

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Оригинал-макет ООО "Адвансед солюшнз". Отпечатано в ООО "Адвансед солюшнз".
119071, г. Москва, Ленинский пр-т, д. 19, стр. 1. Сайт: **www.aov.ru**

Рисунок к статье П. А. Шиловских

«ПРОТОКОЛ ЦИФРОВОЙ ПОДПИСИ НА ОСНОВЕ КРИПТОКОДОВЫХ КОНСТРУКЦИЙ»

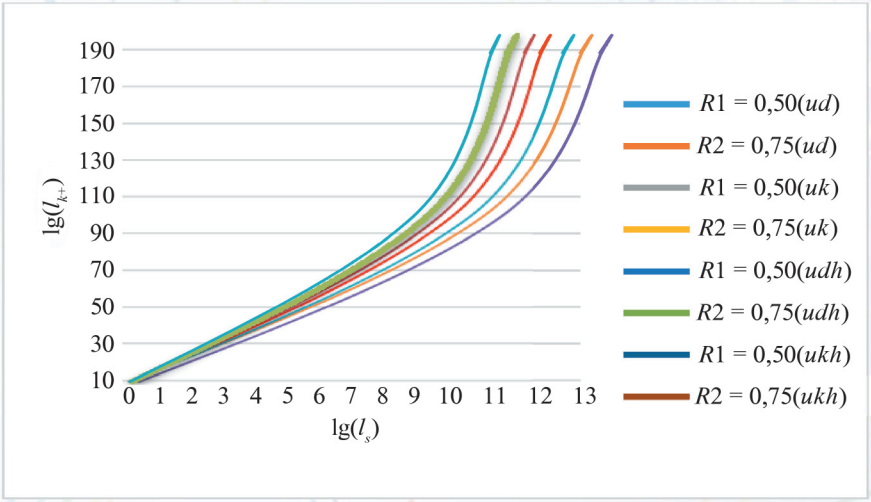


Рис. 12. Сводная диаграмма сложности взлома и сложности кодирования ГККК для различных скоростей МЕС

Рисунок к статье М. А. Колокольцева, У. А. Михалёвой

«АВТОМАТИЗИРОВАННАЯ ИНФОРМАЦИОННАЯ СИСТЕМА «РАСПРЕДЕЛЕНИЕ УЧЕБНОЙ НАГРУЗКИ ПРЕПОДАВАТЕЛЯ»

Рис. 2. Схема данных предлагаемой АИС «Распределение учебной нагрузки преподавателя»

Издательство «НОВЫЕ ТЕХНОЛОГИИ»

выпускает *научно-технические журналы*

Ежемесячный теоретический
и прикладной научно-технический журнал

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

В журнале освещаются современное состояние, тенденции и перспективы развития основных направлений в области разработки, производства и применения информационных технологий.

Подписной индекс по Объединенному каталогу «Пресса России» – 72656

Подписной индекс по Объединенному каталогу «Пресса России» – 79963

Научно-практический
и учебно-методический журнал

БЕЗОПАСНОСТЬ ЖИЗНЕДЕЯТЕЛЬНОСТИ

В журнале освещаются достижения и перспективы в области исследований, обеспечения и совершенствования защиты человека от всех видов опасностей производственной и природной среды, их контроля, мониторинга, предотвращения, ликвидации последствий аварий и катастроф, образования в сфере безопасности жизнедеятельности.

Междисциплинарный
теоретический и прикладной
научно-технический журнал

НАНО- и МИКРОСИСТЕМНАЯ ТЕХНИКА

В журнале освещаются современное состояние, тенденции и перспективы развития нано- и микросистемной техники, рассматриваются вопрос разработки и внедрения нанои микросистем в различные области науки, технологии и производства.

Подписной индекс по Объединенному каталогу «Пресса России» – 79493

Подписной индекс по Объединенному каталогу «Пресса России» – 79492

Ежемесячный теоретический
и прикладной
научно-технический журнал

МЕХАТРОНИКА, АВТОМАТИЗАЦИЯ, УПРАВЛЕНИЕ

В журнале освещаются достижения в области мехатроники, интегрирующей механику, электронику, автоматику и информатику в целях совершенствования технологий производства и создания техники новых поколений. Рассматриваются актуальные проблемы теории и практики автомати ческого и автоматизированного управления техническими объектами и технологическими процессами в промышленности, энергетике и на транспорте.

Теоретический
и прикладной
научно-технический журнал

ПРОГРАММНАЯ ИНЖЕНЕРИЯ

В журнале освещаются состояние и тенденции развития основных направлений индустрии программного обеспечения, связанных с проектированием, конструированием, архитектурой, обеспечением качества и сопровождением жизненного цикла программного обеспечения, а также рассматриваются достижения в области создания и эксплуатации прикладных программно-информационных систем во всех областях челове ческой деятельности.

Подписной индекс по Объединенному каталогу «Пресса России» – 22765

Адрес редакции журналов для авторов и подписчиков:
107076, Москва, Стромынский пер., 4. Издательство "НОВЫЕ ТЕХНОЛОГИИ".
Тел.: (499) 269-55-10, 269-53-97. Факс: (499) 269-55-10. E-mail: antonov@novtex.ru

ISSN 1684-6400. Информационные технологии. 2021. Том 27. № 3. 113–168. Индекс: 72656 (Объединенный каталог «Пресса России»)