Научно-технический журнал

H&ES RESEARCH

Scientific and Technical Journal

НАУКОЕМКИЕ ТЕХНОЛОГИИ В КОСМИЧЕСКИХ ИССЛЕДОВАНИЯХ ЗЕМЛИ

HIGH TECHNOLOGIES IN EARTH SPACE RESEARCH

Журнал **H&ES Research** издается с 2009 года, освещает достижения и проблемы российских инфокоммуникаций, внедрение последних достижений отрасли в автоматизированных системах управления, развитие технологий в информационной безопасности, исследования космоса, развитие спутникового телевидения и навигации, исследование Арктики. Особое место в издании уделено результатам научных исследований молодых ученых в области создания новых средств и технологий космических исследований Земли.

Журнал H&ES Research входит в перечень изданий, публикации в которых учитываются Высшей аттестационной комиссиейРоссии (ВАК РФ), в систему российского индекса научного цитирования (РИНЦ), а также включен в Международный классификатор периодических изданий.

Тематика публикуемых статей в соответствии с перечнем группспециальностей научных работников по Номенклатуре специальностей:

- 05.11.00 Авиационная и ракетно-космическая техника
- 05.12.00 Радиотехника и связь
- 05.13.00 Информатика, вычислительная техника и управление.

H&ES Research is published since 2009. The journal covers achievements and problems of the Russian infocommunication, introduction of the last achievements of branch in automated control systems, development of technologies in information security, space researches, development of satellite television and navigation, research of the Arctic. The special place in the edition is given to results of scientific researches of young scientists in the field of creation of new means and technologies of space researches of Earth.

The journal H&ES Research is included in the list of scientific publications, recommended Higher Attestation Commission Russian Ministry of Education for the publication of scientific works, which reflect the basic scientific content of candidate and doctoral theses. IF of the Russian Science Citation Index.

Subject of published articles according to the list of branches of science and groups of scientific specialties in accordance with the Nomenclature of specialties:

- 05.07.00 Aviation, space-rocket hardware
- 05.12.00 RF technology and communication
- 05.13.00 Informatics, computer engineering and control.

ИНДЕКСИРОВАНИЕ ЖУРНАЛА H&ES RESEARCH

JOURNAL H&ES RESEARCH INDEXING

• NEICON • CyberLenika (Open Science) • Google Scholar • OCLC WorldCat • Ulrich's Periodicals Directory •
• Bielefeld Academic Search Engine (BASE) • eLIBRARY.RU • Registry of Open Access Repositories (ROAR)

Все номера журнала находятся в свободном доступе на сайте журнала **www.hes.ru** и библиотеке **elibrary.ru.**

Всем авторам, желающим разместить научную статью в журнале, необходимо оформить ее согласно требованиям и направить материалы на электронную почту: **HT-ESResearch@yandex.ru.**
С требованиями можно ознакомиться на сайте: **www.H-ES.ru.**

Язык публикаций: русский, английский.
Периодичность выхода – 6 номеров в год.
Свидетельство о регистрации СМИ ПИ № ФС 77-60899 от 02.03.2015
Территория распространения: Российская Федерация, зарубежные страны

Тираж 1000 экз. Цена 1000 руб.
Плата с аспирантов за публикацию рукописи не взимается.

All issues of the journal are in a free access on a site of the journal **www.hes.ru** and **elibrary.ru.**

All authors wishing to post a scientific article in the journal, you must register it according to the requirements and send the materials to your email: **HT-ESResearch@yandex.ru.**
The requirements are available on the website: **www.H-ES.ru.**

Language of publications: Russian, English.
Periodicity – 6 issues per year.
Media Registration Certificate PI No. FS77-60899. Date of issue: March 2, 2015.
Distribution Territory: Russian Federation, foreign countries

Circulation of 1000 copies. Price of 1000 Rur.
Postgraduate students for publication
of the manuscript will not be charged

ISSN 2409-5419 (Print) ISSN 2412-1363 (Online) WWW.H-ES.RU Дата выхода: 25 апреля 2021

2 2021 Т. 13

Учредитель:
ООО «ИД Медиа Паблишер»

Издатель:
ДЫМКОВА С.С.

Главный редактор:
ЛЕГКОВ К.Е.

Адрес издателя:
111024, Россия, Москва, ул. Авиамоторная, д. 8, офис 512-514.

Адрес редакции:
194044, Россия, Санкт-Петербург, Лесной Проспект, 34-36, к. 1,
Тел.: +7(911) 194-12-42.

Адрес типографии:
Россия, Москва, ул. Складочная, д. 3, кор. 6.

СОДЕРЖАНИЕ

АВИАЦИОННАЯ И РАКЕТНО-КОСМИЧЕСКАЯ ТЕХНИКА

РАДИОТЕХНИКА И СВЯЗЬ

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

CONTENTS

Founder:
"Media Publisher", LLC

Publisher:
DYMKOVA S.S.

Editor in chief:
LEGKOV K.E.

Address of publisher:
111024, Russia, Moscow,
st. Aviamotornaya, 8, office 512-514;

Address of edition:
194044, Russia, St. Petersburg,
Lesnoy av., 34-36, h.1,
Phone: +7 (911) 194-12-42.

Address of printing house:
Russia, Moscow, st. Skladochnaya, 3, h. 6

Doi: 10.36724/2409-5419-2021-13-2-4-14

МЕТОДИКИ АГРЕГИРОВАНИЯ ОЦЕНОК ЗНАНИЙ ОБУЧАЮЩИХСЯ ДЛЯ СЛУЧАЕВ ПРЕДСТАВЛЕНИЯ ВНЕШНЕГО КРИТЕРИЯ В ПОРЯДКОВЫХ ШКАЛАХ

ЛИФЕРЕНКО
Виктор Данилович[1]

БАГРЕЦОВ
Сергей Алексеевич[2]

ЧИСТЯКОВ
Денис Владимирович[3]

АННОТАЦИЯ

Введение: вынужденный переход вузов, осуществляющих подготовку специалистов ракетной и авиационной техники на дистанционное обучение стимулировал дальнейшее развитие систем автоматизированного контроля знаний, являющихся одной из важнейших подсистем автоматизированных обучающих систем. Среди множества проблем в системах автоматизированного контроля знаний существует проблема агрегирования оценок, полученных обучающимися за множество занятий (контролей) за период обучения. Такая итоговая (интегральная) оценка должна учитывать неоднородность контролируемых занятий при изучении разделов и тем учебной дисциплины (модуля) и неопределенность полученных ранее оценок. Указанная неопределенность оценок может быть вызвана неудовлетворительным качеством проведения контроля из-за отсутствия, например, необходимого резерва времени, недостаточным качеством организации и применяемых средств контроля знаний обучающихся. **Цель исследования:** цель исследования выражается в анализе и реализации методик агрегирования оценок знаний обучающихся по программам высшего профессионального образования при выполнении ими комплекса задач, имеющих единое смысловое содержание на основе выполнения обучающимися контрольных задач, либо на основе внешних наблюдений результатов деятельности обучающихся преподавателем. **Результаты:** Агрегирование оценок осуществляется в порядковых шкалах. Определен порядок перехода к шкале отношений. Результаты оценок выполнения отдельных этапов учебных заданий представляются в порядковых шкалах или в шкале наименований. Интегральная оценка рассматривается как линейная скалярная функция от исходных параметров, определяющих результаты выполнения контрольных заданий. Параметры скалярной функции определяются на основе оценки расстояний между распределениями оценок знаний обучающихся из состава обучающей выборки, которые задаются опытными педагогами-экспертами. **Методы:** предложенные методики позволяют решить задачу агрегирования оценок, полученных обучающимися за множество занятий (контролей) за период обучения. Исследуются два вида неопределенностей оценок выполнения отдельных этапов учебных задач, а именно на основе их вероятностных или нечетких представлений. **Результаты:** предложенные методики учитывают, прежде всего, целостный характер изучаемого материала, контроль качества выполнения которого на отдельных этапах может быть осуществлен только на основе внешнего наблюдения за деятельностью обучающихся. Подобный подход имеет достаточно общий характер и может быть применим при оценке качества функционирования объектов различной физической природы.

КЛЮЧЕВЫЕ СЛОВА: контроль знаний; шкалы оценок знаний; скалярная функция; нечеткое и вероятностное распределения; агрегирование; метод локальных вариаций; критериальное пространство.

Сведения об авторах:

[1]д.т.н., профессор, ОАО «СУПЕРТЕЛ», г. Санкт-Петербург, Россия

[2]д.т.н., профессор, профессор Военно-космической академии имени А.Ф. Можайского доктор технических наук, г. Санкт-Петербург, Россия, sergeibagrecov@bk.ru

[3]преподаватель Военно-космической академии имени А.Ф. Можайского, г. Санкт-Петербург, Россия

Для цитирования: *Лиференко В.Л., Багрецов С.А., Чистяков Д.В.* Методики агрегирования оценок знаний обучающихся для случаев представления внешнего критерия в порядковых шкалах // Наукоемкие технологии в космических исследованиях Земли. 2021. Т. 13. № 2. С. 4–14. Doi: 10.36724/2409-5419-2021-13-2-4-14

Введение

Необходимость качественной подготовки специалистов по направлению подготовки 24.00.00 Авиационная и ракетно-космическая техника обусловлена требованиями, предъявляемыми к выпускникам технических вузов и формированию у обучающихся профессиональных компетенций, определенных Федеральными государственным и образовательными стандартами (ФГОС). Эффективное освоение учебной программы позволит выпускникам иметь глубокие знания принципов построения и применения авиационной и ракетно-космической техники, необходимые навыки ее эксплуатации, способность выполнять задачи в сложных условиях обстановки, обладать высокими организаторскими способностями, самостоятельностью действий, готовых в короткие сроки освоить выполнение функциональных обязанностей.

Вынужденный переход вузов на дистанционное обучение стимулировал дальнейшее развитие систем автоматизированного контроля знаний, являющихся одной из важнейших подсистем автоматизированных обучающих систем [1,2]. Среди множества проблем контроля знаний рассматривается и проблема агрегирования оценок, полученных обучаемыми за множество занятий (контролей) в цикле обучения. Такая интегральная оценка должна учитывать неоднородность контролируемых занятий при изучении тем и разделов изучаемых дисциплин и неопределенность полученных им ранее оценок. Указанная неопределенность оценок может быть вызвана неудовлетворительным качеством проведения контроля из-за отсутствия, например, необходимого резерва времени, недостаточным качеством организации контроля или применяемых средств контроля знаний. Кроме этого, для целого ряда практических занятий объективный контроль отдельных этапов их выполнения становится вообще невозможным в связи с их целостной формой представления в итогах выполнения учебного задания. Примерами таких занятий могут быть занятия по слаживанию расчетов или дежурных смен по управлению сложных технических систем [7–7]. При проведении таких занятий руководитель на основе имеющейся части объективных данных и во многом, руководствуясь субъективными оценками, определяет степень достижения целей на отдельных этапах выполнения учебного задания. Принимая во внимание во многом неопределенный характер частных оценок таких занятий, можно говорить об их субъективной вероятностной или нечеткой основе. В данном случае, это разделение неопределенностей на субъективно ощущаемую вероятностную или нечеткую основу со стороны руководителя занятия отражает имеющийся опыт оценки этапов выполнения данного типа учебных заданий. Вероятностная аксиоматика оценок, как правило, базируется на имеющемся у руководителя опыте оценок выполнения отдельных этапов подобных задач, достаточном для субъективного определения полной группы событий (факторов), способных повлиять на результат выполнения того или иного этапа учебного задания. В случае нечеткой оценки у руководителя такая основа может быть представлена лишь фрагментарно и не составляет полную группу событий, определяющих исход выполнения рассматриваемого этапа учебного задания, поэтому такая оценка во — многом базируется только лишь на нечеткой интуитивной основе. Оценки качества выполнения отдельных этапов учебного задания далее должны быть агрегированы с учетом их неопределенности и неоднородности их влияния на итоговый результат.

Итоговая интегральная оценка выполнения всего задания может быть представлена в одной из шкал оценок, а именно: номинальной, порядковой, ранговой или шкале отношений. Формирование интегральной шкалы оценок предполагает определение взаимосвязи частных оценок (в данном случае оценок выполнения отдельных этапов задания) в их интегральном представлении. Наиболее часто такая взаимосвязь представляется как аддитивная свертка частных оценок с весовыми коэффициентами, определение которых осуществляется на основе имеющейся внешней по отношению к формируемой шкале системы отношений результатов выполнения учебных заданий в рассматриваемой сфере деятельности. Ниже в статье будем рассматривать только порядковую шкалу оценок, имеющую наиболее широкое распространение на практике.

К настоящему времени накоплено много формальных постановок задач такого типа. В обзорах [8–12] предлагается общая конструкция для единой характеристики различных представлений об интегральном показателе подобного вида. Содержание этой конструкции применительно к оценке знаний обучающихся таково: интегральная оценка знаний в порядковой шкале оценок — это пара ($\boldsymbol{D}$, $\boldsymbol{Z}$), где $\boldsymbol{Z}$ — скалярная функция (обычно линейная) от исходных параметров, определяющих результаты выполнения контрольных заданий, а $\boldsymbol{D}$ — размытое в смысле понятий, введенных Заде, отношение уровней знаний ($\boldsymbol{k}$, $\boldsymbol{r}$) из некоторого множества, образующего обучающую выборку, заданных как точки на оси $\boldsymbol{Z}$. Таким образом, $\boldsymbol{D}$ — это функция от двух переменных ($\boldsymbol{Z}_k$, $\boldsymbol{Z}_r$). Способ определения значений этой функции, задается заранее и существенно зависит от обрабатываемой части данных.

Для формирования интегральной оценки знаний наряду с представлением результатов выполнения контрольных задач как точек в исходном пространстве необходимо задать информацию о взаимосвязях интегральных оценок знаний из состава обучающей выборки. Эта информация задается в виде матрицы $\boldsymbol{Q}=\|\boldsymbol{q}_{rk}\|$ соответствующих коэффициентов связей. В этом случае значения интегрального показателя в порядковой шкале, соответствующие заданным ($\boldsymbol{X}$, $\boldsymbol{Q}$) и выбранным ($\boldsymbol{D}$, $\boldsymbol{Z}$), будут иметь смысл сход-

ства $\boldsymbol{Q}$ с $\boldsymbol{D}$. В задачах оценки знаний $\boldsymbol{Z}$ — линейная функция от $\boldsymbol{X}$, т.е. $Z = \sum_{i=1}^{n} a_i x_i$. При этом в качестве допустимого множества $\widehat{A}$ для векторов $A = (a_i;\ i = \overline{1,n})$ рассматривается ограничение $\sum_{i=1}^{n} a_i = 1, a_i \geq 0$. Задача определения подобного агрегированного показателя с учетом указанных ограничений[1] [13].

Особенностью применения такой методики для диагностики знаний обучающихся с использованием порядковой шкалы является наличие неопределенности в оценках элементов матрицы отношений. Наличие этой неопределенности связано, прежде всего, со сложностью и многообразием связей между характеристиками ($\boldsymbol{x}_i$) исходного пространства результатов выполнения контрольных заданий обучаемыми, характер которых, как правило, не имеет четкого аналитического выражения. Все это приводит к тому, что содержанием отдельных элементов матрицы $\boldsymbol{Q}$ будет являться расстояние между распределениями интегральных оценок знаний обучающихся из состава обучающей выборки, которые задаются опытными педагогами-экспертами. Если каждую точку на оси $\boldsymbol{Z}$ — можно представить в данном случае элементом универсального множества оценок, представляемых при данных характеристиках исходного, пространства $\boldsymbol{X}$, то в отношении типа неопределенности оценок знаний обучающихся можно утверждать, что она носит вероятностный характер. В том случае, если такое утверждение невозможно, то эксперты имеют дело с нестатистической неопределенностью, т.е. с нечеткостью интегральных оценок. Ниже в данном разделе последовательно рассматриваются оба случая формирования порядковых шкал агрегированных оценок знаний обучающихся.

Построение агрегированной оценки знаний при наличии внешнего критерия, имеющего вероятностное распределение

Задача определения вида интегральной оценки знаний в порядковой шкале для типа исходных данных формулируется следующим образом. Пусть $\boldsymbol{X} = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m\}$ — множество результатов выполнения контрольных заданий обучаемыми, включенными в состав обучающей выборки, $X_j = \{x_{ji} : i = \overline{1,n}\}, X = \|x_{ji}\| m,n$ — матрица данных. Характеристики $\boldsymbol{X}_j$ критериальны. Это означает, что, во-первых, $E_{ji} \geq 0, \quad j = \overline{1,m}$, во-вторых, если для пары обучающихся $\boldsymbol{r}$ и $\boldsymbol{k}$ выполняется

$$x_{rp} = x_{kp};\ \ p = \overline{1,n};\ \ x_{rj} > x_{kj}$$

то обучающийся $\boldsymbol{r}$ имеет солее высокую оценку знаний, чем обучающийся $\boldsymbol{k}$.

Обозначим через $Y_r = \{Y_{r\gamma} : \gamma = \overline{1,R_r}\}$ множество допустимых градаций оценок знаний, которые определены, например, экспертами, для каждого объекта $\boldsymbol{r}$ из состава $r \in \{\overline{1,m}\}$ обучающей выборки. При определении оценок эксперты руководствуются исходными данными $\boldsymbol{X}_r$ и своим опытом, определяя возможный диапазон оценок (в порядковой шкале) знаний обучающихся при данном составе исходных данных. При этом множеству допустимых градаций приписываются определенные вероятности

$$P_r = \{P_{r\gamma} : \gamma = \overline{1,R_r}\}(\sum_{\gamma=1}^{R_r} R_{r\gamma} = 1).$$

Аналогичным образом определяется множество оценок знаний r-го обучаемого из ОВ. Тогда элемент матрицы $\boldsymbol{Q} = \|\boldsymbol{q}_{rk}\|$ парных взаимосвязей между обучаемыми из обучающей выборки определяется величиной расстояния между двумя распределениями[2].

$$q_{rk} = Z_{rk} \cdot W_{rk} \cdot H_{rk,} \tag{1}$$

где $$Z_{rk} = \frac{\sum_{\alpha=1}^{R_r}\sum_{\beta=1}^{R_k} \left|y_{r\alpha} - y_{k\beta}\right| \cdot (P_{r\alpha} \cdot P_{r\beta})}{\sum_{\alpha=1}^{R_r}\sum_{\beta=1}^{R_k} (P_{r\alpha} \cdot P_{k\beta})};$$

$$H_{rk} = \frac{H_{rk\max} - H_{rk12}}{H_{rk\max}};$$

$$H_{rk\max} = \ln \theta;$$

θ — число градаций в оценке знаний;

$$H_{rk12} = \frac{H_{rk1} + H_{rk2}}{2};$$

$$H_{rk1} = \sum_{\alpha=1}^{R_r} P_{r\alpha} \cdot \ln P_{r\alpha}; H_{rk2} = \sum_{\beta=1}^{R_k} P_{k\beta} \cdot \ln P_{k\beta};$$

$$W = 1 - \frac{\left|Y_r \cap Y_k\right|}{\left|Y_r \cup Y_k\right|}$$

Как видно из формулы, расстояние $\boldsymbol{q}_{rk}$ тем больше, чем больше неопределенность высказываний экспертов. Причем диапазон изменений $\boldsymbol{q}_{rk}$ соответствует диапазону изменений оценок в порядковой шкале. Отметим, что если для определения $\boldsymbol{Z}_{rk}$ воспользоваться формулой для нормирования оценки расстояния $\boldsymbol{Z}_{rk}$

$$Z_{rk} = \frac{\sum_{\alpha=1}^{R_r}\sum_{\beta=1}^{R_k} \left|Y_{r\alpha} - Y_{k\beta}\right| \cdot (P_{r\alpha} \cdot P_{k\beta})}{2\sum_{\alpha=1}^{R_r} \left|Y_{r\alpha} - Y_{k\beta}\right|}$$

то расстояние $\boldsymbol{q}_{rk}$ будем определять отношением объекта r к объекту $\boldsymbol{k}$ в шкале их соответствия ранее заданным тре-

[1] Авен П.О. Построение интегрального показателя в критериальном пространстве // АиТ. 1985. № 4. С. 87–91.

[2] Загоруйко Н. Г., Бунидев М. В. Меры расстояния в пространстве знаний // Анализ данных в экспертных системах / Под ред. Н. Г. Загоруйко. Новосибирск, 1986. 175 с.

бованиям. В этом случае формируемая далее шкала оценок будет шкалой отношений. Таким образом, на одной информационной основе может быть осуществлен переход из порядковой шкалы в шкалу отношений.

Если показатель $\boldsymbol{Z}$ определяется как линейная комбинация исходных результатов выполнения контрольных заданий, то наблюдаемая структура системы оценок на оси $\boldsymbol{Z}$, определяемая элементами матрицы $\boldsymbol{D}$, будет равна:

$$d_{rk}(A) = (Z_r - Z_k)^2 = \left[\sum_{i=1}^{n}(x_{ri} - x_{ki}) - a_i\right]^2.$$

Соответствие требуемой и наблюдаемой структуры опенок знаний оценивается с помощью функционала $\boldsymbol{J}(\boldsymbol{A})$, определяемого следующим образом:

$$J(A) = \sum_{r,k=1}^{m}(\chi^2 d_{rk}(A) - q_{rk})^2, \qquad (2)$$

где æ — линейный масштабный коэффициент.

Показатель $\boldsymbol{Z}(\boldsymbol{A})$, для которого величина $\boldsymbol{J}(\boldsymbol{A})$ минимальна, а вектор $\boldsymbol{A}$ удовлетворяет условиям:

$$\sum_{j=1}^{n} a_i = 1, \qquad (3)$$

назовем структурным фактором[3] в критериальном пространстве. Задача его отыскания состоит вначале в нахождении вектора $\boldsymbol{A}^*$, удовлетворяющего названным условиям, и минимизирующего функционал при вычисленной матрице $\boldsymbol{Q}$, а затем в определении величины масштабного коэффициента æ*, обеспечивающего глобальный минимум анализируемого функционала.

Задача максимизации функционала (2) при ограничениях (3) относится к классу задач минимизации гладких функций на симплексе[4]. Суть его сводится к следующему. В рассмотрение вводится последовательность $\{\alpha^{(T)}\}$ чисел, отвечающая следующим условиям:

$$0 \le \alpha^{(T)} \le 1, \alpha^{(T)} \to 0 \text{ при } T \to \infty, \sum_{T=1}^{\infty}\alpha^{(T)} = \infty.$$

Последовательности $\{\alpha^{(T)}\}$ соответствует параметрическое семейство операторов $\boldsymbol{M}(\alpha^{(T)})$, где для каждого $\alpha \in \{\alpha^{(T)}\} M(\alpha^{(T)})$ определяется выражением вида:

$$\tilde{a}_t = a_t(1-\alpha) + \alpha;$$
$$a_p = a_p(1-\alpha);\ p \ne t;\ \ 1 \le p \le n.$$

Поиск локального минимума $\boldsymbol{J}(\boldsymbol{A})$ осуществляется следующим образом:

Шаг 1. Положим $\boldsymbol{T} = 1$. Задаём точку $A^{(1)} \in A$. Вычисляем $\boldsymbol{J}(\boldsymbol{A}^{(1)})$.

Шаг 2. Положим $\alpha = \alpha^{(T)}$.

Шаг 3. Применяем в точке $\boldsymbol{A}^{(T)}$ поочерёдно все $\boldsymbol{n}$ операторов системы $\boldsymbol{M}(\alpha^{(T)})$, вычисляя каждый раз величину

$$\Delta_i J(A^{(T)}) = J(M_t(\alpha)\cdot A^{(t)}) - J(A^{(t)}), i = \overline{1,n}.$$

Шаг 4. Определяем $\Delta_i^* J(A^{(t)}) = \min\{\Delta_i J(A^{(t)})\}$.

Шаг 5. Если $\Delta_i^* J(A^{(t)}) \ge 0$, положим $\boldsymbol{A}^{(t+1)} = \boldsymbol{A}^{(t)}$ и переходим к п. 7.

Шаг 6. Если $\Delta_i^* J(A^{(t)}) < 0$, положим $\boldsymbol{A}^{(t+1)} = \boldsymbol{M}_t(\alpha)\cdot\boldsymbol{A}^{(t)}$.

Шаг 7. Увеличиваем $\boldsymbol{t}$ на единицу и переходим к п. 2.

Алгоритм останавливает работу, если при построении очередной точки $\boldsymbol{A}^{(t)} \ne \boldsymbol{A}^{(t+1)}$ достигнута заданная точность $\varepsilon > 0$, т.е. $|\boldsymbol{J}(\boldsymbol{A}^{(t)}) - \boldsymbol{J}(\boldsymbol{A}^{(t-1)})| < \varepsilon$.

В случае изменения требований к знаниям обучающихся необходимо изменить состав обучающей выборки и вновь определить вид интегральной оценки. Для оценки адекватности полученной интегральной характеристики $\boldsymbol{Z}$ требованиям, предъявляемым к уровню знаний, воспользуемся методами статистического анализа. Для этого экспертам предлагается сформулировать новые требования к знаниям обучающихся. В результате экспертами формируются новые интервалы оценок обучающей выборки. Обозначим их через:

$$\{\widehat{Y}_r\} = \{\widehat{Y}_{1r}, \widehat{Y}_{2r}, ..., \widehat{Y}_{R_r r}\}\ \ (r = \overline{1,m})$$

Интегральные оценки обучающихся ($\boldsymbol{Z}$) обозначим через $\widehat{Z}_r = \{\widehat{Z}_1, ..., \widehat{Z}_m\}$. Для оценки степени адекватности модели интегральной оценки обучающихся воспользуемся $\boldsymbol{t}$-критерием, равным[5]:

$$t = \frac{\overline{d}}{S_d / \sqrt{\theta}}, \qquad (4)$$

где $\overline{d} = \sum_{r=1}^{m}\sum_{\alpha=1}^{R_r}(\hat{Y}_{\alpha r} - \hat{Z}_r)\Big/\theta;$

$\theta = \left|U_r \hat{Y}_r\right|;$

$S_d = \sqrt{\sum_{r=1}^{m}\sum_{\alpha=1}^{R_r}(d_r - \overline{d})^2\Big/(\theta - 1)};$

$d_{\alpha r} = \hat{J}_{\alpha r} - \hat{Z}_r.$

Выражение (4) позволяет оценить гипотезу $\boldsymbol{H}_0$, состоящую в том, что разности между оценками экспертов и оценками, полученными по модели (1), есть случайная выборка из нормально распределенной совокупности со

[3]Фу К. С. Структурные методы в распознавании образов: Пер. с англ. Н. В. Завалишина и др. / Под ред. М. А. Айзермана. М.: Мир, 1977. 487 с.

[4]Шевцов Г.С. Линейная алгебра: теория и прикладные аспекты: учеб. пособие. 3-е изд. испр. и доп. М.: Магистр: ИНФРА- М., 2014. 544 с.

[5]Вентцель Е.С., Овчаров Л.А. Теория вероятностей и её инженерные приложения. М.: ACADEMA, 2003.

средним, равным нулю. Если гипотеза $\boldsymbol{H}_0$ верна, то параметр $\boldsymbol{t}$ в выражении (4) будет подчиняться $\boldsymbol{t}$-распределению Стьюдента с $\theta-1$ степенями свободы. Если справедливой окажется противоположная гипотеза $\boldsymbol{H}_1$ о наличии существенных отличий в оценках педагогов-экспертов и оценках полученных по модели, то в выражении (4) будет описываться распределением, идентичным $\boldsymbol{t}$-распределению Стьюдента с $\theta-1$ степенями свободы, однако со средним, отличным от нуля.

Критическое значение для проверки гипотезы $\boldsymbol{H}_0$ против гипотезы $\boldsymbol{H}_0$ на уровне значимости ε с помощью $\boldsymbol{t}$-статистики будет равно[6]

$$\{-(1-\varepsilon/2)^{t_{\theta-1}}, (1-\varepsilon/2)^{t_{\theta-1}}\}. \qquad (5)$$

Если полученное значение $\boldsymbol{t}$-критерия будет меньше наименьшего критического табличного значения (5), то нуль-гипотеза отклоняется, т.е. расчет параметров интегральной оценки знаний должен быть произведен заново на основе новой обучающей выборки. Структурная схема алгоритма формирования порядковой и интегральной шкал оценок знаний обучающихся приведена на рис. 1.

Рассмотрим пример. Допустим, что для определения параметров порядковой шкалы оценок по одной из дисциплин формируется обучающая выборка, в состав которой включены данные о пяти обучающихся. Контрольное задание содержит восемь задач. Результаты выполнения контрольных заданий обучаемыми из состава ОВ представлены в табл. 1.

Таблица 1

Результаты выполнения контрольных заданий обучаемыми из состава ОВ

Номера обучающихся	Результаты выполнения контрольных заданий обучаемыми							
1	0,2	0,4	0,5	0,8	0,4	0,7	0,3	0,4
2	0,5	0,5	0,8	0,8	0,7	0,9	0,6	0,7
3	0,4	0,5	0,5	0,5	0,6	0,6	0,4	0,6
4	0,2	0,2	0,1	0,1	0,3	0,3	0,1	0,2
5	0,4	0,4	0,6	0,7	0,5	0,5	0,4	0,5

Каждый обучающийся из состава ОВ оценивался экспертами соответствующими баллами. В целом оценки экспертов носили неопределенный характер и представлялись в виде интервалов оценок с соответствующим вероятностным распределением. Указанные данные представлены в табл. 2.

Таблица 2

Вероятностное распределение интервалов оценок обучающихся из состава ОВ

Номера обучающихся	Оценка	Распределение вероятностей	
1	5, 4	0,8	0,2
2	3, 4	0,3	0,7
3	3, 2	0,4	0,6
4	4, 5	0,5	0,5
5	2, 3	0,2	0,8

В результате расчетов по описанному выше алгоритму параметры порядковой шкалы оценок указанного контрольного задания будут определяться значениями весовых коэффициентов множества A, равными: $\boldsymbol{a}_1 = 2{,}141$; $\boldsymbol{a}_2 = 2{,}527$; $\boldsymbol{a}_3 = \boldsymbol{a}_4 = \boldsymbol{a}_5 = \boldsymbol{a}_6 = \boldsymbol{a}_7 = 0{,}0564$; $\boldsymbol{a}_8 = 0{,}05$. При этом коэффициент ранговой коррекции ($\boldsymbol{r}_S$) Спирмена для полученной по обучающей выборке модели порядковой шкалы будет равен[6]:

$$r_S = 1 - \frac{6\sum_{r\gamma}(Y_{r\gamma} - Z_r)^2}{\theta(\theta-1)} = 0{,}543.$$

Это несколько ниже критического значения ($\boldsymbol{r}_S = 0{,}595$) коэффициента коррекции для данного числа степеней свободы $\theta = 10$ и уровня значимости 0,05, но достаточно близко к нему. Увеличение точности расчета параметров шкалы требует пересмотра состава ОВ. Определенные выше весовые коэффициенты могут быть внесены в подсистему контроля АОС и могут, далее использоваться для формирования оценок знаний обучающихся по результатам выполнения ими контрольных задач данного цикла обучения.

Таким образом, методика, рассмотренная выше, позволяет на основе анализа данных обучающей выборки, полученных от опытных педагогов, рассчитать характеристики интегральной оценки знаний обучающихся и далее использовать их в практической деятельности.

Построение агрегированной оценки при наличии внешнего критерия, имеющего нечеткое распределение

В предыдущем разделе агрегация оценок обучающихся в порядковой шкале осуществлялась исходя из предложения о наличии статистической неопределенности в определении уровней знаний обучающихся из состава ОВ. Исходя из этого представления исходных данных, в рассмотренной выше методике осуществлялось построение матрицы парных взаимосвязей между элементами ОВ. На практике значительно более часто встречаются случаи, когда неопределенность классификация уровней знаний обучающихся является существенно нестатистической. Задача определения элементов матрицы $\boldsymbol{Q}$ коэффициентов ($\boldsymbol{q}_{rk}$) связей между элементами $\boldsymbol{r}$ и $\boldsymbol{k}$ ОВ в этом случае сводится к опре-

[6]Вентцель Е.С., Овчаров Л.А. Теория вероятностей и её инженерные приложения. М.: ACADEMA, 2003.

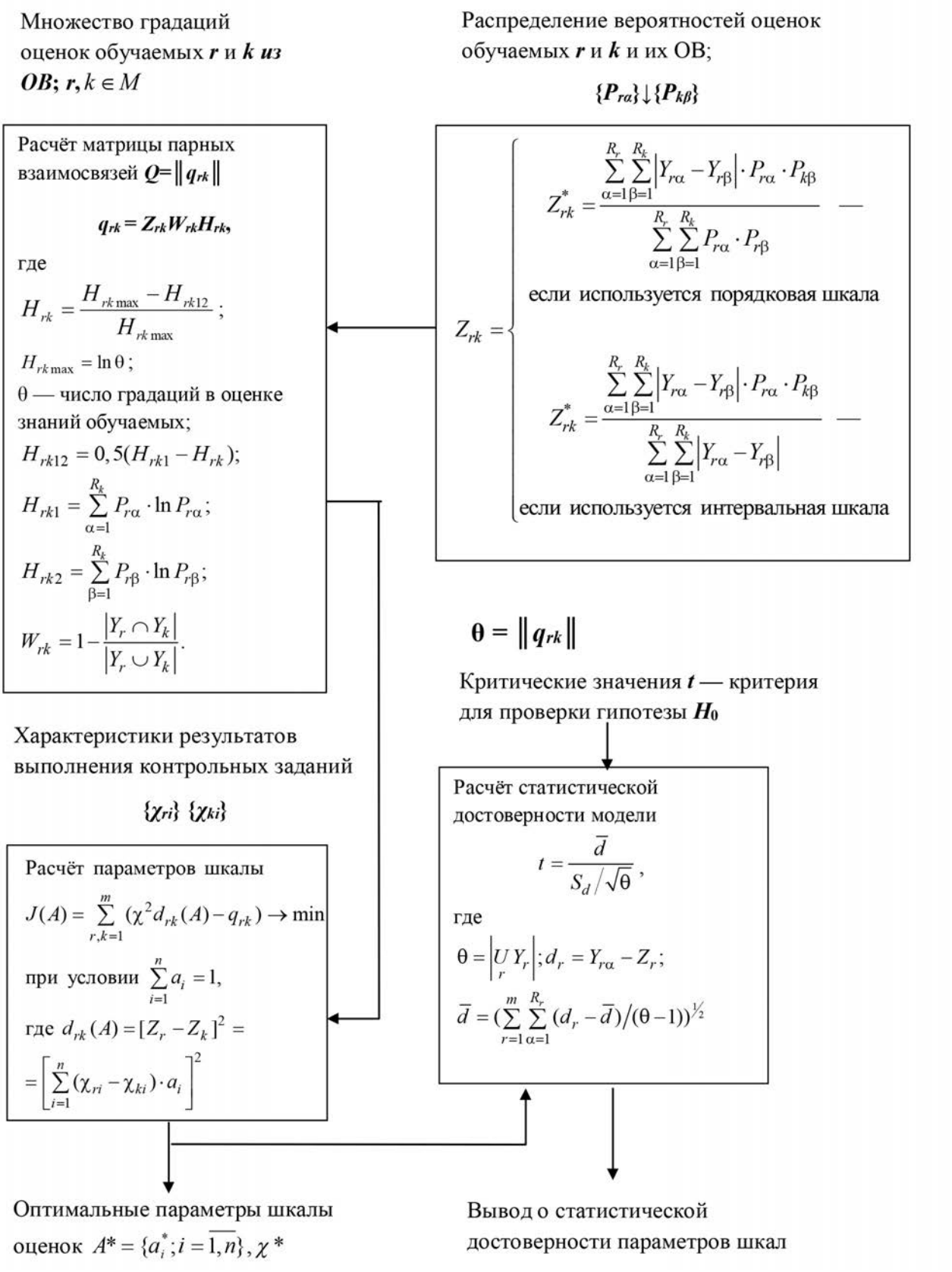

Рис. 1. Структурная схема формирования порядковой и интервальной шкал оценок знаний обучаемых при наличии внешнего критерия, имеющего вероятностное распределение

делению расстояния между двумя соответствующими нечеткими распределениями. Затем по отношению к определяемой таким образом матрице (***Q***) коэффициентов связей применяется изложенная выше процедура аппроксимации, в результате которой определяется структура агрегированного показателя знаний.

Аналогично, как и в предыдущем разделе, будем полагать известными множество результатов ***m*** выполнения контрольных заданий, включенных в состав обучающей выборки, и представляемых в форме матрицы $\boldsymbol{X}=\|\chi_{ij}\|_{m,n}$ данных. Кроме этого, в отношении каждого обучаемого ($r\in\{\overline{1,m}\}$) из 0В будем полагать известными множество $Y_r=\{Y_{r\gamma};\gamma=\overline{1,R_r}\}$ допустимых градаций опенок их знаний в конкретной предметной области, которые определяются экспертами. Отличием от ранее рассмотренной методики является то, что множество допустимых градаций оценок обучающихся из ОВ имеют не вероятностные, а нечеткие распределения, определяемые как функции принадлежности $\mu_2\left(Y_{r\gamma}\right)$ градаций оценок обучающихся[7]. Для построения матрицы парных взаимных связей (матрицы отношений) между обучаемыми воспользуемся соответствующей методикой оценки меры расстояния между двумя нечеткими распределениями. В соответствии с выполненными исследованиями элемент $\boldsymbol{q}'_{rk}$ матрицы ***Q*** определяется следующим образом:

$$\boldsymbol{q}'_{rk}=\boldsymbol{Z}_{rk}\boldsymbol{W}_{rk}\boldsymbol{H}_{rk}, \qquad (6)$$

где $Z_{rk}=\dfrac{1}{\sum\limits_{\alpha=1}^{R_r}\sum\limits_{\beta=1}^{R_k}q\left(Y_{r\alpha},Y_{k\beta}\right)}\cdot\left[\sum\limits_{\alpha=1}^{R_r}\sum\limits_{\beta=1}^{R_k}\left|Y_{r\varepsilon}-Y_{k\beta}\right|g(Y_{r\alpha},Y_{k\beta})\right];$

$\boldsymbol{g}(\boldsymbol{Y}_{r\alpha},\ \boldsymbol{Y}_{k\beta})$ — скалярная оценочная функция нечётких множеств[3];

$\{Y_{r\alpha};\alpha=\overline{1,R_r}\}\ \delta\ \{Y_{k\beta};\beta=\overline{1,R_k}\};$

$g(Y_{r\alpha},Y_{k\beta})=\max\{\mu(Y_{r\alpha})\mu(Y_{k\beta})\};$

$\mu(Y_{r\alpha}),\mu(Y_{k\beta})$ — функции принадлежности градаций α и β оценок обучающихся ***r*** и ***k*** из состава ОВ;

$W_{rk}=1-\dfrac{\left|Y_r\cap Y_k\right|}{\left|Y_r\cup Y_k\right|};$

$\boldsymbol{H}_{rk}$ — степень несоответствия нечёткости градаций α и β оценок обучающихся;

$H_{rk}=1-\sup\{(\min(\mu(Y_{r\alpha}),\mu(Y_{k\beta}))\cdot\delta(Y_{r\alpha},Y_{k\beta}))\};$

$Y_{r\alpha}\in Y_r;Y_{k\beta}\in Y_k;$

$\delta(Y_{r\alpha},Y_{k\beta})=\begin{cases}1,\text{при } Y_{r\alpha}=Y_{k\beta};\\0,\text{в противном случае.}\end{cases}$

[7]Орловский С.А. Проблемы принятия решений при нечеткой информации. М.: Наука. Главная редакция физико-математической литературы, 1981. 208 с.

Как видно из формулы (6), элемент связи $\boldsymbol{q}_{rk}$ будет тем больше, чем больше нечеткость оценок знаний обучающихся ***r*** и ***k*** и чем больше их взаимное соответствие. Полученная на основе такого подхода матрица расстояний далее, аналогично, как и в предыдущей методике, используется для определения параметров интегральной оценки знаний в порядковой шкале. Как и в предыдущей методике формирования порядковой шкалы оценок, в данном алгоритме возможен переход из шкалы порядка в шкалу отношений. Для этого необходимо изменить содержательный смысл параметра $\boldsymbol{Z}_{rk}$ путем расчета его по формуле:

$$Z_{rk}=\frac{1}{\sum\limits_{\alpha=1}^{R_r}\sum\limits_{\beta=1}^{R_k}\left(Y_{r\alpha}-Y_{k\beta}\right)}\cdot\left[\sum_{\alpha=1}^{R_r}\sum_{\beta=1}^{R_k}\left|Y_{r\alpha}-Y_{k\beta}\right|g(Y_{r\alpha},Y_{k\beta})\right].$$

В этом случае элемент $\boldsymbol{q}_{rk}$ приобретает смысл отношения двух объектов ОВ в плане их соответствия некоторым единым требованиям. Характерно, что при таком подходе появляется возможность представлять исходные данные в обучающей выборке не только в количественной порядковой шкале, но и в качественной шкале и в шкале наименований. Например, знания обучающихся из состава ОВ могут быть оценены экспертом в качественной шкале такими выражениями как: вполне соответствуют (занимаемой должности); соответствуют; недостаточно соответствуют; не соответствуют и т.д. Если число различаемых упорядоченных состояний уровней знаний обучающихся ***n***, то все они могут быть пронумерованы числами от 1 до ***n***, после чего расстояние между высказываниями в шкале порядка можно определить по той же формуле, что и для более сильной шкалы отношений [14].

Рассмотрим теперь шкалу наименований. Диагностическая информация, выражаемая в шкале наименований, может содержать, например, данные о характере общественной работы, о достижениях в профессиональной деятельности, об особенностях характера обучаемого и т.п. Применение данной шкалы оправдано особенно в тех случаях, когда множество наименований достаточно полно отражает цель диагностических измерений. В определенных условиях деятельности обучающихся шкала наименований может быть определена ключевыми словами, выражающими характер внешнего поведения человека, например: принципиален, вежлив, честен, предан, военную тайну хранить умеет и т.п. [15].

Количество «имен» личностных характеристик обучающихся и градаций их знаний в ОВ ограничено. Для определения элементов матрицы (***Q***) отношений в шкале наименовани можно воспользоваться вышеприведенной формулой с той лишь поправкой, что расстояние $|\boldsymbol{Y}_{r\alpha}-\boldsymbol{Y}_{k\beta}|=0$, если имя $\boldsymbol{Y}_{r\alpha}$ совпадает с именем $\boldsymbol{Y}_{k\beta}$. Структурная схема формирования порядковой и интервальной шкал оценок знаний обучающихся при наличии внешнего критерия, имеющего нечеткое распределение, представлена на рис. 2.

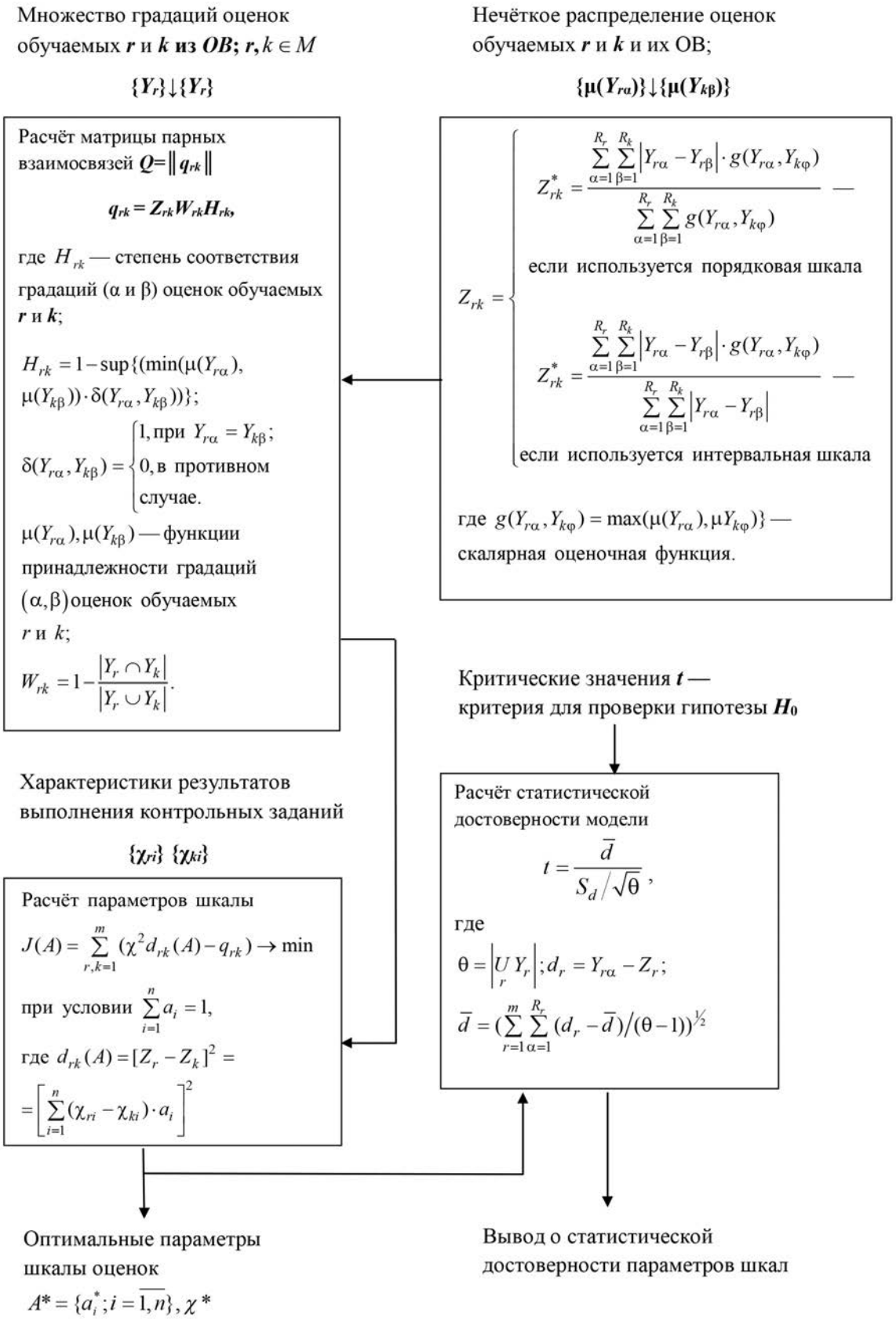

Рис. 2. Структурная схема формирования порядковой и интервальной шкал оценок знаний обучаемых при наличии внешнего критерия, имеющего нечёткое распределение

Аналогично, как и в предыдущем алгоритме, контроль достоверности формирования шкалы оценок знаний обучающихся может быть осуществлен либо по результатам расчета ***t***-критерия Стьюдента, либо по величине ранговой корреляции Спирмена[8].

В качестве примера рассмотрим вариант формирования порядковой балльной шкалы оценок для контроля итогового практического занятия одной из дисциплин. Исходные данные для формирования обучающей выборки следующие: число обучающихся ОВ — 5; количество контрольных задач — 14 (число результатов контроля); количество градаций знаний обучающихся из 0В — 2 (нижняя и верхняяоценка знаний); число классов (баллов) в порядковой шкале — 5. Результаты выполнения контрольных задач обучаемыми из состава ОВ представлены в табл. 3.

Таблица 3

Результаты выполнения контрольных задач обучаемыми из состава ОВ

Обучаемые	Результаты контроля по задачам													
1	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,8	0,7	0,6	0,5	0,4
2	0,1	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,2	0,3	0,4	0,5
3	0,1	0,1	0,2	0,2	0,3	0,3	0,4	0,4	0,5	0,5	0,6	0,6	0,7	0,7
4	0,9	0,9	0,8	0,8	0,7	0,7	0,6	0,6	0,5	0,5	0,4	0,4	0,3	0,3
5	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,9	0,8	0,7	0,6	0,5	0,4

Результаты нечетких оценок знаний обучающихся из состава ОВ приведены в табл. 4.

Таблица 4

Результаты оценок знаний обучающихся из состава ОВ

Обучаемые	Оценки знаний		Степень чёткости оценок	
1	5,0	4,0	0,8	0,2
2	3,0	4,0	0,3	0,7
3	3,0	2,0	0,4	0,6
4	4,0	5,0	0,5	0,5
5	2,0	3,0	0,2	0,8

В процессе расчёта была определена матрица взаимных предпочтений обучающихся из состава ОВ, содержание которой представлено ниже

$$Q = \begin{Vmatrix} 0.00 & 1.04 & 2.15 & 0.12 & 2.00 \\ 1.04 & 0.00 & 1.12 & 0.92 & 0.96 \\ 2.15 & 1.13 & 0.00 & 2.05 & 0.15 \\ 0.12 & 0.92 & 2.05 & 0.00 & 1.88 \\ 2.00 & 0.96 & 0.15 & 1.88 & 0.00 \end{Vmatrix}$$

Весовые коэффициенты порядковой шкалы, рассчитанные по параметрам обучающей выборки, имеют вид: $\boldsymbol{a}_1 = 0{,}0059$; $\boldsymbol{a}_2 = 0{,}0059$; $\boldsymbol{a}_3 = 0{,}155$; $\boldsymbol{a}_4 = 0{,}395$; $\boldsymbol{a}_5 = 1{,}028$; $\boldsymbol{a}_6 = \boldsymbol{a}_7 = \boldsymbol{a}_8 = \boldsymbol{a}_9 = \boldsymbol{a}_{10} = 0{,}0059$; $\boldsymbol{a}_{11} = 0{,}40$; $\boldsymbol{a}_{12} = 0{,}675$; $\boldsymbol{a}_{13} = 0{,}144$; $\boldsymbol{a}_{14} = 1{,}26$. Коэффициент ранговой корреляции Спирмена, рассчитанный по обучающей выборке $\boldsymbol{r}_S(\boldsymbol{TAU}) = 0{,}63$. Он существенно больше критического значения $\boldsymbol{r}_{Skp} = 0{,}595$, что доказывает справедливость полученной шкалы оценок.

Заключение

Многообразие форм и методов представления знаний обучаемыми при решении ими практических задач (например, задач специальной подготовки), предполагающих необходимость учета ситуационного характера оценки учебной обстановки и принятия решений, определяют необходимость разработки систем контроля знаний, базирующихся на системном, комплексном анализе качества выполнения обучаемыми отдельных этапов контрольных заданий. Такая оценка, базируясь в целом, на субъективных оценках качества выполнения учебных заданий преподавателями должна учитывать их опыт. В первую очередь это касается методов формирования оценок деятельности обучающихся на каждом этапе выполнения учебного задания. В зависимости от опыта преподавателя он ориентируется на вероятностные (при условии имеющегося достаточного опыта проведения подобных занятий) или нечеткие (при условии наличия только внутренних (нечетких)) суждения о качестве деятельности обучающихся. Итоговый результат оценки деятельности обучающихся базируется на интегрированной оценке педагогов о взаимосвязи результатов выполнения отдельных этапов учебного задания. Рассмотренные выше методики учитывают, прежде всего, целостный характер изучаемого материала, контроль качества выполнения которого на отдельных этапах может быть осуществлен только на основе внешнего наблюдения за деятельностью обучающихся.

Подобный подход имеет достаточно общий характер и может быть применим при оценке качества функционирования объектов различной физической природы.

Литература

1. *Яковлева Е.И., Шобонов Н.А.* Индивидуализация образовательного процесса в рамках модульной модели программы

[8]Вентцель Е.С., Овчаров Л.А. Теория вероятностей и её инженерные приложения. М.: ACADEMA, 2003.

повышения квалификации при применении дистанционных образовательных технологий // Современные проблемы науки и образования. 2016. № 6. С. 1–8.

2. *Грунт Е.В., Беляева Е.А., Лисситса с.* Дистанционное образование в условиях пандемии: новые вызовы российскому высшему образованию // Перспективы науки и образования. 2020. № 5(47). С. 45-58. DOI: 10.32744/pse.2020.5.3

3. *Ким Н.Ф.* Рейтинговая система оценки успеваемости студентов вуза как фактор повышения качества образования // Молодой ученый. 2015. № 17(97). С. 535–537.

4. *Гельман В.Я.* Совершенствование форм контроля успеваемости в вузе // Современное образование. 2019. № 2. С. 52–57. DOI: 10.25136/2409-8736.2019.2.28364

5. *Кривоногов С.В.* Разработка информационной системы для контроля и оценки знаний студентов // Вестник НГИЭИ. 2016. № 8 (63). С. 30-41.

6. *Ларин С.Н., Юдинова В.В., Юрятина Н.Н.* Информационные технологии контроля уровня знаний обучаемых: российский опыт // Педагогические науки. 2017. № 5-2(59). С. 37-41.

7. *Костылев Д.С., Кутепова Л.И., Трутанова А.В.* Информационные технологии оценивания качества учебных достижений обучающихся // Балтийский гуманитарный журнал. 2017. № . С. 190-192.

8. *Васильева Л.В.* Анализ методических подходов к построению интегральных экономических показателей // экономические исследования и разработки. 2017. № 12. С. 8-18. URL: http://edrj.ru/article/18–12–17 (дата обращения 24.03.2021).

9. *Орлов А.И.* Предельная теория решений экстремальных статистических задач // Научный журнал КубГАУ. 2017. № 133. С. 579–600.

10. *Хведченя Л.В.* Оценочные шкалы как инструмент педагогической квалиметрии // Адукацыя і выхаванне. 2017. № 3. С. 29–36.

11. *Бронов С.А., Мартынов А.В.* Анализ системы автоматического управления образовательным модулем на примере модуля «векторная алгебра» // Молодой ученый. 2016. № 20 (124). С. 127-131.

12. *Шашлюк Ю.А., Багрецов С.А., Добрынин В.Н.* Управление безопасностью эксплуатации железнодорожных транспортных систем: монография. М.: ВНИИ геосистем, 2018. 390 с.

13. Звягин Л.С. Системный анализ в оптимизации и принятии решений // Всероссийская научная конференция по проблемам управления в технических системах. 2017. № 1. С. 167–170.

14. *Брумштейн Ю.М., Молимонов Д.А.* Математические модели и методы решения задач информационного обеспечения, управления и оценки качества работы операторов в сложных человеко-машинных системах // Вестник АГТУ. Сер. Управление, вычислительная техника и информатика. 2019. № 3. С. 73–89.

15. *Привалов Н.И., Полянина А.С.* Тестовый контроль знаний студентов // Международный журнал прикладных и фундаментальных исследований. 2018. № 4. С. 140–144.

METHODS FOR AGGREGATING STUDENTS' KNOWLEDGE ASSESSMENTS FOR CASES OF REPRESENTING AN EXTERNAL CRITERION IN ORDINAL SCALES

VICTOR D. LIFERENKO
St. Petersburg, Russia

SERGEY A. BAGRETSOV
St. Petersburg, Russia, sergeibagrecov@bk.ru

DENIS V. CHISTYAKOV
St. Petersburg, Russia, serg_chern@mail

KEYWORDS: control of knowledge; rating scales of knowledge; scalar function; indistinct and likelihood distributions; aggregation; method of local variations; kriterialny space.

ABSTRACT

Introduction: the forced transition of universities to distance learning stimulated the further development of automated knowledge control systems, which are one of the most important subsystems of automated learning systems. Among the many problems in the systems of automated control of knowledge, there is the problem of aggregating the assessments received by students for many classes (controls) during the training period. Such a final (integral) assessment should take into account the heterogeneity of the controlled classes when studying the sections and topics of the academic discipline (module) and the uncertainty of the previously obtained estimates. The speci-

fied uncertainty of the estimates can be caused by the unsatisfactory quality of the control due to the lack, for example, of the necessary time reserve, the insufficient quality of the organization and the applied means of control of the students' knowledge. **Purpose:** the purpose of the study is expressed in the analysis and implementation of methods for aggregating assessments of students 'knowledge of higher professional education programs when they perform a set of tasks that have a single semantic content based on the performance of control tasks by students, or on the basis of external observations of the results of students' activities by the teacher. Results: Aggregation of estimates is carried out in ordinal scales. The order of transition to the scale of relations is determined. The results of assessments of the fulfillment of individual stages of educational tasks are presented in ordinal scales or in a scale of names. The integral estimate is considered as a linear scalar function of the initial parameters that determine the results of the control tasks. The parameters of the scalar function are determined on the basis of an estimate of the distances between the distributions of knowledge assessments of students from the training sample, which are set by experienced expert teachers. **Methods:** the proposed methods allow solving the problem of aggregating the assessments received by students for a set of classes (controls) during the training period. Two types of uncertainties in assessing the performance of individual stages of educational tasks are investigated, namely, on the basis of their probabilistic or fuzzy representations. **Results:** the proposed methods take into account, first of all, the holistic nature of the material being studied, the quality control of the implementation of which at certain stages can be carried out only on the basis of external observation of the students' activities. This approach has a rather general character and can be applied when assessing the quality of functioning of objects of different physical nature.

REFERENCES

1. Yakovleva E.I., Shobonov N.A. Individualization of educational process in the modular model of advanced training course using distance learning technologies. *Modern Problems of Science and Education. Surgery.* 2016. No. 6. Pp. 1–8. (In Rus)

2. Grunt E.V., Belyaeva E.A., Lissitsa S. Distance education during the pandemic: new challenges to russian higher education. *Perspectives of Science & Education.* 2020. No. 5(47). Pp. 45–58. doi: 10.32744/pse.2020.5.3 (In Rus)

3. Kim N.F. Reytingovaya sistema otsenki uspevaemosti studentov vuza kak faktor povysheniya kachestva obrazovaniya [The rating system for assessing the progress of university students as a factor in improving the quality of education]. *Molodoy uchenyy* [Young Scientist]. 2015. No. 17(97). Pp. 535–537. (In Rus)

4. Gelman V. Improvement of the forms of performance monitoring in a university. *Sovremennoe obrazovanie* [Modern education]. 2019. No. 2. Pp.52–57. DOI: 10.25136/2409-8736.2019.2.28364 (In Rus)

5. Krivonogov S.V. Development of information system for monitoring and evaluation of knowledge student. *Bulletin NGII*. 2016. No. 8 (63). Pp. 30–41. (In Rus)

6. Larin S.N., Yudinova V.V., Yuryatina N.N. Information technologies for the control of knowledge level of trainees: russian experience. *International Research Journal*. 2017. № 5–2(59). C. 37–41. (In Rus)

7. Kostylev D.S., Kutepova L.I., Trutanova A.V. Information technologies for assessment of quality of educational achievements of training. *Baltic humanitarian journal.* 2017. Vol. 6. No. 3 (20). Pp.190–192. (In Rus)

8. Vasilieva L. Analysis of methodical approaches to the development of integral economic indicators. *Economic Development*. 2017. No. 12. Pp. 8–18. URL: http://edrj.ru/article/18–12–17 (date of access 24.03.2021). (In Rus)

9. Orlov A.I. The limit theory of the solutions of extremal statistical problems. *Scientific Journal of KubSAU*. 2017. No. 133. Pp. 579–600. (In Rus)

10. Khvedchenya L.V. Evaluation scales as a tool of pedagogical qualimetry. *Adukatsya and vyhavanne*. 2017. No. 3. Pp. 29–36. (In Rus)

11. Bronov S.A., Martynov A.V. Analiz sistemy avtomaticheskogo upravleniya obrazovatel'nym modulem na primere modulya "vektornaya algebra" [Analysis of the automatic control system of the educational module on the example of the vector algebra module]. *Molodoy uchenyy* [Young Scientist]. 2016. No. 20 (124). Pp. 127–131. (In Rus)

12. Shashlyuk Yu.A., Bagretsov S.A., Dobrynin V.N. *Upravlenie bezopasnost'yu ekspluatacii zheleznodorozhnyh transportnyh sistem: monografiya* [Safety management of railway transport systems operation: monograph]. Moscow: VNII geosystem, 2018. 390 p. (In Rus)

13. Zvyagin L.S. System analysis in optimization and decision making. *Vserossijskaya nauchnaya konferenciya po problemam upravleniya v tekhnicheskih sistemah* [All-Russian scientific conference on control problems in technical systems]. 2017. No. 1. Pp. 167–170. (In Rus)

14. Brumshteyn Yu. M., Molimonov D.A. Mathematical models and methods for solving problems of information support, quality management and assessment of operators' activity in complex human-machine systems. *Vestnik of Astrakhan State Technical University. Series: Management, Computer Science and Informatics*. 2019. No. 3. Pp. 73–89. (In Rus)

15. Privalov N.I., Polyanina A.S. Test control of knowledge of students. *Pedagogical sciences*. 2018. No. 4. Pp. 140–144. (In Rus)

INFORMATION ABOUT AUTHORS:

LIferenko V.D., PhD, Full Professor, Supertel;
Bagretsov S.A., PhD, Full Professor, Senior Researcher of the Military Space Academy named after A.F. Mozhaysky;
Chistyakov D.V., Lecturer of the Military Space Academy named after A.F. Mozhaysky.

For citation: Liferenko V.D., Bagretsov S.A., Chistyakov D.V. Methods for aggregating students' knowledge assessments for cases of representing an external criterion in ordinal scales. *H&ES Research*. 2021. Vol. 13. No. 2. Pp. 4–14. Doi: 10.36724/2409-5419-2021-13-2-4-14 (In Rus)

ЗАО "НПЦ ИРС" НИУ "МИЭТ" ООО "ИНТЕХ"

МОСКВА 20.10.2021

НАУКА И АСУ 2021

ВСЕРОССИЙСКАЯ МЕЖВЕДОМСТВЕННАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ

по теоретическим и прикладным проблемам развития и совершенствования АСУ и связи специального назначения

http://intech-spb.com/conferences/
konferencia_asu_vka@mail.ru

Тематика конференции включает работу следующих секций:

01

Состояние и перспективы развития современных автоматизированных систем управления специального назначения

02

Математическое, программное и информационно-лингвистическое обеспечение автоматизированных систем управления

03

Безопасность в автоматизированных системах управления специального назначения

04

Применение современных инфокоммуникационных технологий и средств при разработке, техническом обеспечении и эксплуатации автоматизированных систем управления специального назначения

05

Состояние и перспективы развития систем, комплексов и средств адиосвязи специального назначения

06

Проблемы развития автоматизированных систем управления технологическим процессом

КРУГЛЫЙ СТОЛ

Цифровая психология: Использование математических методов при прогнозировании развития личности человека

НИУ МИЭТ
Москва, Зеленоград

20 октября

По итогам конференции отобранные оргкомитетом доклады в виде статей будут опубликованы в журналах из Перечня ВАК, РИНЦ

T-comm • Информация и космос • H&ES Research • I-methods • Техника средств связи

Участие в конференции и публикация материалов в сборнике тезисов БЕСПЛАТНО.

Doi: 10.36724/2409-5419-2021-13-2-16-24

МЕТОД МНОГОМЕРНОЙ ДИНАМИЧЕСКОЙ МАРШРУТИЗАЦИИ В РАДИОСЕТИ

ПАВЛИКОВ
Сергей Николаевич[1]

ЗИМАРЕВА
Евгения Андреевна[2]

БОГДАН
Милена Денисовна[3]

ЦЕПЕЛЕВА
Алена Сергеевна[4]

АННОТАЦИЯ

Введение: Внедрение информационных технологий во все сферы деятельности привело к необходимости разработки новых методов по увеличению качественных параметров систем связи. **Цель исследования** состоит в разработки способа динамической многомерной маршрутизации в телекоммуникационной сети с пакетной передачей данных в соответствии с выбранным критерием. **Методы:** достижение поставленной цели осуществляется в следующей последовательности, на первом этапе рассмотрена задача по расширению функций ретрансляторов, на втором – функции ретранслятора совмещены с функциями коммутатора, а затем рассматривается возможность получения сверхсуммарного эффекта за счет системного использования функций маршрутизации-ретрансляции-коммутации-преобразования. **Результаты:** пространственное кодирование трасс доставки сообщений через реальные и виртуальные ретрансляторы позволяют расширить множество траекторий, через которые по заданным критериям происходит передача пакетов, в точках пространства, согласованных абонентами, происходит формирование пакетов в укрупненные группы, преобразование по методам объединения, разделения, дополнения, удаления и пространственной коммутации до следующих точек ретрансляции. Проведено моделирование процесса формирования пространственного поля точек ретрансляции и визуализация разделения трасс доставки радиосообщений. **Практическая значимость:** формирование в точках преобразований нескольких сигналов реализует процессы получения требуемых пакетов и блоков сообщения, при этом совместная обработка блоков и пакетов в нескольких точках ретрансляции обеспечивают получение новых свойств формируемых и считываемых полей по пространству, например для проведения сеанса с абонентом в зоны неуверенного приема. Новизна представлена в совместном применения расширенного комплекса процессов преобразований. **Обсуждение:** реализация новой совокупности принципов информационного управления ресурсами при выполнении задачи доставки сообщения создает условия значительного увеличения одновременно используемых информационных технологий в единице пространства, что позволяет реализовать новый подход, заключающийся в совмещении процессов разделения и кооперации средств информационного обмена в самоорганизующиеся системы, для совместного выполнения расширенного круга задач.

КЛЮЧЕВЫЕ СЛОВА: метод; многомерная динамическая маршрутизация; радиосеть; ретранслятор; коммутатор.

Сведения об авторах:

[1]к.т.н., профессор, профессор Морского государственного университета имени адмирала Г.И. Невельского, г. Владивосток, Россия, psn1953@mail.ru

[2]аспирант Морского государственного университета имени адмирала Г.И. Невельского, г. Владивосток, Россия, fogetmenots@mail.ru

[3]аспирант Морского государственного университета имени адмирала Г.И. Невельского, г. Владивосток, Россия, milkotim@yandex.ru

[4]аспирант Морского государственного университета имени адмирала Г.И. Невельского, г. Владивосток, Россия, alena.tsepeleva@mail.ru

Для цитирования: *Павликов С.Н., Зимарева Е.А., Богдан М.Д., Цепелева А.С.* Метод многомерной динамической маршрутизации в радиосети // Наукоемкие технологии в космических исследованиях Земли. 2021. Т. 13. № 2. С. 16-24. Doi: 10.36724/2409-5419-2021-13-2-16-24

Введение

Внедрение информационных технологий во все сферы деятельности привело к необходимости разработки новых методов по увеличению качественных параметров систем связи, особенно актуально повышение эффективности в мобильных приложениях, для которых значимы не только абсолютные, но и удельные параметры, например пропускной способности на единицу кубического метра, мобильности, на единицу скорости взаимного перемещения абонентов, результативности на единицу энергетических затрат участников взаимодействия и др. Среди основных тенденций развития инфокоммуникационных технологий особое место отводится децентрализации, расширению функций и передачи их части точкам ретрансляции, доступа, коммутации, самоорганизации сетевых технологий, адаптации под меняющиеся условия и проведение обработки сигналов в самом канале. Поэтому значимость применения сигналов с большой базой возрастает, однако для этого приходится использовать сложные сигналы [1]. Среда радиоканала меняется во времени и по пространству, что оказывает существенное влияние на раскорреляцию широкополосных сигналов, однако чем сложнее сигнал, тем труднее сохранить его устойчивые параметры, характеристики и свойства. Этим объясняется применение во всех сферах деятельности относительно узкополосных сигналов, для которых принципы обнаружения, измерения, селекции и др. меняются незначительно. Радиопромышленность не готова самостоятельно менять узкополосные антенны на пространственные преобразователи с низкой добротностью.

Новые требования к широкополосным технологиям формируют новые задачи, модели сигналов, каналов, систем и их элементов [2]. Цель исследования состоит в разработке способа динамической многомерной маршрутизации в телекоммуникационной сети с пакетной передачей данных в соответствии с выбранным критерием. Интенсивное внедрение телекоммуникационных технологий приводит: к быстрому росту количества одновременно работающих абонентов в ограниченном пространственно временном объеме радиосети, построению сетей на основе принципа многоэтажного частотно-временного планирования, когда в одном и том же частотном диапазоне работают одновременно тысячи информационных каналов между абонентами. И увеличение взаимных помех не мешает качеству радиосвязи, а при определенных значениях параметров применяемых сигналов и согласованных фильтров даже увеличивает помехоустойчивость и разведзащищенность. Использование моделей каналов с виртуальными антеннами, случайными антенными решетками [3] расширяет область информационного взаимодействия и является фундаментом расширения теории и практики инфокоммуникационных технологий. Развитие методов несанкционированного использования информации обострило значимость поиска новых методов защиты, скрытности, помехоустойчивости. Значимость исследований возрастает с учетом широкого внедрения Интернет-вещей, новых критериев, новых технологии M2M, P2P, M2I и др. [4, стр. 238 и 108]. В работе [5] рассмотрены варианты объединения сетей различного уровня и с различными протоколами информационного обмена. Показано, что построение интегрированных гибридных многоуровневых и динамически перестраиваемых сетей является перспективным в решении проблем нехватки связного ресурса. Исследованиям адаптивных маршрутизаторов посвящена работа [6] в которой рассмотрена многослойная архитектура и алгоритм управления постепенным расширением пространства для формирования маршрутов в ответ на возникающие перегрузки. Авторы работ [7 и 8] приводят обзор возможностей и перспектив внедрения скоростных вычислений, основанных на гибридной архитектуре с параллельной обработкой данных и пространственным распределением промежуточных результатов и организации их хранения. В работе [9] описаны процедуры обработки запросов и распределенных потоков данных в сети. В работах [10, 11 и 12] рассмотрены особенности применения методов множественного доступа для повышения эффективности разделения каналов с использованием технологий расширения спектра, ортогональных сочетаний сигналов и их параметров в каналах. В работе [13] сеть рассматривается как слоистая архитектура, в которой маршрутизация и планирование реализуются совместно, что позволяет значительно улучшить пропускную способность, облегчает практическое осуществление, путем сведения к минимуму межслойных операций и реализации принципа модульности, когда одновременно реализуются алгоритмы совместного и раздельного выполнения маршрутизации и планирования, что обеспечивает надежность работы, даже если один из них перестал работать. В работе [14] приводятся особенности поколений до 5G включительно, показаны структурные и технологические изменения, позволяющие обеспечить эффективность расширенного спектра инфокоммуникационных услуг с разнообразными требованиями и в долгосрочной перспективе. В статье [15] Гутковской О.Л и Понамарёва Д. Ю. предложен метод оптимального распределения потоков трафика по минимуму пакетов во всей сети. В работе [16] предложено для повышения эффективности управления трафиком использовать степень снижения вычислительной сложности. Автор в работе [17] предлагает оптимизацию отдельно для беспроводных каналов, данных и видео.

Анализ приведенных работ в данной области показал наличие потенциальных направлений по увеличению эффективности управления динамической маршрутиза-

цией в сети путем расширению процессов и методов обработки в узлах и непосредственно в каналах распространения сигналов.

Предлагается решать проблему управления динамической маршрутизацией путем увеличения количества каналов за счет дополнительных квазивиртуальных узлов ретрансляции с возможным расширением функции по маршрутизации, коммутации и преобразований в виде хранения, сжатия, применение других сигналов физического уровня, модуляции, кодирования и др. Для поиска приемлемых сочетаний технологий преобразований сигналов, каналов, маршрутов и их кратности и последовательности в сети был проведен поиск технических решений, результаты приведены в табл. 1.

Таблица 1

Анализ технических решений по маршрутизации пакетов данных в радиосети

Способ маршрутизации пакетов в радиосети	Недостаток
Пат. Республики Казахстан 10964 / Глухих А.В., Опубл. 15.11.2001, Бюл. №11.	Расчет маршрутов осуществляется только на участках пути от одного промежуточного к другому ближайшему узлу. Ретрансляция сообщений от одного узла к другому ближайшему неизбежно приведет к увеличению количества ретрансляций, загруженности радиоканала и увеличению времени доставки сообщения.
Пат. Республики Казахстан №22321, МПК H04J 3/26, Опубл. 15.02.2010, Бюл. №2	Адрес назначения пакета и адреса принимающего и передающего ретрансляторов содержатся в каждом пакете данных, а адрес следующего ретранслятора, минимально удаленного от адреса назначения пакета, вычисляется процессором каждого принимающего ретранслятора, который по результатам непрерывного контроля пакетов, проходящих по каналу радиосети, формирует собственный динамически изменяющийся список адресов других ретрансляторов и строит математическую модель относительных позиций всех ретрансляторов, потенциально способных участвовать в процессе доставки пакета конечному адресату, что значительно увеличивает трафик сети служебной информацией.
Пат. РФ №2608678. Приор. 17.11.2017, Опубл.23.01.2017,Бюл. № 3	Формирование многомерных маршрутов проводится без полного учета факторов взаимного влияния сигналов входящих в него каналов связи. Управляя только скоростью передачи информации трудно добиться эффективности сети связи.

Анализ приведенных способов показал наличие тенденции по расширению функции и задач ретрансляторов, применение процесса реконфигурации структуры сети и сопряжения с технологиями обработки трафика. На первом этапе рассмотрим задачу с применением в маршрутизаторах ретрансляторов (MR), на втором добавим функции коммутатора (MRK) и затем рассматривается возможность получения сверхсуммарного эффекта за счет системного использования в маршрутизаторах ретранслятора-коммутатора-преобразователя (MRKT).

Первый этап исследования, заключающийся в применении и расширении функции ретрансляторов. Анализ возможных направлений решения проблемы показал, что наибольший потенциал связного ресурса связан с пространственным разделением каналов. Известно, что в этой области эксплуатируется схемы, которые условно обозначается MIMO, SIMO и др. [18, стр. 93]. В них участвуют одиночные или многоканальные передающие и приемные стороны. В работе [19] введен дополнительный элемент ретранслятор (R) или множество ретрансляторов (MR), роль которых может выполнять известное техническое решение или природные отражатели в пространстве радиоканала. В данный момент чаще всего используют передачу по MIMO, когда множество излучающих антенн соединяются с множеством приемных антенн. Однако увеличение количества только элементов антенн на передающей и приемной сторонах не позволяют получить необходимое количество каналов с требуемым качеством связи. Предложено рассмотреть варианты доставки сообщений через распределение трасс в трехмерном пространстве. В итоге получим технологию MIMRMO [19]. В табл. 3 приведена сравнительная оценка эффективности методов пространственного преобразования. Самый эффективный метод MIMRMO с параметрами 2x3x2 получил значение C = 0,845. Чем больше элементов MIMRMO, тем больше пропускная способность, выше скрытность и помехоустойчивость, тем сложнее станции противной стороны своевременно собрать всю информацию. При этом роль ретранслятора может быть выполнена за счет метеорных следов, естественных отражателей в виде характерного рельефа или слоев тропосферы, космических аппаратов, а также неоднородности в канале распространения. В предпочтительном варианте техниче-

ского решения мобильный терминал содержит ретранслятор с передачей части функций базовой станции для расширения зоны действия и обеспечение связанности базовой станции с мобильным терминалов на границе зоны уверенного приема. Другой вариант реализации пространственного разделения является метод, в котором происходит сбор информационного пакета из двух и более блоков, один из которых ключ для получения пакета, и так в каждой точке пространства, известных на передающей и приемной сторонах. Алгоритм пространственного распределения трасс и результаты визуализации работы данного технического решения приведены в работе [20].

Таблица 2

Сравнительная оценка эффективности методов MIMRMO

№ варианта	Передатчики, MI	Ретрансляторы, MR	Приемники, MO	Эффективность, C = log(MI + MR + MO)
1	1	1	1	0,477
2	1	1	2	0,602
3	1	2	2	0,699
4	2	1	1	0,602
5	2	1	2	0,699
6	2	2	2	0,778
7	1	2	1	0,602
8	2	2	1	0,699
9	1	3	1	0,699
10	2	3	1	0,778
11	2	3	2	0,845

Второй этап — совмещение функции ретранслятора с функциями коммутатора (MRK) [19]. Рассмотрим возможности коммутации пакетов в расширенном пространстве методов разделения каналов, сигналов, часть которых приведена в выражении для управления пропускной способности каналов [2]:

$$C = \frac{\Delta F}{(\delta f + \Delta f)} \cdot \frac{T}{(\delta t + \Delta t)} \cdot \frac{\Delta \Theta_{\text{г}}^{\circ}}{(\delta \Theta_{\text{г}} + \Delta \Theta З_{\text{г}}^{\circ})} \times$$
$$\times \frac{\Delta \Theta_{\text{в}}^{\circ}}{(\delta \Theta_{\text{в}} + \Delta \Theta З_{\text{в}}^{\circ})} \cdot \frac{\Delta П}{(\delta П + \Delta ПЗ)} \cdot \frac{\Delta П R_i}{(\delta П R_i + \Delta П R З_i)} \times \quad (1)$$
$$\times \frac{\Delta Ф}{(\delta Ф + \Delta Ф)} \cdot \frac{ДД}{(\delta ДД + \Delta ДД)} \cdot \frac{1}{T},$$

где ΔF — полоса радиосвязи;

δf — ширина радиоканала;

Δf — защитная полоса радиоканала;

T — интервал сеанса связи;

δt — временной интервал элемента радиосообщения;

Δt — защитный интервал;

$\Delta \Theta_{\text{г/в}}^{\circ}$ — сектор в горизонтальной и вертикальной плоскостях;

$\delta \Theta_{\text{г/в}}$ — ширина характеристики направленности канала радиосвязи;

$\Delta \Theta З_{\text{г/в}}$ — защитный интервал по угловому пространству;

$\Delta П$ — интервал поляризационных значений;

$\delta П$ — ширина поляризации радиоканала;

$\Delta ПЗ$ — защитный интервал поляризационных изменений;

$\Delta П R_{\text{i}}$ — размерность пространства ортогональных протоколов каналов на физическом $i = 1$ и других уровнях модели радиосистемы, $i = \overline{1,7}$;

$\delta П R_{\text{i}}$ — расстояние между ортогональными протоколами каналов на физическом $i = 1$ и других уровнях модели радиосистемы, $i = \overline{1,7}$;

$\delta П R k m_{\text{i}}$ — защитный интервал ортогональных k из m протоколов обмена на i уровне, например $i = 1$ на физическом уровне;

$\Delta Ф$ — размерность пространства форм ортогональных сигналов;

$\delta Ф$ — градация форм сигналов;

$\Delta ФЗ$ — защитный интервал форм сигналов;

ДД — динамический диапазон мощности сигналов в радиоканале;

$\delta ДД$ — интервал мощности одного радиоканала;

$\Delta ДД$ — защитный интервал мощности между соседними радиоканалами.

Выбор вариантов управления информационными каналами для достижения требуемого качества обслуживания абонентов может быть возложен на систему, ведущую мониторинг загруженности сети, электромагнитной обстанов-

ки, рельефа местности и др. факторов с целью определения сочетания методов преобразования сигналов, разделения каналов, методов коммутации и маршрутизации, а также скорости перемещения по трассам с временным хранением в узлах сети и их использования в данный момент и в течении допустимого периода сеанса связи. При этом учет возможностей противной стороны для несанкционированного съема информации приведет к расширению процедур с сигналами и каналами по расширенному спектру параметров каналов, например, путем разделения исходного пакета на подпакеты и передачи их одновременно через различные узлы различных сетей (через двух и более сотовых операторов). На рис. 1 и 2 представлены процессы формирования необходимого количества информационных каналов. Прохождение линий указывает на значения выбранных параметров информационного пространства, участвующих в формировании траектории информационного канала. Если трассы по параметрам информационного пространства не пересекаются, значит, взаимное влияние минимально, что соответствует — ортогональности друг другу, а в случае пересечений трасс, но не в точках параметров — квазиортогональны. Один из вариантов реализации способа маршрутизации с расширенным пространством коммутации приведен на рис. 3, где обозначены: 1 — контроль качества входящих в узлы каналов сети связи; 2 — обмен между узлами связи; 3 и 4 — формирование одномерных и соответственно многомерных маршрутов; 5 определение целевой функции; 6 — формирование набора вариантов сочетаний: допустимых методов разделения сигналов, переносчиков пакетов по каналам связи, определяемых в каждом узле для каждого канала связи, в каждом маршруте, а также при их совместном использовании и кратности их применения; 7 — уточнение целевой функции; 8 — управление многомерной динамической маршрутизацией; 9 — уточненный контроль качества входящих в узлы каналов сети связи

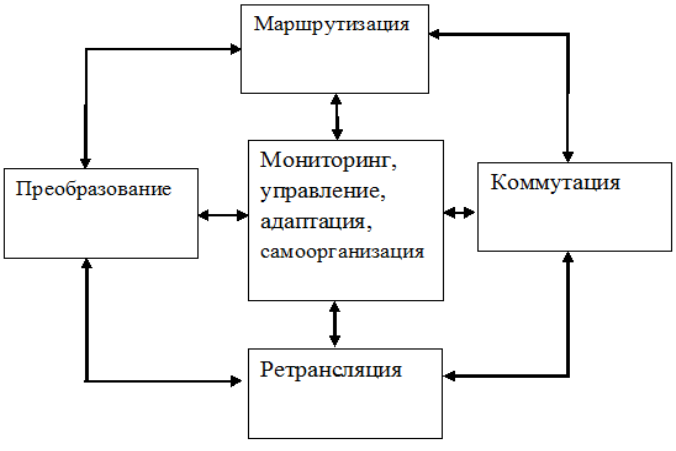

Рис. 1. Процессы управления формирования информационного пространства трасс каналов

Рис. 2. Принцип формирования информационного пространства трасс каналов

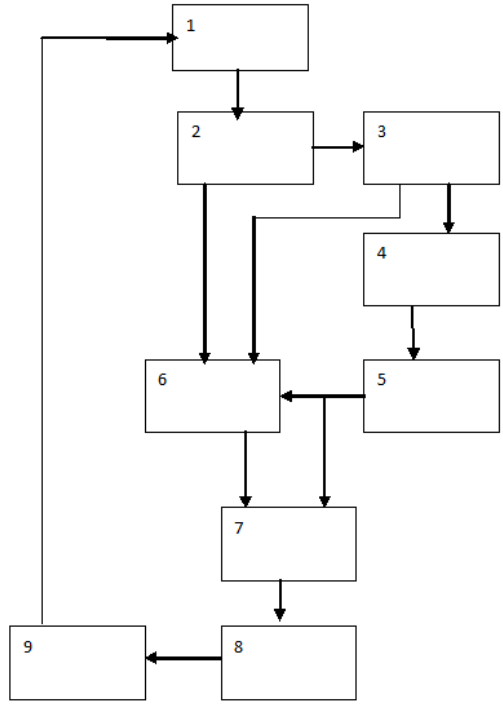

Рис. 3. Способа маршрутизации

Принцип работы. В узлах связи проводят мониторинг входящих в них каналов связи, результаты передаются на доступные узлы связи, формируется набор вариантов маршрутов в виде сочетаний: допустимых методов разделения сигналов, переносчиков пакетов по каналам связи, определяемых в каждом узле для каждого канала связи, в каждом маршруте, а также при их совместном использовании, уточняют целевые функции многомерного динамического распределения маршрутов в сети осуществляют по результатам оценки сформированных вариантов комбинаций (сочетаний) назначенных методов разделения ка-

налов между узлами связи, обеспечивающих выполнение требований при допустимых параметрах и их соотношениях, таких как: вероятность ошибки, отношения сигнал/помеха, скорости передачи, затрат связного ресурса и их комбинаций, а также кратности их использования и выбор оптимальных многомерных динамических трасс связи, при этом управление определяется назначением и согласованием для смежных каналов методов разделения каналов между узлами связи, сочетаний маршрутов и кратности их использования, при этом методы разделения каналов между узлами связи включают технологии использования: различных физических полей, сред распространения сигналов и временного, частотного, поляризационного, кодового, по мощности, по пространственному кодированию, модуляции, кодирования сигнала, скорости передачи, форме сигнала, протоколам обмена, и их комплексирования, а также кратности их использования в соответствии с целевой функцией. Контроль качества входящих в узлы каналов осуществляют по расширенному спектру методов разделения сигналов каналов узлов связи, сочетаний маршрутов и кратности их использования, а также с учетом изменения целевой функции, характеристик и параметров каналов сети связи.

В табл. 3 приведены формулы вычисления количества информационных траекторий, определяемого числом сочетаний из n методов разделения каналов по k. Анализ показывают, что увеличение количества коммутаций сигналов в узлах ретрансляции составляет от несколько десятков до сотен раз.

На третьем этапе планируется расширение функций точек коммутации траекторий до комплексной трансформации в виде ретранслятора-коммутатора-преобразователя (MRKT). Это направление известно как интегрированные, гибридные, комплексные методы формирования трасс доставки сообщений в расширенном классе преобразований физических сигналов и сред распространения: акустические, гидроакустические, электро-магнитные радио и оптические, электрические, а также их разновидностей, например радиосигналы спутниковые, приповерхностные, метеорные и др.

Поэтому работы по поиску технологий и методикам их совместного применения при решении задач доставки сообщений в расширенном понимании эффективности процесса продолжаются. В процессе решения задач были получены следующие результаты. Предложен модернизированный метод многомасштабной многомерной адаптивной маршрутизации, который соответствует протоколам и гибридной и геомаршрутизации.

В отличие от известных технических решений см. табл. 1 методов маршрутизации и методов коммутации предложено следующие.

Структура сети представлена в виде фрактальной математической модели с возможностью планирования трасс от источника к потребителю, в зависимости от суммарной длинны трасс, с несколькими этапами, привязанными к территориям (зонам ответственности).

Для этого в начале на электронной карте размещаются абоненты и узлы сети, и отслеживается их перемещение. Оценивается их подвижность, загруженность, резервы и др. параметры.

Многомасштабное параллельное планирование ведется в мелком масштабе с выбором зон ответственности по территории сети, с учетом критических параметров, влияющих на качество сети, направлений (не менее двух) и ключевых нодов (узлов) через которые предусмотрено обязательное прохождение пакетов с восстановление частей или сообщения в целом, соответствует классу проактивных протоколов. И так для каждой зоны ответственности в направлении от источника информации к получателю.

Многомасштабное планирование продолжается в крупном масштабе внутри зон ответственности, что соответствует реализации класса реактивных протоколов, где реализуется один из вариантов.

Таблица 3

Число информационных траекторий

Числом сочетаний из n методов разделения каналов по k	Формула	Величина при $n = 6$ и $k = 4$	Величина при $n = 7$ и $k = 4$	Величина при $n = 9$ и $k = 4$
Число сочетаний из n по k	$\binom{n}{k} = C_n^k = \frac{n!}{k!(n-k)!}$	15	35	126
Число сочетаний с повторениями из n по k равно	$\binom{n+k-1}{k} = (-1)^k \binom{-n}{k} = = \frac{(n+k-1)!}{k!(n-1)!}$	126	210	495

В отличие от известных в алгоритме заложена возможность отступления (увеличение расстояния до получателя или текущего промежуточного узла, построение обходных путей зон с высоким уровнем шумов или низкой надежности работы узлов).

Пополнение базы данных с описанием прецедентов и вариантов успешного их решения.

Предложено параллельное распределение пакетов и их дублирование по маршрутам с применением реактивных протоколов, с последующей оценкой достоверности и др. качественных параметров, в том числе и затрат в ключевых нодах, как до, так и после преобразований с последующим сравнением полученного результата с результатами от других ключевых нодов, одного и того-же этапа и принятия корректирующих действий в виде регулировки скорости передачи или др.

В роли ключевых нод выбираются или стационарные или малоподвижные узлы на границах

Заключение

Результаты: пространственное кодирование трасс доставки сообщений через реальные и виртуальные ретрансляторы позволяют расширить множество трасс, через которые по заданным критериям происходит передача пакетов, в точках пространства, согласованных абонентами, происходит формирование пакетов в укрупненные группы, преобразование по методам разделения и пространственной направленности передачи до следующих точек MRKT коммутации. Проведено моделирование процесса формирования пространственного поля точек ретрансляции и визуализация разделения трасс доставки радиосообщений. Рассмотренные в работе технологии позволяют компенсировать указанные выше недостатки известных технических решений, например по обеспечению требуемого времени достоверной сообщения абоненту при минимальной вероятности её несанкционированного получения конкурирующей стороной. Практическая значимость заключается в расширении координатного пространства действия сети, в том числе и в зонах неуверенного приема, путем формирование в точках ретрансляции нескольких сигналов над которыми выполняются процессы преобразований.

Реализация новой совокупности принципов информационного управления ресурсами при выполнении задачи доставки сообщения создает условия значительного увеличения одновременно используемых информационных технологий в единице пространства, что позволяет реализовать новый подход не разделения, а кооперации средств информационного обмена в самоорганизующиеся системы, для совместного выполнения расширенного круга задач.

Литература

1. *Скляр Б.* Цифровая связь: Теоретические основы и практическое применение. М.: Вильямс, 2016. 1099 с.

2. *Мочалов А.В., Павликов С.Н., Убанкин Е.И.* Новые направления в развитии телекоммуникационных систем. Владивосток: ВГУЭС. 2016. 116 с.

3. *Винник Л.В., Колесниченко В.И., Литвинов А.В., Мищенко С.Е., Шацкий В.В*. Метод синтеза линейной виртуальной антенной решетки // Журнал радиоэлектроники. 2020. № 1. URL: http://jre.cplire.ru/jre/jan20/2/text.pdf (дата обращения: 20.03.2021). Doi: 10.30898/1684–1719.2020.1.2

4. *Степутин А.Н, Николаев А.Д.* Мобильная связь на пути к 6G. Вологда; Инфра-инженерия, 2018. Т. 2. 420 с.

5. *Yin P., Diamond S., Lin B., Boyd S.* Network Optimization for Unified Packet and Circuit Switched Networks. March 2020. Doi:10.1007/s11081–019–09439–0.

6. *Ping Yin, Sen Yang, Jun Xu, Jim Dai, Bill Lin.* Improving backpressure-based adaptive routing via incremental expansion of routing choices // In Proceedings of the Symposium on Architectures for Networking and Communications Systems. 2017. Pp. 1–12. Doi: 10.1109/ANCS.2017.11

7. *Biswas R., Jiang Z., Kechezhi K., Knysh S., Mandra S., O'Gorman B., Perdomo-Ortiz A., Petukhov A., Realpe-Gomez J., Rieffel E., Venturelli D., Vasko F., Wang Z.* A NASA perspective on quantum computing: Opportunities and challenges; Parallel Computing. 2017. Vol. 64. Pp 81–98. Doi: 10.1016/j.parco.2016.11.002

8. *Weiming Lu, Yaoguang Wang, Jingyuan Juang, Jian Liu, Yapeng Shen, Baogang Wei*. Hybrid storage architecture and efficient MapReduce processing for unstructured data; Parallel Computing. 2017. Vol. 69. Pp. 63–77. Doi: 10.1016/j.parco.2017.08.008

9. *Daichi Amagata, Takashiro Hara, Shojiro Nishio*. Sliding window top-k dominating query processing over distributed data streams; Distributed and Parallel Databases. 2016. Vol. 34. Iss. 4. Pp 535–566. Doi: 10.1007/s10619–015–7187–9

10. *Mahdi Sharifi, Mohammad Jafar Pour Jalali.* Using chaotic sequence in direct sequence spread spectrum based on code division multiple access (DS-CDMA) // ARPN Journal of Engineering and Applied Sciences. 2017. Vol. 12. No. 20. Pp. 5837–5846. Doi: 10.9790/2834–1203021622

11. *Nguyen X., Nguyen C. T., Barlet P., Dojen R.* A novel approach to security enhancement of chaotic DSSS systems // IEEE ICCE2016: 2016 IEEE Sixth International Conference on Communications and Electronics: Novotel, Ha Long, Vietnam. 2016. Pp. 471–476. Doi:10.1109/CCE.2016.7562681

12. *Hordiichuk V.* Method of accuracy increase in radio control systems with orthogonal frequency multiplexising at the consideration of the timer signal constructions use // Advanced Information Systems. 2018. Vol. 2. No. 4. Pp. 108–113. Doi: 10.20998/2522–9052.2018.4.18

13. *Seferoglu H., Modiano E.* Separation of Routing and Scheduling in Backpressure-Based Wireless Networks // *IEEE/ACM Transactions on Networking*. 2016. Vol. 24. No. 3. Pp. 1787–1800. Doi: 10.1109/TNET.2015.2436217

14. *Yazar A., Arslan H.* Flexible Multi-Numerology Systems for 5G New Radio // Journal of Mobile Multimedia. 2018. Vol. 14. No. 4. Pp. 367–394. Doi: https://doi.org/10.13052/jmm1550–4646.1442

15. *Гутковкая О.Л., Пономарёв Д.Ю.* Применение ортогональной модели телекоммуникационной сети для решения задачи оптимального распределения трафика // Кибернетика и программирование. 2017. № 1. С. 11–29. Doi: 10.7256/ 2306–4196.2017.1.21810URL

16. *Симаков Д.В.* Управление трафиком в сети с высокой динамикой метрик сетевых маршрутов // Интернет-журнал «Науковедение». 2016. Т. 8. № 1. URL: http://naukovedenie.ru/PDF/ 60TVN116.pdf (дата обращения: 20.03.2021). Doi: 10.15862/ 60TVN116

17. *Zhang X.* LTE optimization engineering handbook Hoboken, NJ, USA: John Wiley & Sons Singapore Pte. Ltd, 2018. 827 p.

18. Степутин А.Н, Николаев А.Д. Мобильная связь на пути к 6G. Вологда; Инфра-инженерия, 2018, Т 1. 384 с.

19. *Стволовая А.К., Павликов С.Н.* Разработка алгоритма и визуализация пространственного распределения трасс доставки сообщений в условиях угрозы несанкциниорованного съема // Современные наукоемкие технологии. 2018. № 2. С. 104–109. URL: http://top-technologies.ru/ru/article/view?id=36914 (дата обращения: 20.03.2021).

20. Свидетельство на программный продукт для ЭВМ РФ 2018616795. Программа имитации разделения трасс доставки радиосообщений / Павликов С.Н., Стволовая А.К., Котович Е.Е. Заявл. 06.06.2018.

MULTIDIMENSIONAL DYNAMIC ROUTING METHOD IN RADIO NETWORK

SERGEJ N. PAVLIKOV
Vladivostok, Russia, psn1953@mail.ru

EVGENIA A. ZIMAREVA
Vladivostok, Russia, fogetmenots@mail.ru

MILENA D. BOGDAN
Vladivostok, Russia, milkotim@yandex.ru

ALENA S. CEPELEVA
Vladivostok, Russia, alena.tsepeleva@mail.ru

KEYWORDS: method multidimensional dynamic routing radio network; the relay switch.

ABSTRACT

Introduction: The introduction of information technology has accelerated the development of methods to improve communication systems. The aim of the study is to improve the efficiency of the network with batch data transmission. **Methods:** achieving the goal is carried out in three stages, the first stage is considered the task of expanding the functions of repeaters, the second – the functions of the relay are combined with the functions of the switch, and then considered the possibility of obtaining a super-sumar effect through the system use of the repeater-switch-converter. **Results:** spatial coding of message delivery tracks through real and virtual repeaters allows you to expand the many routes through which, according to the specified criteria, packages are transferred, in the points of space agreed by subscribers is the formation of packages in enlarged groups, transformation by separation methods and spatial switching to the following switching points. The process of forming a spatial field of relay points and visualization of the separation of routes of radio transmissions was carried out. **Practical significance:** the formation of multiple signals at the focus points implements transformation processes, and the joint processing of processes at multiple relay points provides a switching through the space, for example, for the coordinate space, the lighting of the area of uncertain reception. **Discussion:** The implementation of a new set of resource management principles in the delivery of communication creates conditions for a significant increase in the simultaneous use of information technologies in a unit of space, allowing for the implementation of a new approach, which is to combine the processes of separation and cooperation of the means of information exchange into self-organizing systems, to jointly perform an expanded range of tasks.

REFERENCES

1. Sklar B. *Digital Communication: Theoretical Basics and Practical Application*. Moscow: Williams, 2016. 1099 p. (In Rus)
2. Mochalov A.V., Pavlikov S.N., Ubankin E.I. *New directions in the development of telecommunications systems: monograph*. Vladivostok: VSUES. 2016. 116 p. (In Rus)
3. Vinnik L.V., Kolesnichenko V.I., Litvinov A.V., Mishchenko S.E., Shatsky V.V. Method of synthesis of linear virtual antenna grille. *The Journal of Electronics*. 2020. No. 1. URL: http://jre.cplire.ru/jre/jan20/2/text.pdf. doi: 10.30898/1684-1719.2020.1.2 (In Rus)
4. Steputin A.N., Nikolaev A.D. *Mobile Communication on the way to 6G*. Infra-engineering, 2018. Vol. 2. 420 p. (In Rus)
5. Yin P., Diamond S., Lin B., Boyd S. *Network Optimization for Unified Packet and Circuit Switched Networks*. March 2020. Doi:10.1007/s11081–019–09439–0.
6. Ping Yin, Sen Yang, Jun Xu, Jim Dai, Bill Lin. Improving backpressure-based adaptive routing via incremental expansion of routing choices. *In Proceedings of the Symposium on Architectures for Networking and Communications Systems.* 2017. Pp. 1–12. Doi: 10.1109/ANCS.2017.11
7. *Biswas R., Jiang Z., Kechezhi K., Knysh S., Mandra S., O'Gorman B., Perdomo-Ortiz A., Petukhov A., Realpe-Gomez J., Rieffel E., Venturelli D., Vasko F., Wang Z. A NASA perspective on quantum computing: Opportunities and challenges; Parallel Computing*. 2017. Vol. 64. Pp 81–98. Doi: 10.1016/j.parco.2016.11.002
8. Weiming Lu, Yaoguang Wang, Jingyuan Juang, Jian Liu, Yapeng Shen, Baogang Wei. *Hybrid storage architecture and efficient MapReduce processing for unstructured data*; *Parallel Computing*. 2017. Vol. 69. Pp. 63–77. Doi: 10.1016/j.parco.2017.08.008
9. Daichi Amagata, Takashiro Hara, Shojiro Nishio. *Sliding window top-k dominating query processing over distributed data streams; Distributed and Parallel Databases*, 2016. Vol. 34. Iss. 4. Pp 535–566. Doi: 10.1007/s10619–015–7187–9
10. Mahdi Sharifi, Mohammad Jafar Pour Jalali. Using chaotic sequence in direct sequence spread spectrum based on code division multiple access (DS-CDMA). *ARPN Journal of Engineering and Applied Sciences*. 2017. Vol. 12. No. 20. Pp. 5837–5846. Doi: 10.9790/2834–1203021622
11. Nguyen X., Nguyen C.T., Barlet P., Dojen R. A novel approach to security enhancement of chaotic DSSS systems. *IEEE ICCE2016: 2016 IEEE Sixth International Conference on Communications and Electronics: Novotel*, Ha Long, Vietnam. 2016. Pp. 471–476. Doi:10.1109/CCE.2016.7562681
12. Hordiichuk V. Method of accuracy increase in radio control systems with orthogonal frequency multiplexising at the consideration of the timer signal constructions use. *Advanced Information Systems*. 2018. Vol. 2. No. 4. Pp. 108–113. Doi: 10.20998/2522–9052.2018.4.18
13. Seferoglu H., Modiano E. Separation of Routing and Scheduling in Backpressure-Based Wireless Networks. *IEEE/ACM Transactions on Networking*. 2016. Vol. 24. No. 3. Pp. 1787–1800. Doi: 10.1109/TNET.2015.2436217
14. Yazar A., Arslan H. Flexible Multi-Numerology Systems for 5G New Radio. *Journal of Mobile Multimedia*. 2018. Vol. 14 . No. 4. Pp. 367–394. Doi: https://doi.org/10.13052/jmm1550–4646.1442
15. Gutkovka O.L., Ponomar in D.J. Applying an orthogonal model of the telecommunications network to solve the problem of optimal traffic distribution. *Cybernetics and programming*. 2017. No. 1. Pp. 11-29. Doi: 10.7256/2306–4196.2017.1.21810URL (In Rus)
16. Simakov D.V. Traffic Management in a network with high dynamics of network route metrics. *Internet magazine "SCIENCE"*. 2016. Vol. 8. No. 1. URL: http://naukovedenie.ru/PDF/60TVN116.pdf (date of access 09.03.2021). Doi: 10.15862/60TVN116 (In Rus)
17. Zhang X. LTE optimization engineering handbook Hoboken, NJ, USA: John Wiley & Sons Singapore Pte. Ltd, 2018. 827 p.
18. Steputin A.N., Nikolaev A.D. *Mobile Communication on the way to 6G*. Infra-engineering, 2018. Vol. 1. 384 p. (In Rus)
19. Stvolovaj A.K., Pavlikov S.N. Development of an algorithm and visualization of spatial distribution of routes of delivery of messages in the face of the threat of unauthorized removal. *Modern Science-Intensive Technologies.* 2018. No. 2. Pp. 104–109. URL: http://top-technologies.ru/ru/article/view?id=36914 (date of access 09.03.2021). (In Rus)
20. Testimony on the software product for computer RF 2018616795. The program simulates the separation of routes of radio delivery / Pavlikov S.N., Stvolovaj A.K., Kotovich E.E. Pablished 06.06.2018. (In Rus)

INFORMATION ABOUT AUTHORS:

Pavlikov S.N., PhD, Professor Full, Professor of Admiral Nevelskoy Maritime State University;
Zimareva E.A., postgraduate student of Admiral Nevelskoy Maritime State University;
Bogdan M.D., postgraduate student of Admiral Nevelskoy Maritime State University;
Cepeleva A.S., postgraduate student of Admiral Nevelskoy Maritime State University;

For citation: Pavlikov S.N., Zimareva E.A., Bogdan M.D., Cepeleva A.S. Multidimensional dynamic routing method in radio network. *H&ES Research*. 2021. Vol. 13. No. 2. Pp. 16–24. Doi: 10.36724/2409-5419-2021-13-2-16-24 (In Rus)

Doi: 10.36724/2409-5419-2021-13-2-25-34

ПРОГРАММНЫЙ КОМПЛЕКС ОЦЕНКИ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ БЕЗ ИСХОДНЫХ ТЕКСТОВ

САМАРИН
Николай Николаевич

АННОТАЦИЯ

Введение: цифровизация затрагивает все отрасли человеческой деятельности, вследствие чего создается разнообразное программное обеспечение, реализующее бизнес-логику и технические процессы в сложных системах. В таких условиях возрастает актуальность задачи распознавания вредоносного программного обеспечения. Решение данной задачи осложняется отсутствием исходных текстов программ и необходимостью оперативного принятия решения о наличии либо отсутствии вредоносного функционала. **Цель исследования:** целью исследования является создание подхода к оценке информационной безопасности программного обеспечения без исходных текстов. Предлагается в основу подхода положить использование гипервизора, обеспечивающего контроль за работой программного обеспечения с памятью как ключевой характеристики наличия/отсутствия его вредоносного функционала. В качестве метрики безопасности предлагается вычислять оценку безопасности программного обеспечения. **Методы:** решение поставленной задачи основано на использовании механизма виртуализации, обеспечивающего контроль всех операций над памятью, реализуемых программным обеспечением, и на использовании методов теории вероятностей для получения комплексной оценки безопасности, учитывающей надежность функционирования программного обеспечения и его безопасность. **Результаты:** разработана методика получения комплексной оценки безопасности функционирования программного обеспечения, учитывающая надежность функционирования программного обеспечения; сетевую безопасность – выявленные в результате сканирования уязвимости и сетевые порты; потенциально небезопасные изменения в файловой системе и реестре, а также потенциально опасные операции, связанные с использованием памяти. Описана архитектура макета программного комплекса, реализующего предложенный подход, выполнена его экспериментальная апробация, в результате которой высокую оценку безопасности получили только образцы штатного программного обеспечения. **Практическая значимость:** разработанный макет может быть использован для автоматизированного анализа программного обеспечения, функционирующего в различных сложных системах. Важным достоинством макета является его масштабируемость и доверенность, обеспечиваемая за счет использования средств виртуализации, не позволяющих нанести вред работе вычислительной системы в случае обнаружения вредоносного программного обеспечения.

КЛЮЧЕВЫЕ СЛОВА: информационная безопасность; вредоносное программное обеспечение; гипервизор; модульная архитектура.

Сведения об авторе:

начальник научно-исследовательского отделения № 6 Научно-исследовательского института «Квант», г. Москва, Россия, samarin_nik@mail.ru

Для цитирования: *Самарин Н.Н.* Программный комплекс оценки информационной безопасности программного обеспечения без исходных текстов // Наукоемкие технологии в космических исследованиях Земли. 2021. Т. 13. № 2. С. 25–34. Doi: 10.36724/2409-5419-2021-13-2-25-34

Введение

Особую сложность представляет задача распознавания ВПО в случаях, когда исходные тексты ПО отсутствуют [1–4]. В таких случаях ручной анализ требует больших временных затрат, следовательно, становится малоэффективным при защите от кибератак. Статический анализ с использованием дизассемблера позволяет лишь приблизительно восстановить высокоуровневый код, а большинство вариантов применения динамического анализа требуют больших вычислительных затрат. Потребление вычислительных ресурсов требуется минимизировать, поскольку современные вычислительные системы, интегрированные с отраслями деятельности человека, включают большое число малоресурсных компонентов — датчики и контроллеры.

В таких условиях актуальной является задача создания программного решения, обеспечивающего автоматизированный анализ безопасности ПО без открытых текстов и являющегося расширяемым и доверенным, то есть, гарантирующим, что в случае обнаружения вредоносного ПО не будет нанесен вред работе целевой системы.

1. Подход, составляющий основу программного комплекса

В настоящее время не существует универсального и оптимального решения задачи обнаружения ВПО без исходных текстов, особенно если речь идет о потребности в минимизации используемых вычислительных ресурсов [5–8]. В основе предлагаемого к разработке программного комплекса автоматизированного анализа безопасности ПО без открытых текстов лежат следующие положения:

1. Для повышения эффективности обнаружения ВПО и обеспечения доверенной работы комплекса при анализе безопасности следует использовать механизмы виртуализации [9–13].

2. Анализ безопасности должен быть комплексным и включать как оценку надежности работы ПО, так и оценку того, насколько безопасно он функционирует [14].

3. Анализ безопасности должен основываться, в том числе, на оценке ресурсов, потребляемых при работе исследуемого образца ПО, и на контроле работы с памятью [15–17].

Процесс в виртуальной операционной среде с полным контролем действий ПО и отслеживанием алгоритма его работы осуществляется более эффективно. Поэтому целесообразно использовать такой механизм виртуализации как гипервизор. Под гипервизором принято понимать программное или микропрограммное обеспечение, позволяющее виртуализировать системные ресурсы вычислительных систем.

Оценка надежности ПО важна с точки зрения той системы, с которой данное ПО интегрируется. Все чаще появляются сложные системы, в которых программные механизмы управляют техническими и бизнес-процессами. Отказы надежности таких программных механизмов приведут к значительным финансовым потерям из-за простоя систем и необходимости наладки поврежденной инфраструктуры [18–23].

Вместе с тем, ориентироваться только на надежность нельзя, поскольку современное ВПО обладает широким функционалом по сетевому взаимодействию и также способно перехватывать контроль над управлением всей системы.

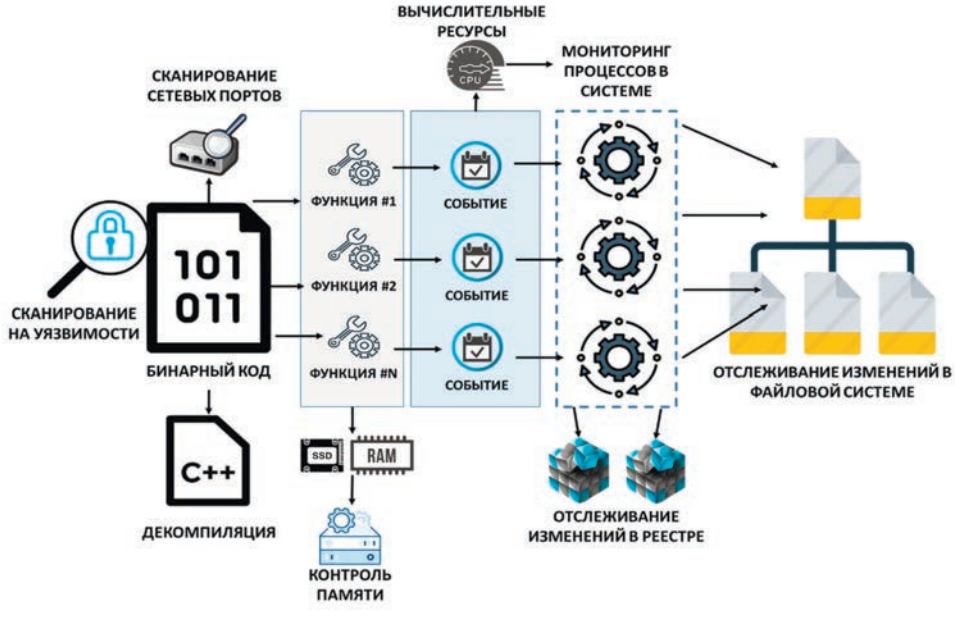

Рис. 1. Общая схема подхода к анализу безопасности ПО без исходных текстов

На (рис. 1) представлена общая схема подхода к анализу безопасности ПО без исходных текстов, на ней четко показаны объекты и параметры, которые необходимо контролировать для получения итоговой оценки безопасности.

Предложена оценка безопасности, включающая следующие частные оценки:

1. Оценку надежности функционирования ПО в терминах теории вероятностей. ПО тестируется *n* раз, число успешно пройденных испытаний обозначается как n^+, а число отказов и случаев невыполнения/некорректного выполнения функций как n^-. Получаем набор $\{n_1^-, n_2^-, \ldots, n_k^-\}$, разделив все случаи отказов/невыполнения/некорректного выполнения функций на *k* групп. Введем коэффициент $\delta_i, i = \overline{1,k}$ для каждой группы. Он характеризует вероятность устранения проблемы безопасности. Оценка надежности *R* функционирования ПО вычисляется по формуле:

$$R = \frac{n^+ + \sum_{i=1}^{k} \delta_i n_i^-}{n}.$$

2. Оценку безопасности на основе выявленных в результате сканирования уязвимостей и сетевых портов, а также выявленных потенциально небезопасных изменений в файловой системе и реестре. Введены два атрибута, характеризующие безопасное функционирование ПО: A_1 характеризует сетевую безопасность, A_2 характеризует безопасность работы системы. Для оценки атрибута A_1 используем вероятностную метрику P_1: $P_1 = P_T \bullet P_U \bullet L$, где P_T — вероятность реализации угрозы безопасности, связанной с данным типом уязвимости; P_U — вероятность использования данной уязвимости; L — степень тяжести последствий от реализации угрозы безопасности, связанной с данным типом уязвимости (пятиуровневая шкала). Для оценки A_2 используется формула

$$P_2 = \frac{n^-}{n} \bullet P_R \bullet L,$$

где $\frac{n^-}{n}$ — отношение числа выявленных потенциально небезопасных изменений к общему числу изменений (операций), произошедших за время анализа в системе,

P_R — вероятность возникновения угрозы безопасности в результате данного изменения.

3. Оценку безопасности на основе выявленных с использованием гипервизора потенциально опасных операций, связанных с памятью. При оценке используется два типа проблем безопасности: k_1, приводящий к небезопасной работе с памятью, и k_2, приводящий к замедлению работы системы из-за чрезмерной загрузки памяти процессами ВПО. С типом k_1 будем связывать вероятность использования данной потенциально опасной операции при реализации кибератаки P_{k_1}, значение вероятности берется с учетом вероятности реализации подобных инцидентов безопасности в рассматриваемой индустрии. С типом k_2 — вероятность чрезмерной загрузки памяти к следующему моменту времени P_{k_2}. Общая оценка безопасности вычисляется по формуле

$$P_H = \left(1 - \frac{k_1 \bullet P_{k_1}}{n}\right) + \frac{k_2 \bullet e^{-\lambda_i t_i}}{n},$$

где *n* — общее число прогонов ПО;

i–1 — номер последнего на данный момент времени события типа k_2;

C_D — коэффициент пропорциональности, вычисляемый по формуле

$$C_D = \frac{\sum_{i=1}^{k-1} \frac{1}{n-i+1}}{\sum_{i=1}^{k-1} t_i},$$

(t_i – временные отсчеты, в которые были зафиксированы события типа k_2).

Итоговая вероятностная оценка P_{sec} вычисляется как среднее значение вероятностей:

$$P_{sec} = \frac{R + P_V + P_H}{3} = \frac{\frac{n^+ + \sum_{i=1}^{k} \delta_i n_i^-}{n} + \left(1 - \frac{n^-}{n} \bullet L^2 \bullet P_T \bullet P_U \bullet P_R\right) + \left(\left(1 - \frac{k_1 \bullet P_{k_1}}{n}\right) + \frac{k_2 \bullet e^{-\lambda_i t_i}}{n}\right)}{3}$$

2. Архитектура программного комплекса автоматизированного анализа безопасности ПО без открытых текстов

Архитектура программного комплекса может быть рассмотрена с двух сторон:

− с функциональной стороны;

− с аппаратно-технической стороны.

Несомненно, оба эти взгляда будут соответствовать друг другу, однако для лучшего понимания того, как должен работать программный комплекс, и какие для его реализации потребуются технические и программные средства, необходимо описать обе точки зрения.

Важно отметить, что выбрана модульная архитектура программного комплекса, поскольку она является расширяемой и тем самым обеспечивает широкие возможности по развитию ее функционала.

Начнем с функциональной архитектуры программного комплекса оценки информационной безопасности ПО без исходных текстов. Она представляет собой набор функциональных блоков системы, связанных между собой и направленных на решение поставленной задачи анализа безопасности ПО. Она включает следующие функциональные модули:

1. Виртуальный модуль — среда исследований образцов ПО на предмет безопасности. Модуль включает в себя

гипервизор, набор виртуальных машин, базу данных сценариев тестирования ПО.

2. Модуль мониторинга, обеспечивающий после запуска ПО:

− отслеживание изменений в файловой системе и реестре;

− мониторинг процессов вычислительной системы;

− анализ сетевой безопасности системы — сбор и анализ сетевого трафика в процессе функционирования ПО, сканирование сетевых портов и сканирование на предмет наличия уязвимостей.

3. Модуль анализа и контроля, обеспечивающий:

− декомпиляцию бинарного образца ПО и получение программного кода на языке высокого уровня, приближенного к реальному коду ПО;

− контроль обращений процессора к памяти при запуске ПО;

− контроль использований памяти при запуске ПО;

− формирование визуального представления такого использования памяти.

4. Модуль оценки безопасности функционирования ПО, реализующий оценку вероятности того, что исследуемый образец ПО реализует надежное и безопасное функционирование в вычислительной системе.

5. Модуль представления результатов, обеспечивающий представление проведенного анализа безопасности ПО в удобном для оператора/исследователя форме.

Также в рамках функциональной архитектуры присутствуют различные базы данных, в которых содержатся промежуточные результаты анализа, список исследуемых образцов ПО и результаты оценки безопасности.

Функциональная архитектура программного комплекса оценки информационной безопасности ПО без исходных текстов представлена на (рис. 2).

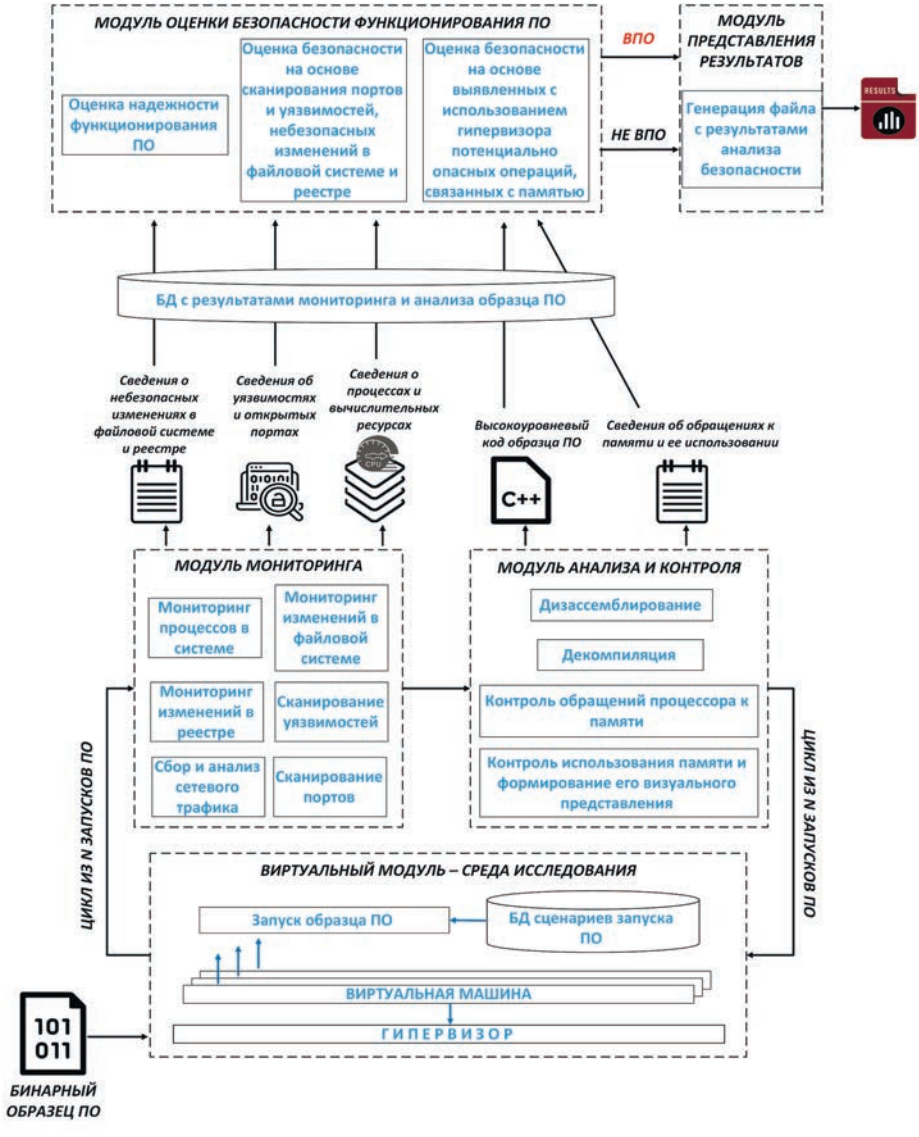

Рис. 2. Функциональная архитектура программного комплекса

Техническое представление архитектуры программного комплекса оценки информационной безопасности ПО без исходных текстов опирается преимущественно на аппаратные и программные средства, которые необходимо использовать для реализации и функционирования комплекса (рис. 3).

К необходимым средствам для реализации программного комплекса относятся:

1. Сервер, на котором будет выполняться анализ ПО, включающий следующие компоненты:

− установленная операционная система хоста;

− гипервизор;

− набор виртуальных машин, на каждой из которых установлена гостевая операционная система;

− системные файлы и библиотеки, необходимые для работы гостевой операционной системы;

− приложение или скрипт, реализующий автоматический запуск нескольких прогонов ПО в соответствии с заранее сформированными сценариями;

2. Сервер баз данных, на котором расположены:

− база данных, в которой хранятся сведения об исследуемых образцах ПО;

− база данных со сценариями тестирования образцов ПО;

− базы данных для хранения промежуточных и конечных результатов мониторинга;

− база данных для хранения декомпилированных образцов кода ПО;

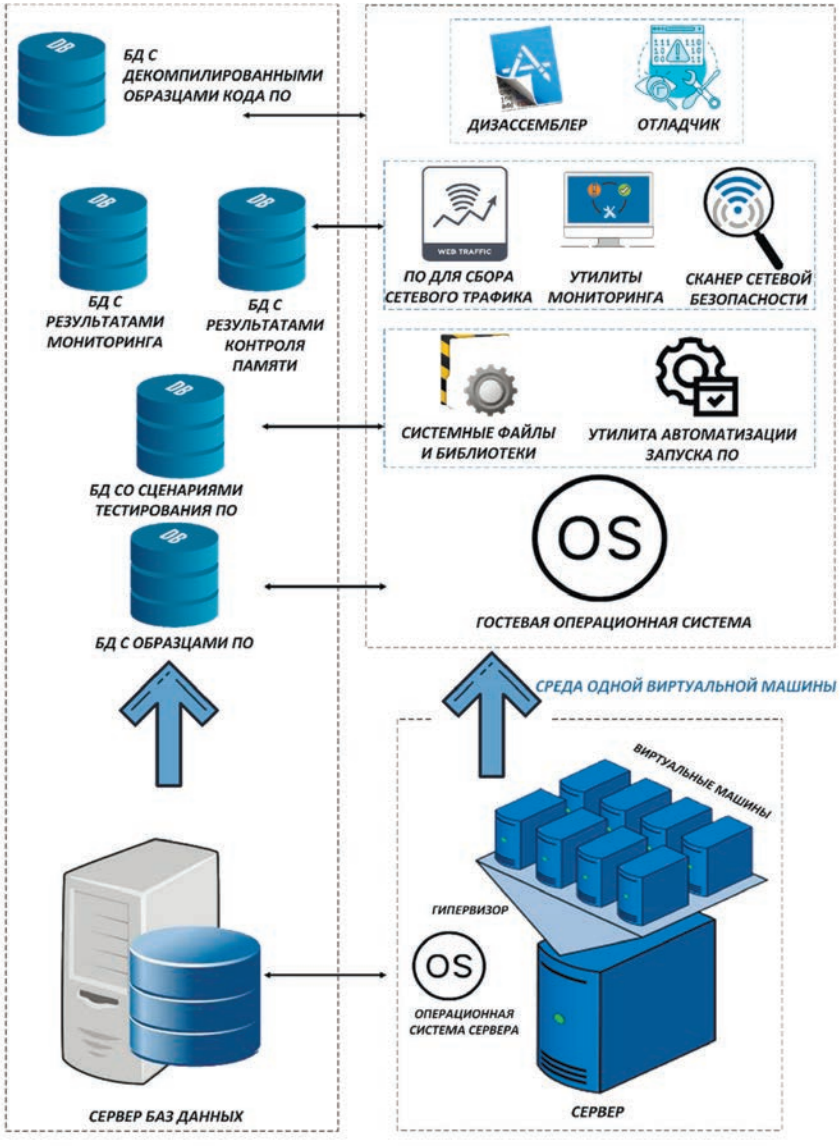

Рис. 3. Графическое представление модели безопасного функционирования ПО

– базы данных для результатов работы алгоритмов контроля обращений процессора к памяти при запуске ПО, контроля использований памяти при запуске ПО и формирования визуального представления такого использования памяти.

3. Утилиты мониторинга процессов, отслеживания изменений в файловой системе и реестре.

4. Программное обеспечение для сбора сетевого трафика.

5. Сканер сетевой безопасности, обеспечивающий поиск открытых портов и сканирование на предмет наличия уязвимостей.

6. Дизассемблер и прочие инструменты статического анализа.

7. Отладчик, с возможностью получения динамических параметров исполнения ПО.

2. Реализация макета программного комплекса

Реализация макета программного комплекса также предполагает выбор необходимого инструментария. В качестве среды исследования выбрана операционная система DOS, поскольку при ее хорошей изученности, ее процессы также могут быть интерпретированы на более сложные системы семейства Windows.

Для мониторинга изменений в тестируемой системе с установленным исследуемым ПО используются свободно распространяемые утилиты от разработчиков Microsoft:

– Process Hacker — исследования и мониторинга процессов;

– FileMon — отслеживания изменений в файловой системе;

– RegMons — отслеживания изменений в реестре.

Сканирование портов реализуется утилитой nmap, в автоматизированном режиме — с помощью python скрипта, использующего библиотеку python-3 nmap. Сканирование на уязвимости реализуется с помощью OpenVas.

В качестве механизма виртуализации выбран гипервизор Bochs. Bochs — это бесплатный эмулятор, на основе которого можно воссоздать на своем устройстве операционную систему OS, VM, VSE. Данный программный продукт включает в себя эмуляцию процессоров архитектуры x86, различных видеоадаптеров и версий прошивки BIOS. Bochs может эмулировать такие процессоры, как AMD64, Pentium, Pentium Pro, 386, 486 и другие. Также поддерживает MMX, SSE3, SSE4, 3DNow, SSE, SSE2. Поскольку Bochs эмулирует множество операционных систем, его используют для отладки новинок и запуска старых игр и приложений.

К основным возможностям эмулятора Bochs можно отнести следующее:

– эмуляция операционных систем Windows, Linux, DOS;

– проверка состояния памяти и регистров процессора;

– поиск ошибок в программах и приложениях с помощью графического отладчика;

– эмуляция процессоров 386, 486, Pentium всех поколений или x86–64, включая опциональные инструкции MMX, SSEx i 3DNow;

– поддержка вычисления типов памяти;

– поддержка Broadwell ULT в CPUDB;

– проверка и отладка VGA;

– сохранение и восстановления расширенных настроек интерфейса операционной систем;

– поддержка образов Oracle VM VirtualBox;

– поддержка MIDI UART.

При стендовых имитационных испытаниях анализ трафика осуществляется с использованием Wireshark (tshark с автоматизацией на базе pyshark), имитация внешней сети, включая параметры ntp, производится с использованием пакета inetsim. Для автоматизации испытаний запуск и установка параметров осуществляется с помощью управляющего скрипта netmanager.

Для проверки теоретических положений, изложенных в данной главе, на конкретных программных продуктах, предложена следующая структурная схема макета (рис. 4).

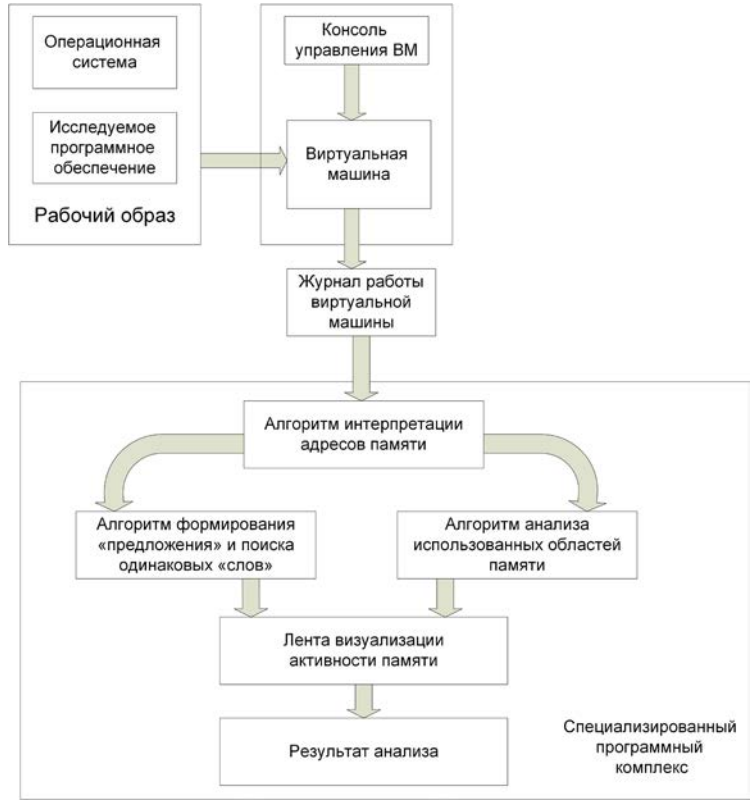

Рис. 4. Структурная схема макета программного комплекса

Рабочий образ представляет собой операционную систему MSDOS6.22 и исследуемый образец ПО. Указанный образ формируется исследователем перед проведением работ.

Важно отметить, что за счет использования гипервизора в составе программного комплекса (заявленного еще на этапе разработки модели) в случае внештатной ситуации можно «откатить» систему в последнее работоспособное состояние и продолжить работу.

На части программного комплекса получены свидетельства о государственной регистрации программы для ЭВМ.

3. **Экспериментальная апробация макета программного комплекса**

При проведении экспериментальных исследований использовались 10 образцов ВПО и 2 штатные программы, функционирующие корректно. Частные оценки надежности и безопасности получились следующими (табл. 1–3).

Из табл. 1 видно, что в основном исследуемое ПО функционировало корректно, и отказы ПО были чрезвычайно редкими. Очевидно, что одной оценки надежности функционирования ПО недостаточно для того, чтобы сделать выводы о том, является ли ПО вредоносным или нет, поскольку оценка надежности штатных программ незначительно отличается от надежности ВПО.

Оценка безопасности по итогам сетевого сканирования и мониторинга изменений в файловой системе и реестре является неоднозначной. Для ряда образцов ВПО она действительно оказалась весьма низкой (Ukraine, Omsk622, Yosha, Ah), однако для некоторых вредоносных программ оценка безопасности довольно высока (Abbas, Bomzh, Green), также она высока и для штатных программ, функционирующих корректно.

Из табл. 3 видно, что те образцы ВПО, которые на втором этапе оценки характеризовались довольно высокой вероятностью, на третьем этапе оценки проявили себя как однозначно вредоносные: их оценка безопасности низкая. При этом, для штатных программ она высокая, что позволяет четко выделить два класса образцов — класс легитимного ПО и класс вредоносного ПО.

Из табл. 4 видно, что высокую оценку получили только штатные программы, что подтверждает успешность апробации разработанного программного комплекса. При этом, задействование вычислительных ресурсов виртуальной машины при анализе безопасности было незначительным, что говорит о высокой производительности комплекса.

Таблица 1

Результаты оценки надежности

AVV	Abbas	Adi	Ah	Bomzh	Green	Omsk622	Ukraine	Yosha	Keyrus	ШП №1	ШП №2
0,75	0,62	0,74	0,76	0,72	0,84	0,82	0,84	0,85	0,81	0,94	0,96

Таблица 2

Результаты оценки безопасности по итогам сетевого сканирования и мониторинга изменений в файловой системе и реестре

AVV	Abbas	Adi	Ah	Bomzh	Green	Omsk622	Ukraine	Yosha	Keyrus	ШП №1	ШП №2
0,12	0,48	0,26	0,11	0,48	0,41	0,09	0,05	0,14	0,04	0,87	0,96

Таблица 3

Результаты оценки безопасности по итогам контроля работы с памятью

AVV	Abbas	Adi	Ah	Bomzh	Green	Omsk622	Ukraine	Yosha	Keyrus	ШП №1	ШП №2
0,09	0,03	0,11	0,09	0,09	0,08	0,07	0,15	0,18	0,14	0,97	0,99

Таблица 4

Итоговая оценка безопасности

AVV	Abbas	Adi	Ah	Bomzh	Green	Omsk622	Ukraine	Yosha	Keyrus	ШП №1	ШП №2
0,32	0,38	0,37	0,32	0,43	0,44	0,33	0,35	0,39	0,33	0,93	0,97

Заключение

В статье описан подход, лежащий в основе программного комплекса оценки информационной безопасности ПО без исходных текстов. Описаны этапы анализа безопасности и предложена комплексная вероятностная оценка безопасности ПО, учитывающая и надежность его функционирования, и возможные проблемы безопасности — что полностью соответствует новым условиям, возникшим в результате цифровизации и разработки ПО.

Описана архитектура программного комплекса, как с точки зрения его функциональных возможностей, так и с точки зрения необходимого программно-технического состава. Реализован макет программного комплекса, выполнена его оценка эффективности, подтвержденная результатами экспериментальных исследований.

Разработанный макет позволяет выполнить автоматизированный анализ безопасности ПО без исходных текстов, с учетом следующих аспектов:

– контроль областей памяти системы, с которыми взаимодействует исследуемый образец ПО;

– мониторинг процессов и изменений, происходящих в системе с момента запуска сценариев тестирования ПО;

– анализ сетевой безопасности при запуске исследуемого ПО;

– контроль использования памяти и вычислительных ресурсов системы, что особенно важно в условиях ограниченных вычислительных мощностей;

– проведение исследований в реальном масштабе времени с возможностью записи результатов для повторного анализа.

Наряду с этим, программный комплекс оценки информационной безопасности ПО без исходных текстов является масштабируемым за счет модульной архитектуры и доверенным за счет использования средств виртуализации, не позволяющих нанести вред работе вычислительной системы в случае обнаружения вредоносного поведения, исследуемого ПО.

Литература

1. *Самарин Н.Н.* Виды потенциально-опасных возможностей, реализуемых вредоносным кодом // Успехи современной науки и образования (Международный научно-исследовательский журнал). 2016. Т. 4. № 9. С.199–202.

2. *Родионов А.В.* Исследование способов распространения вредоносного программного обеспечения // Фундаментальные проблемы системной безопасности: Материалы школы-семинара молодых ученых, посвященной 60-летию запуска первого в мире искусственного спутника Земли. 2017. С. 249–253.

3. *Тоторкулов М.З.С.* Защита от вредоносных программных обеспечений // Традиции и инновации в системе образования. 2017. С. 213–216.

4. *Родионов А.В.* Сравнительный анализ методов распространения вредоносного программного обеспечения // Системы управления, технические системы: устойчивость, стабилизация, пути и методы исследования: Материалы молодежной секции в рамках IV Международной научно-практической конференции. 2018. С. 189–193.

5. *Затульветер В.О., Фролов А.Е.* Исследование средств автоматизации анализа вредоносного программного обеспечения // Ломоносовские чтения на Алтае: фундаментальные проблемы науки и техники. Сборник научных статей международной конференции: электронный ресурс. Ответственный редактор: Родионов Е.Д., 2018. С. 790–796.

6. *Переберина А.А., Костюшко А.В.* Проектирование программно-аппаратного комплекса для запуска вредоносного программного обеспечения // Труды Московского физико-технического института (национального исследовательского университета). 2018. Т. 10. № 2 (38). С. 114–130.

7. *Переберина А.А., Костюшко А.В.* Разработка инструментария для динамического анализа вредоносного программного обеспечения // Труды Московского физико-технического института (национального исследовательского университета). 2018. Т. 10. № 3 (39). С. 24–44.

8. *Маршев И.И., Жуковский Е.В., Александрова Е.Б.* Использование ассемблерного кода программ для защиты средств обнаружения вредоносных программного обеспечения от состязательных атак // Методы и технические средства обеспечения безопасности информации. 2020. № 29. С. 85–86.

9. *Стасьев Д.О.* Контроль целостности компонентов виртуальных машин, созданных на базе гипервизора KVM // Безопасность информационных технологий. 2020. Т. 27. № 2. С. 118–131.

10. *Осипов А.В.* Метод решения задачи поиска оптимального распределения вычислительных ресурсов при применении механизмов виртуализации // Радиоэлектроника, электротехника и энергетика: Тезисы докладов. 2018. С. 269.

11. *Бутин А.А.* Технология защиты программного обеспечения // Информационные технологии и математическое моделирование в управлении сложными системами. 2019. № 2 (3). С. 53–63.

12. *Лапшин Д.В., Кабанцов Ю.Е., Баулин А.В., Летунова Ю.О., Тарасов В.С., Каленик Д.С.* Анализ защищенности систем виртуализации // Colloquium-journal. 2020. № 10–2 (62). С. 63–66.

13. *Мустафин Т.Р., Алехин А.И., Кравцунов Е.М., Макаев Б.О.* Безопасная среда исполнения критических приложений во встраиваемых системах на базе вычислительных средств семейства «Эльбрус» // Радиопромышленность. 2019. № 1. С. 16–22.

14. *Ахтямов Д.Р., Зегжда Д.П.* Обнаружение скомпрометированных мониторов виртуальных машин методами искусственного интеллекта // Методы и технические средства обеспечения безопасности информации. 2020. № 29. С. 28.

15. *Гордеев А.В., Горелик Д.В.* Сравнительное тестирование контейнерной и гипервизорной виртуализации // Информационно-управляющие системы. 2018. № 2 (93). С. 60–66.

16. *Колясников П.В., Силаков И.Н., Ильин Д.Ю., Гусев А.А., Никульчев Е.В.* Повышение эффективности виртуального рабо-

чего окружения распределенной разработки программ // Современные информационные технологии и ИТ-образование. 2019. Т. 15. № 1. С. 72–80.

17. *Томаев М.Х.* Средства автомитизации оптимизационныз преобразований исходных кодов программных систем // Программные продукты, системы и алгоритмы. 2018. № 3. С. 3.

18. *Мусин С.М., Дорохов Д.Г., Сутормин Д.А.* Надежность программного обеспечения авиационного оборудования летательных аппаратов государственной авиации // Научные чтения по авиации, посвященные памяти Н.Е. Жуковского. 2018. № 6. С. 413–422.

19. *Расулов М.М.* Оценка надежности программного обеспечения // Актуальные научные исследования в современном мире. 2020. № 6–2 (62). С. 112–116.

20. *Чепцов М.Н., Сребная И.Г.* Метод повышения надежности программного обеспечения в системах управления движением поездов // Сборник научных трудов Донецкого института железнодорожного транспорта. 2019. № 52. С. 27–31.

21. Пошивалов В.П., Даниев Ю.Ф. О моделях надежности программного обеспечения эргатических систем // Техническая механика. 2017. № 4. С. 89–95.

22. *Гусеница Я.Н.* Имитационно-аналитическая модель надежности программного обеспечения // Информационные системы и технологии. 2019. № 5 (115). С. 10–17.

23. *Левина Т.М., Фомина В.В., Переверзева А.И., Полянская В.И.* Оценка надежности при написании программного обеспечения применяемого в нефтяной отрасти // Нефтегазовое дело. 2018. Т. 16. № 6. С. 107–114.

SOLUTION FOR A SOURCE CODE-LESS SOFTWARE INFORMATION SECURITY ASSESSMENT

NIKOLAY N. SAMARIN

Moscow, Russia, samarin_nik@mail.ru

KEYWORDS: information security; malware; hypervisor; modular architecture.

ABSTRACT

Introduction: Digitalisation affects all sectors of human activity, resulting in the creation of a variety of software that implements business logic and technical processes in complex systems. Under these conditions, the issue of identifying malware becomes even more important. The solution to this problem is complicated by the lack of source code and the need to quickly make a decision on the presence or absence of malicious functionality. **Research Aim:** The aim of the research is to create an approach to assess the information security of software without source code. It is proposed that the approach is based on the use of a hypervisor that provides control over the operation of software with memory as a key characteristic of the presence/absence of its malicious functionality. It is proposed to calculate a software security score as a security metric. **Methods:** the solution of the set issue is based on the use of virtualization mechanism providing control over all operations over the memory realized by the software and on the use of probability theory methods to get a complex security estimate which takes into account the reliability of the software functioning and its security. **Results:** the methodology of getting a complex estimation of software functioning security is developed which takes into account the security of software functioning; network security – vulnerabilities and network ports detected by scanning; potentially insecure changes in file system and register and also potentially dangerous operations connected with the use of memory. The architecture of the software prototype that implements the proposed approach is described and its experimental testing is carried out, as a result of which only regular software samples received high security assessment. **Practical significance:** the developed system can be used for automated analysis of software operating in various complex systems. An important advantage of the software prototype is its scalability and trustworthiness ensured through the use of virtualization tools that do not allow damaging the work of a computer system in case of detection of malicious software.

REFERENCES

1. Samarin, N.N. Types of potentially dangerous opportunities implemented by malicious code. *Uspekhi sovremennoy nauki i obrazovaniya (International Scientific Research Journal).* 2016. Vol. 4. No. 9. Pp. 199–202. (In Rus)

2. Rodionov A.V. Research of methods of distribution of malicious software. *In the collection: Fundamental problems of system security.*

Materials of the school-seminar of young scientists dedicated to the 60th anniversary of the launch of the world's first artificial Earth satellite. 2017. Pp. 249–253. (In Rus)

3. Totorkulov M.Z.S. Protection from malicious software. *In the collection: Traditions and innovations in the education system*. 2017. Pp. 213–216. (In Rus)

4. Rodionov A.V. Comparative analysis of methods of distribution of malicious software. *In the collection: Control systems, technical systems: stability, stabilization, ways and methods of research. Materials of the youth section within the IV International Scientific and Practical Conference*. 2018. Pp. 189–193. (In Rus)

5. Zatulveter V.O., Frolov A.E. Research of automation tools for the analysis of malicious software. *In the collection: Lomonosov Readings in the Altai: fundamental problems of science and technology. Collection of scientific articles of the international conference: electronic resource*. 2018. Pp. 790–796. (In Rus)

6. Pereberina A.A., Kosciusko A.V. Design of a software and hardware complex for launching malicious software. *Proceedings of the Moscow Institute of Physics and Technology (National Research University)*. 2018. Vol. 10. No. 2 (38). Pp. 114–130. (In Rus)

7. Pereberina A.A., Kosciusko A.V. Development of tools for dynamic analysis of malicious software. *Proceedings of the Moscow Institute of Physics and Technology (National Research University).* 2018. Vol. 10. No. 3 (39). Pp. 24–44. (In Rus)

8. Marshev I.I., Zhukovsky E.V., Alexandrova E.B. Using the assembler code of programs to protect the means of detecting malicious software from adversarial attacks. *Methods and technical means of ensuring information security*. 2020. No. 29. Pp. 85–86. (In Rus)

9. Stasyev D.O. Monitoring the integrity of components of virtual machines created on the basis of the KVM hypervisor. *Information Technology security.* 2020. Vol. 27. No. 2. Pp. 118–131. (In Rus)

10. Osipov A.V. A method for solving the problem of finding the optimal distribution of computing resources when using virtualization mechanisms. *In the book: Radioelectronics, electrical engineering and power engineering. Abstracts of reports.* 2018. P. 269. (In Rus)

11. Butin A.A. Technology of software protection. *Information technologies and mathematical modeling in the management of complex systems*. 2019. No. 2 (3). Pp. 53–63. (In Rus)

12. Lapshin D.V., Kabantsov Yu. E., Baulin A.V., Letunova Yu.O., Tarasov V.S., Kalenik D.S. Analysis of the security of virtualization systems. *Colloquium-journal.* 2020. No. 10–2 (62). Pp. 63–66. (In Rus)

13. Mustafin T.R., Alyokhin A.I., Kravtsunov E.M., Makaev B.O. Safe execution environment for critical applications in embedded systems based on computing tools of the "Elbrus" family. *Radio industry*. 2019. No. 1. Pp. 16–22. (In Rus)

14. Akhtyamov D.R., Zegzhda D.P. Detection of compromised virtual machine monitors by artificial intelligence methods. *Methods and technical means of ensuring information security.* 2020. No. 29. P. 28.

15. Gordeev A.V., Gorelik D.V. Comparative testing of container and hypervisor virtualization. *Information and control systems.* 2018. No. 2 (93). Pp. 60–66. (In Rus)

16. Kolyasnikov P.V., Silakov I.N., Ilyin D. Yu., Gusev A.A., Nikulchev E.V. Improving the efficiency of the virtual working environment of distributed software development. *Modern information technologies and IT education.* 2019. Vol. 15. No. 1. Pp. 72–80. (In Rus)

17. Tomaev M.H. Means of automitization of optimization transformations of source codes of software systems. *Software products, systems and algorithms.* 2018. No. 3. P. 3. (In Rus)

18. Musin S.M., Dorokhov D.G., Sutormin D.A. Reliability of software for aviation equipment of state aviation aircraft. *Scientific readings on aviation, dedicated to the memory of N.E. Zhukovsky.* 2018. No. 6. Pp. 413–422. (In Rus)

19. Rasulov M.M. Evaluation of software reliability. *Actual scientific research in the modern world.* 2020. No. 6–2 (62). Pp. 112–116. (In Rus)

20. Cheptsov M.N., Srebnaya I.G. Method of improving the reliability of software in train traffic control systems. *Collection of scientific papers of the Donetsk Institute of Railway Transport.* 2019. No. 52. Pp. 27–31. (In Rus)

21. Poshivalov V.P., Daniev Yu.F. On models of reliability of software for ergatic systems. *Technical Mechanics.* 2017. No. 4. Pp.89-95. *(In Rus)*

22. Caterpillar Ya.N. Simulation and analytical model of software reliability. *Information systems and technologies.* 2019. No. 5 (115). Pp. 10–17. (In Rus)

23. Levina T.M., Fomina V.V., Pereverzeva A.I., Polyanskaya V.I. Evaluation of reliability in writing software used in the oil industry. *Oil and gas business.* 2018. Vol. 16. No. 6. Pp. 107–114. (In Rus)

INFORMATION ABOUT AUTHOR:

Samarin N.N., Head of the Research Department No. 6, Federal State Unitary Enterprise "Kvant Research Institute".

For citation: Samarin N.N. Solution for a source code-less software information security assessment. *H&ES Research*. 2021. Vol. 13. No. 2. Pp. 25–34. Doi: 10.36724/2409-5419-2021-13-2-25-34 (In Rus)

Doi: 10.36724/2409-5419-2021-13-2-35-43

НАУЧНО-ТЕХНИЧЕСКИЕ ПРЕДЛОЖЕНИЯ ПО ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОМУ ОБЕСПЕЧЕНИЮ КОМПЛЕКСОВ РАДИОМОНИТОРИНГА

СМИРНОВ
Андрей Александрович[1]

ИВАНОВ
Андрей Анатольевич[2]

ЗАИКА
Павел Валентинович[3]

КУЛИКОВ
Максим Владимирович[4]

АННОТАЦИЯ

Введение: в рамках научного направления по совершенствованию информационно-управляющих систем комплексов радиомониторинга рассмотрены вопросы формирования структуры и содержания информационных ресурсов для решения задач информационно-аналитического обеспечения. Отправной точкой послужили требования к информационно-управляющим системам, к реализации в них процессов сбора, обработки и представления информации. основой выполнения которых позиционируется нахождение рациональной структуры информационных ресурсов и разработка механизмов их взаимодействия, наполнения и обновления. **Цель исследования**: на основе анализа требований к информационно-управляющим системам комплексов радиомониторинга выработать научно-технические предложения по реализации научных разработок по формированию их информационных ресурсов, обеспечивающих выполнение указанных требований. **Результаты.** Определены целевая функция системы, показатели качества ее функционирования. Представлены результаты построения концептуальной модели единого информационного ресурса комплекса радиомониторинга как совокупности структурной, функциональной и онтологической модели. В качестве основных структурных элементов рассматриваются хранилище данных радиомониторинга, поступающих от технических средств обнаружения, пеленгования радиоэлектронных средств, база данных признаковых описаний объектов радиомониторинга, фактографический информационно-справочный ресурс, база данных с результатами обработки данных радиомониторинга, а также непосредственно алгоритмы обработки. Отражены вопросы наполнения информационных ресурсов, определения периодичности их обновления. Конкретизирован состав алгоритмов оперативной (текущей) и тематической (отложенной) обработки данных радиомониторинга, детализированы основные задачи подсистем управления и обеспечения информационно-управляющих систем. Предложена программная архитектура информационно-аналитического обеспечения комплексов радиомониторинга. Показано применение технологии интеграционной сервисной шины в качестве метасистемной основы объединения информационных ресурсов, информационных и аналитических сервисов, реализации механизмов их взаимодействия, доступа к ним. Сделаны выводы о применимости разработанных научно-теоретических положений на практике, показаны примеры построения реализующего их программного обеспечения. **Обсуждение**. Результаты исследований могут быть применены при проектировании информационных и аналитических подсистем многоуровневых систем и комплексов радиомониторинга, информационный процесс в которых предполагает многоэтапную обработку данных и их представление в реальном масштабе времени различным потребителям.

КЛЮЧЕВЫЕ СЛОВА: информационные ресурсы; радиомониторинг; обработка данных радиомониторинга; интеграционная сервисная шина; информационно-управляющая система; информационно-аналитическое обеспечение.

Сведения об авторах:

[1]к.т.н., докторант Военной академии связи им. С. М. Буденного, г. Санкт-Петербург, Россия, andrew_work@list.ru

[2]к.т.н., доцент, профессор Военной академии связи им. С. М. Буденного, г. Санкт-Петербург, Россия, a-a-iv@yandex.ru

[3]преподаватель Военной академии связи им. С. М. Буденного, г. Санкт-Петербург, Россия, pashasever@mail.ru

[4]к.т.н., докторант Военнойя академии связи им. С. М. Буденного, г. Санкт-Петербург, Россия, mr.maximus85@mail.ru

Для цитирования: *Смирнов А.А., Иванов А.А., Заика П.В., Куликов М.В.* Научно-технические предложения по информационно-аналитическому обеспечению комплексов радиомониторинга // Наукоемкие технологии в космических исследованиях Земли. 2021. Т. 13. № 2. С. 35–43. Doi: 10.36724/2409-5419-2021-13-2-35-43

Введение

Наряду с развитием методов и средств обнаружения, измерения параметров радиоизлучений важной задачей совершенствования комплексов радиомониторинга является повышение эффективности процессов сбора данных от технических средств, их обработки и представления потребителям [1]. Эффективность в данном случае оценивается системой показателей, характеризующих выполнение требований к результатам радиомониторинга — их своевременности, полноте и достоверности. В современных комплексах радиомониторинга реализация процессов сбора, обработки и представления данных возлагается на информационно-аналитическое обеспечение их информационно-управляющих систем [2, 3]. Проведенные исследования показали, что успешность их функционирования во многом определяется структурированностью, наполнением, актуальностью и интеллектуальностью информационных ресурсов, как совокупности данных, участвующих в информационном процессе комплексов, а также алгоритмов их обработки. В статье рассматриваются требования к информационно-управляющим системам комплексов радиомониторинга, научные разработки по формированию информационных ресурсов как основы информационно-аналитического обеспечения комплексов и научно-технические предложения по их реализации.

Основная часть

Принципиальные требования к информационно-управляющим системам комплексов радиомониторинга, как средствам управления и информационно-аналитического обеспечения автоматизированных систем радиомониторинга, могут быть разделены на три группы [4]: требования к результатам функционирования — представляемым потребителям данным; требования к структуре системы; требования к функционалу программного обеспечения.

В соответствии с назначением информационно-управляющих систем основными требованиями к результатам радиомониторинга являются их полнота, достоверность и своевременность. Имея в виду, что целью функционирования комплекса радиомониторинга как системы является отслеживание радиоэлектронной обстановки (РЭО) в заданном районе, интегральным показателем его эффективности является расстояние между множествами, составляющими реальную РЭО и ее отображение, полученное в ходе радиомониторинга. Представим реальную РЭО как данные об истинном местоположении и характеристиках источников радиоизлучения (ИРИ) в виде матрицы $\mathbf{M}_G = \{\mathbf{g}_i\}$, $i = 1, K$, где вектор характеристик ИРИ $\mathbf{g}_i = \{f, t, (x, y, z), U_G\}$, f — частота (пул частот) излучения (МГц), t — характеристика времени излучения (его начала, продолжительности), (x, y, z) — характеристики положения ИРИ в пространстве (широта, долгота, относительная высота), U_G — технические параметры излучения (вид модуляции, мощность сигнала и др.), K — количество ИРИ в заданном районе. Аналогично, наблюдаемая РЭО представляется матрицей $\mathbf{M}_D = \{\mathbf{d}_i\}$, $i = 1, K$, $\mathbf{d}_i = \{f, t, (x, y, z), U_D\}$. Заполнение матрицы $\mathbf{M}_D$ *осуществляется в ходе радиомониторинга.* Тогда функциональная модель системы определяется отображением $\varepsilon = \mathbf{M}_G \rightarrow \mathbf{M}_D$. Показателем эффективности является количество ошибок идентификации ИРИ, определения их местоположения, других характеристик ИРИ: $\sigma = K^{-1}\sum[\mathbf{g}_i \neq \mathbf{d}_i]$, где квадратные скобки (в нотации Айверсона) переводят значение в число по правилу [ложь] = 0, [истина] = 1. Таким образом, выражение для σ подчиняется традиционным требованиям полноты, достоверности и своевременности обработки [4]. Целью совершенствования системы радиомониторинга является минимизация этого показателя.

С точки зрения теории систем в структуре информационно-управляющей системы должны быть выделены три подсистемы [5]: подсистема основного процесса (добывание, сбор, обработка и представление данных радиомониторинга), подсистема управления и подсистема обеспечения. Организационно-техническая структура подсистемы основного процесса должна иметь в своем составе (рис. 1):

– хранилище данных радиомониторинга, поступающих от технических средств;

– фактографические ресурсы: способствующие семантической обработке справочники, словари, инструкции, карты, модели и признаки объектов, оснащаемых ими радиоэлектронных средств (РЭС);

– структурированное хранилище результатов обработки данных радиомониторинга в виде массивов данных по идентифицированным РЭС, радиосетям и объектам;

– сервис обработки данных радиомониторинга, реализующий комплекты задач оперативной и тематической обработки, обеспечивающий преобразование массива данных радиомониторинга в результаты в виде данных об объектах и РЭС, РЭО;

– сервис представления, обеспечивающий формирование отчетно-информационных документов по результатам радиомониторинга для потребителей, их визуализацию.

Функционал подсистемы основного процесса должен включать в себя комплекты задач оперативной (текущей) и тематической (отложенной) обработки данных радиомониторинга.

Комплект задач оперативной обработки должен обеспечивать выполнение следующих функций в реальном масштабе времени:

– определение местоположения РЭС и объектов по результатам пеленгования, проверка и корректировка данных местоположения;

– идентификация РЭС, контроль соответствия параметров его работы полученным лицензиям, разрешениям;

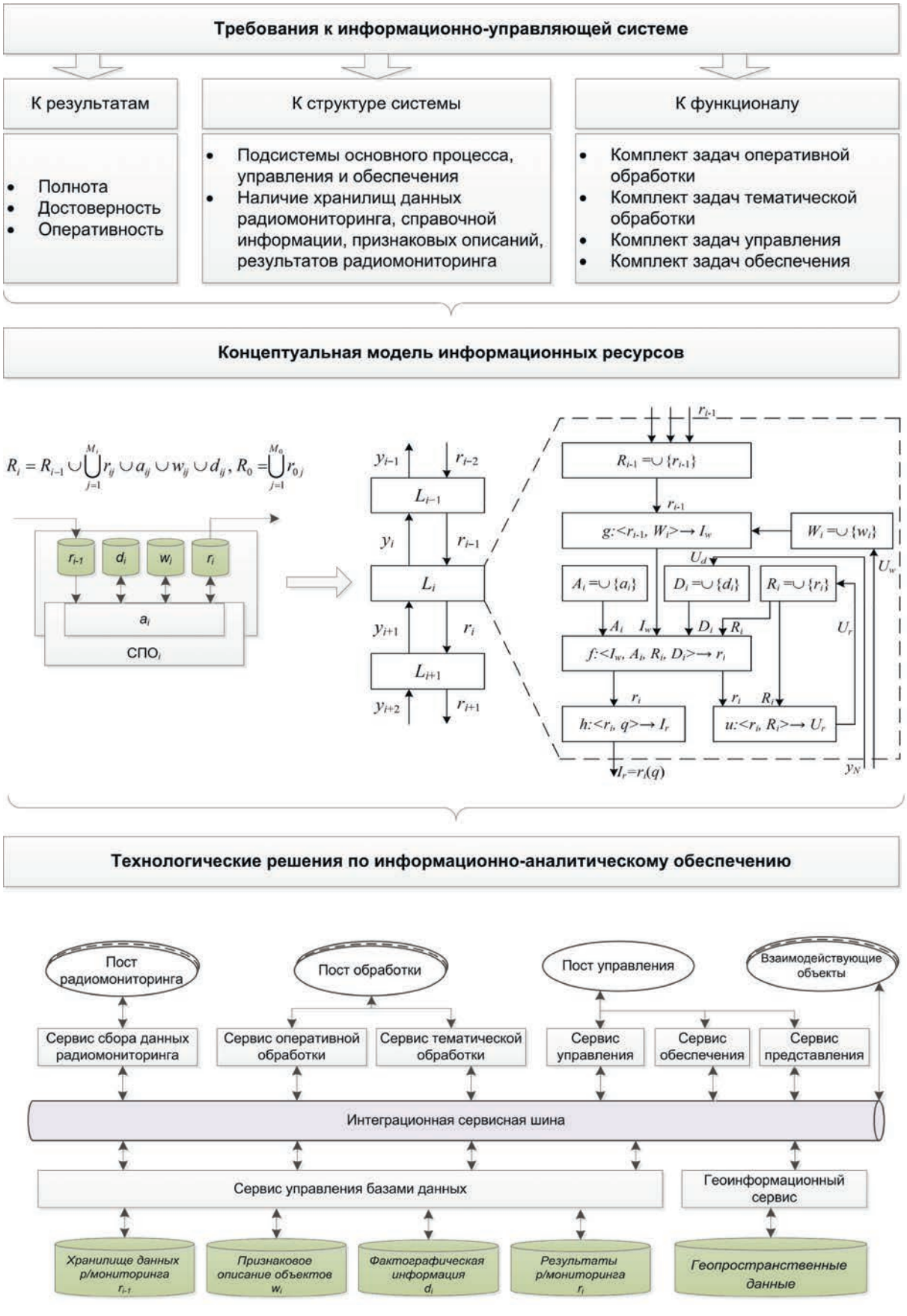

Рис. 1. Структурно-логическая схема исследования

– определение принадлежности радиоэлектронных средств к радиосетям, объектам;

– формирование и отправка докладов о функционировании радиоэлектронных средств потребителям, отображение РЭО на электронной карте района.

Комплект задач тематической обработки должен обеспечивать:

– оценку радиоэлектронной обстановки в районе радиомониторинга;

– оценку электромагнитной доступности РЭС;

– выявление новых признаков в работе РЭС, структурно-статистических характеристик РЭО, обеспечивающих ее категоризацию, отслеживание изменений (особенно в районе проведения массовых мероприятий).

Подсистема управления предназначена для выработки управляющих воздействий на средства радиомониторинга, средства сбора и обработки данных радиомониторинга.

Основными задачами, решаемыми подсистемами управления информационно-управляющих систем комплексов радиомониторинга, должны быть следующие:

– задачи управления состоянием и процессами радиомониторинга в реальном масштабе времени;

– задачи по перераспределению ресурсов технических средств на наиболее важные направления радиоконтроля, частотные диапазоны;

– задачи по учету и отображению состояния технических средств;

– задачи по организации перемещения комплексов радиомониторинга, выбора позиционных районов, маршрута выдвижения.

Подсистема обеспечения предназначена для решения задач по подготовке к применению всех видов информационно-вычислительных ресурсов и должна включать в себя:

– задачи геоинформационного обеспечения;

– задачи информационно-лингвистического обеспечения;

– задачи обеспечения безопасности информации;

– задачи долговременного хранения данных в целях их ретроспективного анализа для последующего решения задач тематической обработки.

В соответствии с указанными требованиями разработаны структурная (рис. 1, 2) и функциональная (рис. 1) модели информационных ресурсов комплексов радиомониторинга [6, 7], методика их наполнения и обновления.

При разработке структурной модели информационных ресурсов комплексов радиомониторинга было выявлено следующее:

– структура информационных ресурсов R_i, используемых операторами технических средств, информационно-управляющих систем комплексов, ситуационных центров радиомониторинга является одинаковой (рис. 2) и состоит из баз данных результатов обработки информации (r_i), эталонных признаковых описаний объектов радиомониторинга (w_i), алгоритмов, процедур обработки данных радиомониторинга (a_i), фактографических справочных информационных ресурсов (d_i) и исходных данных для обработки (r_{i-1});

– ресурсы различных уровней i отличаются наполнением (на практике к тому же и исполнением), при этом выходные данные ресурса i-го уровня являются исходными для $i+1$-го уровня (рис. 1, 2), что должно учитываться для обеспечения интероперабельности систем и комплексов радиомониторинга;

– взаимодействие элементов информационного ресурса внутри одного уровня и по иерархии подчинения обеспечивается алгоритмами, реализующими конкретные бизнес-политики обработки данных, управления, лежащие в основе соответствующих информационно-аналитических сервисов.

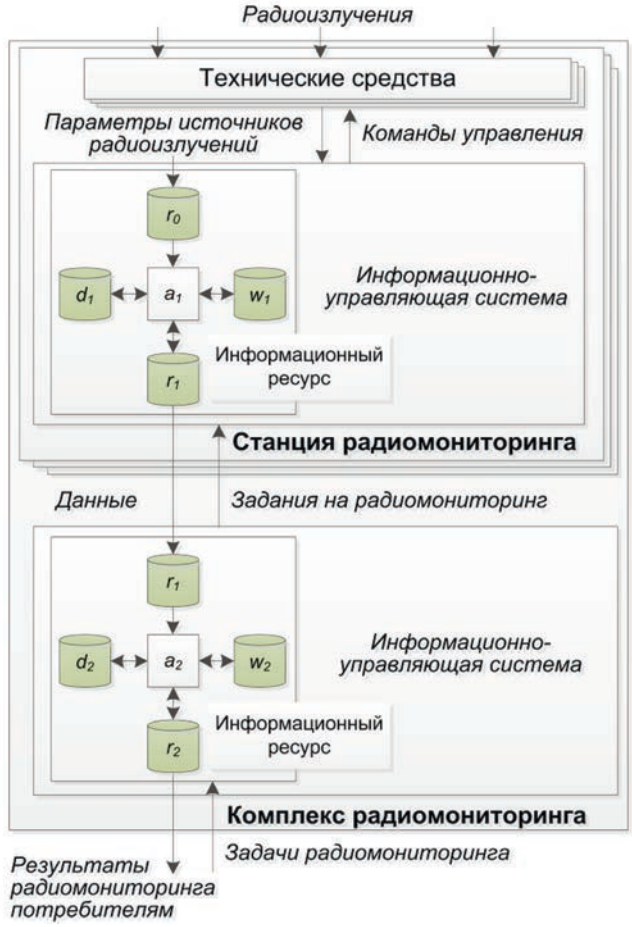

Рис. 2. Место информационных ресурсов в информационном процессе комплекса радиомониторинга

Наполнение баз данных r_{i-1} и r_i осуществляется во время функционирования комплекса в процессе сбора и обработки данных. Данные r_{i-1} и r_i относятся к переменной информации и подвержены интенсивному старению. В отличие от них базы справочных данных d_i, признаковых описаний w_i и алгоритмы, процедуры обработки a_i разрабатываются заранее и долгое время остаются актуальными.

Учет актуальности, степени доверия к используемым в обработке данным лежит в основе выполнения требования к достоверности результатов радиомониторинга. Поэтому для определения периодичности обновления ресурсов с переменной и условно-постоянной информацией была разработана методика, основанная на экспоненциальном законе старения информации $T=(-\ln P(t))/S$, где T — период обновления данных для обеспечения степени доверия к ним не хуже $P(t)$ при интенсивности их старения S. Достоверность сделанных предположений о законе старения подтверждается множеством специальных исследований других авторов, а также результатами проведенного имитационного моделирования [8].

Наиболее трудоемкой задачей, требующей научного обоснования, явилось определение состава и разработка методики наполнения фактографического информационного ресурса, обеспечивающего представление справочной информации по вопросам предметной области радиомониторинга в условиях автономного функционирования комплекса в заданном географическом районе без подключения к глобальной информационной сети. Так как справочная информация о предметной области в таких условиях очень обширна, а современные поисковые системы предназначены, главным образом, для поиска конкретных ответов на короткие текстовые запросы, определение оптимального набора коротких запросов, обеспечивающего получение точной и полной поисковой выдачи и формирование базы справочной информации, в такой ситуации является весьма нетривиальной задачей. Поэтому было предложено применить технологию разведочного (исследовательского, обзорного) информационного поиска (exploratory search) [9], целью которого является получение ответов на сложные слабо формализуемые в виде ключевых слов вопросы. Основными этапами решения этой задачи являются следующие:

– построение тематической модели коллекции предварительно выгруженных документов-кандидатов (размер коллекции может составлять сотни тысяч документов);

– формирование текстового запроса-документа (может состоять из сотен и тысяч слов, отражающих смысловое содержание предметной области);

– построение тематической модели для запроса-документа при фиксированной матрице термов тем для коллекции;

– вычисление меры близости столбцов матрицы тем документов коллекции с вектором тем запроса, ранжирование результатов и формирование коллекции релевантных документов (мера близости которых выше установленного порога);

– построение индекса классической поисковой системы для полученной в ходе разведочного поиска коллекции релевантных документов.

Дальнейшая обработка коллекции релевантных документов позволила по матрице термов тем осуществить отбор основных терминов предметной области и разработать ее онтологическую модель (рис. 3), используемую для

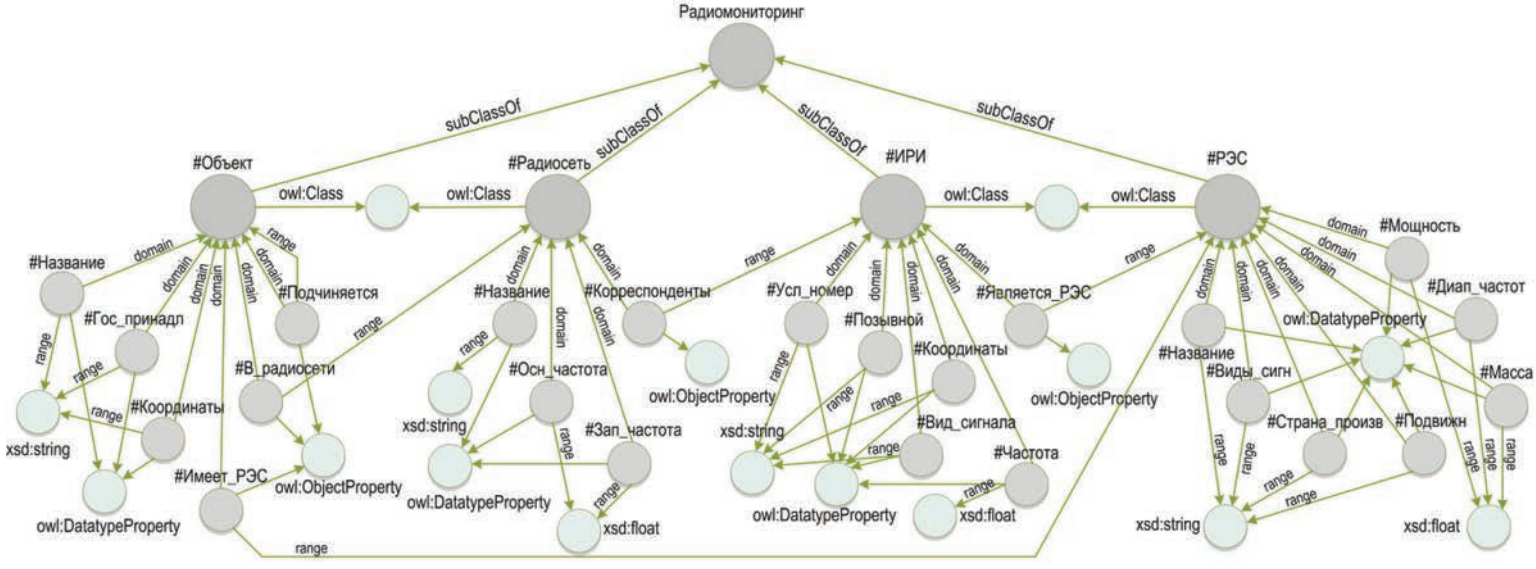

Рис. 3. Фрагмент онтологии предметной области радиомониторинга

поиска значений при заполнении признаковых описаний радиоэлектронных средств и объектов радиомониторинга.

Научный задел в решении указанной задачи обеспечивается активным развитием теоретических и прикладных вопросов тематического моделирования [9–11], информационного поиска [12–13], автоматической обработки текстов [14–17].

В целях практической реализации научных разработок по информационно-аналитическому обеспечению комплексов радиомониторинга предлагается:

1. Ввиду высокой интенсивности поступления данных от технических средств радиомониторинга хранилище данных радиомониторинга организовать на основе NoSQL базы данных. Это позволит при систематизации данных эффективно и с минимальными затратами времени на процессы обращения к памяти применить технологии обработки больших данных, например, Map Reduce [18];

2. Для формирования фактографического информационного ресурса методом тематического информационного поиска применять открытую программную библиотеку *artm*, разрабатываемую и поддерживаемую в рамках проекта *BigARTM*. Результаты тематического моделирования коллекции релевантных документов использовать для автоматического тегирования документов при формировании индекса поисковой системы в локальном фактографическом информационном ресурсе. Поиск по документам ресурса с учетом морфологии и геопривязки на основе запросов из строки поиска обеспечивается поисковой системой, например, Elastic Search [19], Sorl и др.;

3. Представление и хранение признакового описания радиоэлектронных средств и объектов радиомониторинга, результатов радиомониторинга реализовать с использованием объектно-реляционной системы управления базами данных (например, PostgreSQL), дополненной средствами хранения и обработки геопространственных данных (например, PostGIS). В качестве инфологической схемы базы данных результатов радиомониторинга использовать построенную онтологию предметной области;

4. Реализацию функционирования информационных сервисов, доступ к различным информационным ресурсам осуществлять с использованием одной из сервис-ориентированных технологий, например, интеграционной сервисной шины (Enterprise Service Bus, ESB). С точки зрения модели взаимодействия открытых систем она функционирует на прикладном уровне. При этом технология интегрированной сервисной шины предоставляет приложениям возможность функционировать через единую точку доступа к ресурсам (рис. 1), находящимся в информационно-коммуникационной сети [20]. Основу реализации функций подсистем сбора данных от различных средств в технологии интеграционной сервисной шины осуществляют агент сообщений и шлюзы (программные модули, обеспечивающие взаимодействие с приложениями в том формате, который для них приемлем). Они представляют информацию от интегрируемых элементов (средств радиомониторинга, других подсистем комплексов радиомониторинга) в унифицированном формате интеграционной сервисной шины, воспринимаемом агентом сообщений. Пример интерфейса программного обеспечения комплекса радиомониторинга, построенного с применением технологии интеграционной сервисной шины представлен на рис. 4.

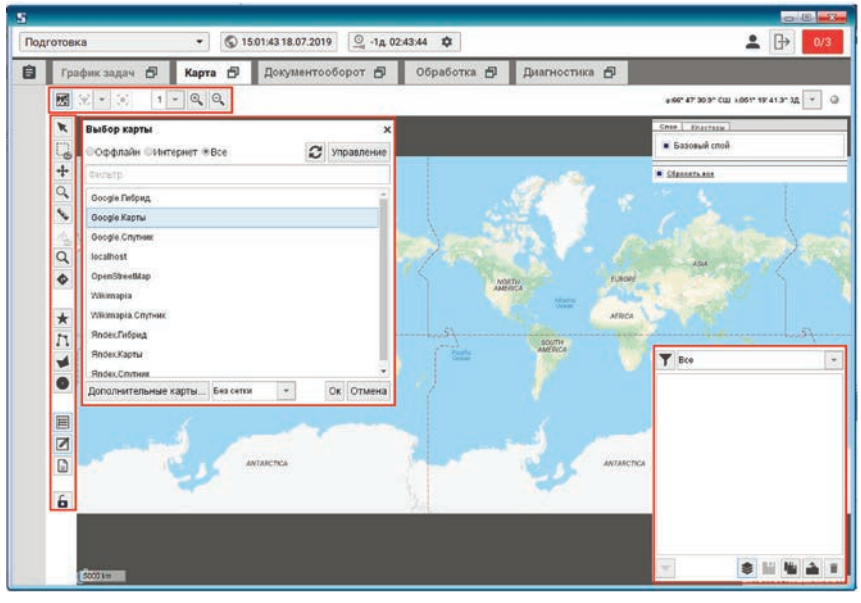

Рис. 4. Пример интерфейса программного обеспечения комплекса радиомониторинга, объединяющего сервисы обработки, управления и обеспечения по технологии ESB

Заключение

Таким образом, представленные в статье научно-технические предложения по информационно-аналитическому обеспечению комплексов радиомониторинга показывают реализуемость разработанных моделей информационных ресурсов, методик наполнения и обновления. Примененные программные решения не являются единственно возможными и были выбраны в качестве наиболее доступных программных средств с открытым исходным кодом среди прочих средств, удовлетворяющих рассмотренным в статье требованиям.

Литература

1. *Липатников В.А., Царик О.В.* Методы радиоконтроля. Теория и практика: Монография. СПб.: ГНИИ «НАЦРАЗВИТИЕ», 2018. 608 с.

2. *Белов С.Г., Белуга Г.И., Верба В.С.* Информационно-измерительные и управляющие радиоэлектронные системы и комплексы: Монография / Под ред. В.С. Вербы. М.: Радиотехника, 2020. 490 с.

3. *Xu W., Xu J., Li J., Liu W., Gong S., Zeng K.* Robust Spectrum Monitoring in Cognitive Radio Networks With Uncertain Traffic Information // IEEE Access. 2018. Vol. 6. Pp. 34696–34706.

4. *Иванов А.А., Кудрявцев А.М., Смирнов А.А.* Концептуальные проблемы информационно-аналитической работы в современном военном противостоянии // Военная мысль. 2020. № 9. С. 79–85.

5. *Щекочихин О.В.* Объектно-процессная модель данных в управляющих информационных системах // Научно-технический вестник информационных технологий, механики и оптики. 2017. Т. 17. № 2. С. 318–323.

6. *Заика П.В., Смирнов А.А., Галов С.Ю.* Формирование информационного ресурса в цикле управления радиомониторингом // Известия Тульского государственного университета. Технические науки. 2019. № 7. С. 223–229.

7. *Смирнов А.А., Кудрявцев А.М., Заика П.В.* Модель информационного ресурса автоматизированного комплекса радиомониторинга // Электросвязь. 2020. № 10. С. 42–48.

8. *Смирнов А.А., Кудрявцев А.М., Галов С.Ю.* Имитационное моделирование радиоэлектронной обстановки в районе действий воинских формирований // Электросвязь. 2020. № 10. С. 36–41.

9. *Ianina A., Vorontsov K.* Hierarchical Interpretable Topical Embeddings for Exploratory Search and Real-Time Document Tracking // International Journal of Embedded and Real-Time Communication Systems. 2020. Vol. 11. Issue 4. Pp. 134–152.

10. *Apishev M., Vorontsov K.* Learning topic models with arbitrary loss // Conference of Open Innovations Association, FRUCT. 2020. № 26. Pp. 30–37.

11. *Frei O., Apishev M.* Parallel non-blocking deterministic algorithm for online topic modeling // Proceedings of the AIST Conference (Analysis of Images, Social networks and Texts). 2016. Vol. 661. Pp. 132–144.

12. *Roy D., Ganguly D.S., Bhatia S., Bedathur S., Mitra M.* Using word embeddings for information retrieval: How collection and term normalization choices affect performance // Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM'18), New York: ACM. 2018. Pp. 1835–1838.

13. *Vuong T., Jacucci G., Ruotsalo T.* Proactive information retrieval via screen surveillance // Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2017. Pp. 1313–1316.

14. *Лукашевич Н.В., Добров Б.В., Павлов А.М., Штернов С.В.* Онтологические ресурсы и информационно-аналитическая система в предметной области «Безопасность» // Онтология проектирования. 2018. Т. 8. № 1(27). С. 74–95.

15. *Wang W., Kennedy R. , Lazer D. , Ramakrishnan N.* Growing pains for global monitoring of societal events // Science. 2016. Vol. 353(6307). Pp. 1502–1503.

16. *Bansal A.* Advanced Natural Language Processing with TensorFlow 2. Birmingham: Packt Publishing, 2021. P. 360.

17. *Сидорова Е.А.* Подход к моделированию процесса извлечения информации из текста на основе онтологии // Онтология проектирования. 2018. Т. 8. № 1(27). С. 134–151.

18. *Bolla S., Anandan R.* Contemporary review on technologies and methods for converting unstructured data to structured data // International Journal of Engineering and Technology (UAE). 2018. Vol. 7. No. 3. Special Iss. 27. Pp. 527–530.

19. *Shaik Subhani.* A Conceptual Review of Elastic Search — Survey Paper. International Journal for Research in Applied Science and Engineering Technology. 2017. Vol. V. Pp. 1703–1710.

20. *Gosewehr F., Wermann J., Borsych W., Colombo A.W.* Specification and design of an industrial manufacturing middleware // Proceedings of the IEEE15th International Conference on Industrial Informatics, INDIN-2017. 2017. Pp. 1160–1166.

SCIENTIFIC AND TECHNICAL PROPOSALS FOR THE RADIOMONITORING COMPLEXES INFORMATION AND ANALYTICAL SUPPORT

ANDREI A. SMIRNOV,

St. Petersburg, Russia, andrew_work@list.ru

ANDREI A. IVANOV,

St. Petersburg, Russia, a-a-iv@yandex.ru

PAVEL V. ZAIKA,

St. Petersburg, Russia, pashasever@mail.ru

MAKSIM V. KULIKOV,

St. Petersburg, Russia, mr.maximus85@mail.ru

KEYWORDS: information resources; radiomonitoring; radiomonitoring data processing; enterprise service bus; information management systems; information and analytical support.

ABSTRACT

Introduction: within the framework of the scientific direction on improving the information and control systems of radio monitoring complexes, the issues of forming the structure and content of information resources that ensure the solution of the problems of collecting, processing and presenting radio monitoring data are considered. The starting point was the requirements for information management systems, the basis for the implementation of which is to find a rational structure of information resources and develop mechanisms for their interaction, filling and updating. **Purpose:** on the basis of an analysis of the requirements for radio monitoring complexes information and control systems, develop scientific and technical proposals for the implementation of scientific developments on the formation of their information resources that ensure the fulfillment of these requirements. **Results:** the target function of the system, indicators of the quality of its functioning have been determined. The article presents the results of constructing a conceptual model of a single information resource of a radio monitoring complex as a set of structural, functional and ontological models. The main structural elements are the storage of radio monitoring data coming from technical means of detection, direction finding of radio electronic means, a database of indicative descriptions of radio monitoring objects, a factual information resource, a database with the results of processing radio monitoring data, as well as directly processing algorithms. The issues of filling information resources, determining the frequency of their updating are reflected. The composition of the algorithms for operational (current) and thematic (deferred) processing of radio monitoring data is concretized, the main tasks of the control subsystems and the provision of information and control systems are detailed. The software architecture of information support of radio monitoring complexes is proposed. The application of the integrated service bus technology as a metasystem basis of information resources is shown. Conclusions are made about the applicability of the developed scientific and theoretical provisions in practice, examples of the construction of software that implement them are shown. **Discussion:** the research results can be used in the formation of information resources of multi-level systems and radio monitoring complexes, the information process in which involves multi-stage data processing and their presentation in real time to various consumers.

REFERENCES

1. Lipatnikov V.A., Tsarik O.V. *Metody radiokontrolja. Teorija i praktika: Monografija* [Radio monitoring methods. Theory and Practice: Monograph]. St-Petersburg: GNII "NACRAZVITIE", 2018. 608 p. (In Rus)

2. Belov S.G., Beluga G.I., Verba V.S. *Informatsionno-izmeritelnyye i upravlyayushchiye radioelektronnyye sistemy i kompleksy. Monografiya* [Information-measuring and control radio-electronic systems and complexes. Monograph]. Edited by V.S. Verba]. Moscow: Radiotechnika, 2020. 490 p. (In Rus)

3. Xu W., Xu J., Li J., Liu W., Gong S., Zeng K. Robust Spectrum Monitoring in Cognitive Radio Networks With Uncertain Traffic Information. IEEE Access. 2018. Vol. 6. Pp. 34696–34706.

4. Ivanov A.A., Kudryavtsev A.M., Smirnov A.A. Information and analytical work conceptual problems in contemporary military confrontation. Voyennaya mysl [Military Thought]. 2020. No. 9. Pp. 79–85. (In Rus)

5. Schekochikhin O.V. Object-process data model in management information systems. *Nauchno-tehnicheskij vestnik informacionnyh tehnologij, mehaniki i optiki* [Scientific and Technical Journal of Information Technologies, Mechanics and Optic]. 2017. T. 17. No 2. Pp. 318–323. (In Rus)

6. Zaika P.V., Smirnov A.A., Galov S.YU. Information resource organization in radiomonitoring management cycle. *Izvestiya Tulskogo gosudarstvennogo universiteta. Tekhnicheskiye nauki* [News of Tula State University. Technical science]. 2019. No. 7. Pp. 223–229. (In Rus)
7. Smirnov A.A., Kudryavtsev A.M., Zaika P.V. Model of informational resource of automated radiomonitoring complex. *Elektrosvyaz* [Telecommunications and Radio Engineering]. 2020. No. 10. Pp. 42–48. (In Rus)
8. Smirnov A.A., Kudryavtsev A.M., Galov S.IU. Simulation modeling of radio-electronic environment in area of military formations activities. *Elektrosvyaz* [Telecommunications and Radio Engineering]. 2020. No. 10. Pp. 36–41. (In Rus)
9. Ianina A., Vorontsov K. Hierarchical Interpretable Topical Embeddings for Exploratory Search and Real-Time Document Tracking. *International Journal of Embedded and Real-Time Communication Systems*. 2020. Vol. 11. Issue 4. Pp. 134–152.
10. Apishev M., Vorontsov K. Learning topic models with arbitrary loss. *Conference of Open Innovations Association, FRUCT*. 2020. No. 26. Pp. 30–37.
11. Frei, O., Apishev, M. Parallel non-blocking deterministic algorithm for online topic modeling. *Proceedings of the AIST Conference* (*Analysis of Images, Social networks and Texts*). 2016. Vol. 661. Pp. 132–144.
12. Roy, D., Ganguly, D.S., Bhatia, S., Bedathur, S., Mitra, M. Using word embeddings for information retrieval: How collection and term normalization choices affect performance. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM'18)*. New York: ACM. 2018. Pp. 1835–1838.
13. Vuong T., Jacucci G., Ruotsalo T. Proactive information retrieval via screen surveillance. *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2017. Pp. 1313–1316.
14. Lukashevich N.V., Dobrov B.V., Pavlov A.M., Shternov S.V. Ontological resources and information-analytical system in security domain. *Ontologiya proektirovaniya* [Ontology of designing]. 2018. Vol. 8. No. 1(27). Pp. 74–95. (In Rus)
15. Wang W., Kennedy R., Lazer D., Ramakrishnan N. Growing pains for global monitoring of societal events. *Science*. 2016. Vol. 353(6307). Pp. 1502–1503.
16. Bansal A. Advanced Natural Language Processing with TensorFlow 2. Birmingham: Packt Publishing, 2021. 360 p.
17. Sidorova E.A. An approach to modeling the process of extracting information from text based on ontology. *Ontologiya proektirovaniya* [Ontology of designing]. 2018. Vol. 8. No. 1(27). Pp. 134–151. (In Rus)
18. Bolla, S., Anandan, R. Contemporary review on technologies and methods for converting unstructured data to structured data. International Journal of Engineering and Technology (UAE). 2018. Vol. 7. No. 3. Special Issue 27. Pp. 527–530.
19. Shaik, Subhani. A Conceptual Review of Elastic Search – Survey Paper. International Journal for Research in Applied Science and Engineering Technology. 2017. Vol. V. Pp. 1703–1710.
20. Gosewehr F., Wermann J., Borsych W., Colombo A.W. Specification and design of an industrial manufacturing middleware. Proceedings of the IEEE15th International Conference on Industrial Informatics, INDIN-2017. 2017. Pp. 1160–1166.

INFORMATION ABOUT AUTHORS:

Smirnov A.A., PhD, Doctoral Candidate of the Military Telecommunications Academy named after S. M. Budennyi;
Ivanov A.A., PhD, Docent, Professor at the Department of the Military Telecommunications Academy named after S. M. Budennyi;
Zaika P.V., Lecturer of the Military Telecommunications Academy named after S. M. Budennyi;
Kulikov M.V., PhD, Doctoral Candidate of the Military Telecommunications Academy named after S. M. Budennyi.

For citation: Smirnov A.A., Ivanov A.A., Zaika P.V., Kulikov M.V. Scientific and technical proposals for the radiomonitoring complexes information and analytical support. *H&ES Research*. 2021. Vol. 13. No. 2. Pp. 35–43. Doi: 10.36724/2409-5419-2021-13-2-35-43 (In Rus)

Doi: 10.36724/2409-5419-2021-13-2-44-51

РАЗРАБОТКА ЭКОНОМИЧЕСКИХ МОДЕЛЕЙ РАЗВИТИЯ ТЕХНИЧЕСКИХ СИСТЕМ ПО ЭТАПАМ ЖИЗНЕННОГО ЦИКЛА

МИКИТЕНКО
Игорь Иванович

АННОТАЦИЯ

Введение: при прогнозировании и одновременном сопровождении планового развития совокупности больших технических систем требуется разработка комплексного механизма учета состояния и затрат ресурсов. Известные для этих целей подходы характеризуются высокой сложностью реализации и разрозненностью по этапам цикла систем. **Цель исследования:** разработать комплекс взаимосвязанных экономических моделей для анализа и развития больших технических систем. **Результаты:** на базе научных направлений экономического анализа рассмотрены аспекты разработки экономических моделей и на их основе проведено моделирование сложных составных технических систем по этапам их жизненного цикла в интересах последующего формирования и принятия управленческих решений как по выбору и реализации таких систем, так и по сопровождению их развития. Совокупность экономических моделей объединена в обобщенную модель развития систем в интересах комплексного оценивания текущих и предстоящих затрат, дальнейшего прогнозирования и, на основе результатов имитационного моделирования и расчетов,– принятия эффективных решений. Так как особенностью производства больших технических систем является их серийный характер продукции, то в комплексной модели проводится оценивание затрат как по отдельным системам и по этапам их жизненного цикла, так и для всего комплекса систем в целом. Предсказано моделирование различных режимов финансирования развития одной системы или их комплекса, учитываются особенности для каждого этапа развития. Получена возможность оценивания структурной перестройки взаимосвязи предприятий участвующих в процессе производства и изменения возможной очередности технологических процессов производства образцов. **Практическая значимость:** предлагаемая модель (модели) при серийном или штучном производстве изделий рекомендуется в конкретных наукоемких направлениях машиностроения и применима, например, в авиационной, космической, судостроительной, горной и др. отраслях промышленности. Особенно актуально ее применение при проектировании, разработке и производстве изделий по государственным заказам. **Обсуждение:** представленное решение предлагается оформить программным комплексом и создать автоматизированную систему управления процессами разработки, производства, эксплуатации и утилизации больших технических систем с функциями системы поддержки принятия решений.

Сведения об авторе:
к.т.н., старший научный сотрудник, старший преподаватель Национального исследовательского технологического университета «Московский институт стали и сплавов», г. Москва, Россия, iimiki@bk.ru

КЛЮЧЕВЫЕ СЛОВА: экономическая модель; большая техническая система; этапы жизненного цикла; оценка затрат на систему; имитационное моделирование развития системы; серийное производство продукции; эффективные управленческие решения.

Для цитирования: Микитенко И.И. Разработка экономических моделей развития технических систем по этапам жизненного цикла // Наукоемкие технологии в космических исследованиях Земли. 2021. Т. 13. № 2. С. 44–51. Doi: 10.36724/2409-5419-2021-13-2-44-51

Введение

В условиях спада мировой экономики из-за пандемии коронавируса, а с другой стороны, необходимости дальнейшего совершенствования больших технических систем (БТС) с рациональными затратами [1,6], особую значимость приобретает решение задачи экономической оценки развития возможных проектов [2–5], вплоть до учета поштучного производства готовых образцов и изделий. Данная задача выдвигается при прогнозировании (планировании) разработки и серийного производства БТС в интересах снижения затрат как на отдельных этапах их жизненного цикла, так и всего рассматриваемого процесса в целом [3,7].

Разработка экономических моделей развития БТС объединяет и опирается на четыре научных направления экономического анализа [1–3]:

1. На экономическую статистику. В рамках этого направления в интересах разработки экономических моделей вырабатываются методы сбора, обработки, хранения больших объемов разнородной информации о технике и подсистемах, их разработках, производстве, эксплуатации и утилизации.

2. На техническое прогнозирование, которое изучает принципы и методы научно-обоснованного предсказания перспектив развития систем, возможные направления совершенствования систем и техники, появление принципиально новых видов систем, форм и способов их применения.

3. Непосредственно на экономическое моделирование, представляющее основной научный метод исследований экономической деятельности, методологической концепцией которого является имитационное исследование, т.е. исследование, основанное на методах имитационного моделирования [14,18,19]. Для разработки адекватных экономических моделей развития БТС необходима разработка комплексных имитационных систем, позволяющих в приемлемые сроки формировать модели всех этапов жизненного цикла. Комплекс подобных моделей должен решать задачи оптимизации распределений ассигнований по различным этапам жизненного цикла элементов систем, а на каждом этапе — по предприятиям, входящих в объединение производителей комплектующих и элементов систем.

4. На теорию выработки экономических решений, имеющую своим предметом принципы и методы формирования решающих правил, т.е. критериев, на основе которых вырабатываются варианты готовых решений, а также различного рода системы поддержки принятия решений [5,10,14,19,20] в экономической области развития БТС. Разрабатываются и совершенствуются методы выработки рациональных решений по векторным показателям. При этом показатели могут быть либо строго упорядочены по важности, либо равноценны, либо (в более общем случае) упорядочены по группам равноценности. Критерии такой степени обобщения позволяют формировать широкий круг задач выработки решений при разработке экономических моделей развития БТС.

Перечисленные направления составляют теоретический базис экономического анализа развития БТС и в значительной мере повышают практическую ценность разрабатываемых на их основе экономических моделей, получаемых при исследованиях результатов и выводов. Основная задача реализации указанных направлений при проведении исследований — снижение степени неопределенности основных экономических факторов и повышение точности расчетов, на основе которых принимаются ответственные и эффективные управленческие решения в области технической политики развития систем [3,22].

Формирование обобщенной экономической модели развития БТС

Значительное число элементов (разнотипность комплектующего оборудования, техники), составляющих БТС, множество методов формирования и определения критериев для выработки рациональных решений развития анализируемых систем [1] и ряд других особенностей требуют разработки множества моделей с последующим их объединением в комплексную имитационную систему [6,8]. Более того, в дальнейшем, на базе реализации данного подхода возможно создание автоматизированной системы управления (АСУ) разработкой, производством, эксплуатацией и утилизацией БТС.

Однако, учитывая специфическое предназначение элементов БТС и возможность прогнозирования их развития по этапам жизненного цикла [8], предлагается формирование обобщенной экономической модели развития систем, представленной на рис. 1. Такой подход, не нарушая общности, при моделировании развития каждого отдельного элемента по этапам жизненного цикла системы позволяет учесть специфические особенности: от предназначения элемента и выдвигаемых критериев для оценки его функционирования до возможной рациональной схемы взаимодействия предприятий при производстве элемента, — и, вместе с тем, определять и рассчитывать значения частных и интегральных показателей эффективности функционирования и развития БТС.

Элемент системы (один комплекс оборудования, одно изделие, образец) при своем развитии проходит различные этапы жизненного цикла (ЖЦ) [7,22]: от этапа научно-исследовательских и опытно-конструкторских работ (НИОКР) до этапа снятия данного типа элемента системы с эксплуатации (СЭ) и утилизации. На каждом этапе ЖЦ элемента системы расходуются потребные денежные ресурсы. Этап ЖЦ элемента состоит из ряда временных интервалов, соизмеримых с годами, месяцами, декадами и т.д. Каждый отдельный элемент определенного типа может находиться на своем временном интервале. Выпуск элемен-

тов системы может осуществляться штучно или серийно. Тогда общие затраты на развитие системы в целом и для каждого момента времени t на протяжении всего рассматриваемого временного периода T будут складываться из затрат по отдельным элементам, каждый из которых, в свою очередь, будет находиться на своем этапе жизненного цикла. Если элемент системы прошел этап опытно-конструкторских работ (ОКР), подготовку и производство (ПП), то затраты для этих этапов на каждом такте времени t не учитываются. Необходимы ассигнования только для этапа эксплуатации элементов системы и/или учет накапливаемых затрат на снятие элементов с эксплуатации (СЭ).

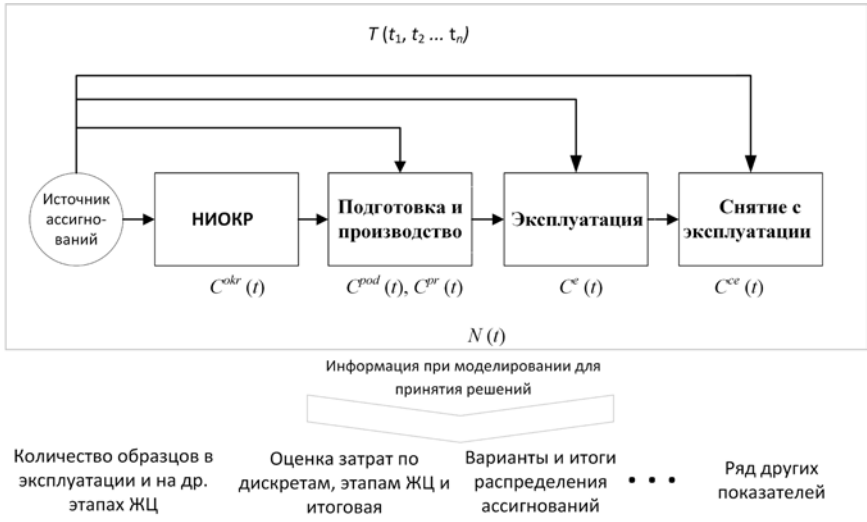

Рис. 1. Обобщенная экономическая модель развития БТС

Соответственно, затраты этапа ОКР C_{et}^{okr}, этапа подготовки к производству C_{et}^{pod}, затраты этапа производства C_{et}^{pr}, этапа эксплуатации C_{et}^{e}, затраты этапа снятия с эксплуатации C_{et}^{ce} на развитие системы по одному типу элемента для каждого этапа ЖЦ вычисляются с помощью выражений, приведенных в таблице (табл. 1):

где $C^{okr}(t)$, $C^{pod}(t)$, $C^{pr}(t)$, $C^{e}(t)$, $C^{ce}(t)$ — затраты на элемент на соответствующем этапе ЖЦ в дискретный текущий момент времени t;

$t_{0\ okr}$, $t_{0\ pod}$, $t_{0\ pr}$, $t_{0\ e}$, $t_{0\ ce}$ — начальный дискретный момент времени для соответствующего этапа ЖЦ;

T_{okr}, T_{pod}, T_{pr}, T_{e}, T_{ce} — время окончания соответствующего этапа ЖЦ;

N_{pr}, N_{e}, N_{ce} — количество элементов системы, прошедших соответствующий этап ЖЦ (количество произведенных элементов системы; количество элементов системы, находящихся в эксплуатации; количество элементов системы, снимаемых с эксплуатации);

K_{pr}^{cz}, K_{e}^{cz}, K_{ce}^{cz} — коэффициенты снижения затрат на соответствующем этапе ЖЦ ($K^{cz} \leq 1$), характеризующие особенности серийного производства, эксплуатации и снятия элементов с эксплуатации, а также другие особенности, понижающие затраты на однотипную группу (серию) элементов системы. Если серийное производство отсутствует и элементы системы выпускаются (участвуют в обработке, эксплуатации) штучно, то $K^{cz} = 1$.

Тогда общий рассматриваемый период времени T развития системы определяется выражением (1):

$$T = T_{okr} + T_{pod} + T_{pr} + T_{e} + T_{ce}. \quad (1)$$

Таблица 1

Выражения для расчета затрат на развитие системы по одному типу элемента

Этап ОКР	**Этап подготовки к производству**	**Этап производства**
$C_{et}^{okr} = \sum_{t=t0okr}^{t=Tokr} C^{okr}(t)$	$C_{et}^{pod} = \sum_{t=t0pod}^{t=Tpod} C^{pod}(t)$	$C_{et}^{pr} = N_{pr} K_{pr}^{cz} \sum_{t=t0pr}^{t=Tpr} C^{pr}(t)$
Этап эксплуатации	**Этап снятия с эксплуатации**	
$C_{et}^{e} = N_{e} K_{e}^{cz} \sum_{t=t0e}^{t=Te} C^{e}(t)$	$C_{et}^{ce} = N_{ce} K_{ce}^{cz} \sum_{t=t0e}^{t=Tce} C^{ce}(t)$	

Значения переменных $C^{okr}(t)$, $C^{pod}(t)$, $C^{pr}(t)$, $C^{e}(t)$, $C^{ce}(t)$, представленных в выражениях табл. 1, получают в результате имитационного моделирования развития систем по тактам модельного времени *t*. При этом моделируются различные режимы финансирования развития проектов, учитываются особенности и уточняются получаемые значения для каждого этапа ЖЦ рассматриваемых элементов систем. Так с помощью датчиков случайных чисел (ДСЧ), настроенных на определенные типы распределений, при моделировании могут использоваться следующие случайные величины и функции [23]:

– функции зависимости полной стоимости процесса ОКР и ПП от их продолжительности;

– случайные величины процесса производства элементов систем — продолжительность изготовления первого образца; показатель, характеризующий уменьшение продолжительности изготовления изделия при увеличении числа изделий в партии; показатель, характеризующий снижение стоимости изделия в зависимости от номера изделия в партии; стоимость первого образца в партии элементов системы;

– случайные величины процесса эксплуатации — продолжительность эксплуатации элементов систем в зависимости от их типа, условий эксплуатации и других факторов;

– случайные величины процесса снятия элементов системы с эксплуатации, характерные для рассматриваемого периода.

В результате неоднократного прогона имитационной модели для каждого интервала периода *T* получают как осредненные значения затрат для элементов системы каждого типа по этапам их ЖЦ, так и общие затраты всех систем в целом. Таким образом, изменяя пошагово модельное время, соответствующее каждому *t*-тому интервалу периода *T* ($t = 1, T$) и фиксируя состояние моделируемого варианта проекта для отдельных элементов системы и этапов их ЖЦ, исходя из выделенных ассигнований и некоторых других ограничений, представляется возможным проводить оценивание затрат на развитие рассматриваемых элементов систем.

Суммарные затраты $C_{sum}(t)$ на каждом такте времени *t* развития элементов систем определяются по зависимости (2):

$$C_{sum}(t) = C^{okr}(t) + C^{pod}(t) + C^{pr}(t) + C^{e}(t) + C^{ce}(t). \qquad (2)$$

Общие затраты C_{sum} за рассматриваемый период времени *T* для системы определяются как сумма затрат по соответствующим этапам C_{et} ЖЦ (3):

$$C_{sum} = C_{et}^{\ okr} + C_{et}^{\ pod} + C_{et}^{\ pr} + C_{et}^{\ e} + C_{et}^{\ ce} \qquad (3)$$

или с помощью выражения (4):

$$C_{sum} = \sum_{t=t0}^{T} (C^{okr}(t) + C^{pod}(t) + C^{pr}(t) + C^{e}(t) + C^{ce}.(t)) \quad (4)$$

Если в системе одновременно находится в обороте (в производстве, в эксплуатации) сразу несколько *j* типов элементов системы из множества *J* ($j \in J$), то затраты на систему в целом C^{sist}_{sum} слагаются из затрат C^{j}_{sum} по всем *j* типам элементов системы (5):

$$C^{sist}{}_{sum} = \sum_{j=1}^{J} C^{j}_{sum} \qquad . (5)$$

Количество элементов системы $N(t)$, находящихся в эксплуатации к моменту времени *t* вычисляется с помощью уравнения баланса (6):

$$N(t) = N_{e}(t–1) + N_{pr}(t) – N_{ce}(t), \qquad (6)$$

учитывающего зависимость числа элементов $N_{e}(t–1)$, находящихся в эксплуатации на предыдущем интервале времени, числа $N_{pr}(t)$ произведенных (выпущенных) на текущем интервале времени и числа элементов $N_{ce}(t)$ снятых с эксплуатации на этом интервале.

Узким местом в предлагаемых экономических моделях является наработка совокупности регулируемых параметров для моделирования, подбора типов распределений для ДСЧ, выбора случайных величин и функций, характерных и значимых для процессов ОКР, ПП, эксплуатации и СЭ типов элементов системы. Важную роль при этом играет накопленный опыт сотрудников и организаций, задействованных на соответствующих этапах ЖЦ систем, созданные и поддерживаемые ими базы данных, статистическая информация о тех самых регулируемых параметрах в соответствующей области промышленности. В данном вопросе возможно также опираться на накопленную информацию по разработке, производству и эксплуатации изделий смежными организациями или от других отраслей промышленности, адаптируя информацию по параметрам для своих конкретных систем, использовать данные для этапов ЖЦ от специализированных по теме программных продуктов, например, Anylogic, Plant Simulation, CrystalInfo и др. [23].

Этапы ЖЦ подготовки к производству и производства элементов системы связаны с кооперацией заводов-изготовителей комплектующих и готовых элементов [8]. Поэтому следует учитывать, что средства и временные затраты на изготовление комплектующих на данных этапах ЖЦ распределяются между заводами (предприятиями), а время выпуска готового единичного образца зависит от технологической схемы производства и взаимодействия заводов-изготовителей, вариант которой может быть представлен рис. 2.

Комплектующие для выпуска готового элемента системы поступают на конечный пункт сборки (предприятие)

по различным схемам. Каждый последующий модуль или единица элемента системы может быть собран лишь при условии полной его комплектации изделиями, поступающими от предыдущих заводов в общей технологической схеме. Назовем данную схему, относящуюся лишь к одному типу элемента системы, технологической линией.

В конкретной наукоемкой отрасли промышленности (машиностроения) возможно одновременное существование нескольких технологических линий, определяемых количеством разрабатываемых и эксплуатируемых типов (видов) элементов системы [8,9]. Подобные схемы и приведенные зависимости характерны и применимы в авиационной, космической, судостроительной, горной и др. отраслях промышленности.

Представленная обобщенная экономическая модель развития системы в конечном итоге позволяет прогнозировать и определять полные затраты в отдельных ответственных областях производства и может служить прототипом крупной системы поддержки принятия решений, объединяющей в единый контролируемый цикл десятки и сотни предприятий кооперации промышленности и других организаций. Особенно актуально ее применение при проектировании, разработке и производстве изделий по государственным заказам.

При практическом использовании описанных моделей следует закладывать и проводить пересчет затрат на развитие систем с учетом дисконтирования цен [2–4]. В этом случае, при определении полных предстоящих затрат на развитие проекта на всем плановом (оцениваемом) периоде T необходимо устранить влияние разновременности использования материальных и финансовых ресурсов. Приведение разновременных затрат к одному моменту времени осуществляется с помощью норматива приведения E_{pr}, косвенно отражающего эффективность капитальных вложений (инвестиций). Экономическое содержание и сущность норматива приведения аналогичны проценту по долгосрочному кредиту, который может быть начислен на отвлекаемые (временно «замороженные») средства. Величина E_{pr} также близка к величине норматива платы за фонды, который представляет собой процент, отчисляемый от основных производственных фондов и нормируемых оборотных средств.

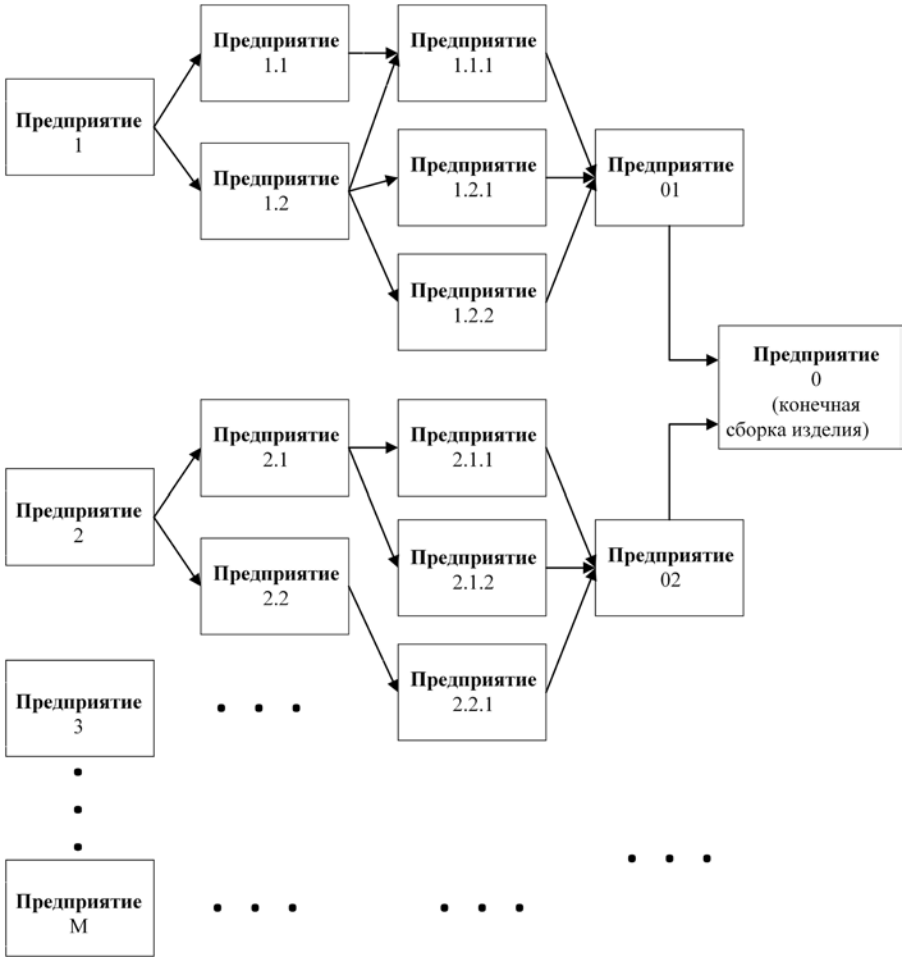

Рис. 2. Пример схемы взаимодействия заводов-изготовителей при производстве изделия

Для приведения разновременных затрат к одному моменту времени показатель затрат определенного такта времени (года, месяца, недели) C_t умножается на коэффициент дисконтирования α_t, определяемый по формуле (7):

$$\alpha_t = (1 + E_{pr}^{\Delta t}), \quad (7)$$

где Δ_t -число интервалов (лет), отделяющее начало расчетного года t_p от года t, в котором осуществляют расход ресурсов: $\Delta_t = t_p - t - 1$.

Тогда полные предстоящие затраты на развитие проекта на всем плановом периоде, приведенные к расчетному такту (году), будут определяться по выражению (8):

$$C = \sum_{t=1}^{T} C_t \alpha_t. \quad (8)$$

Многообразие рассматриваемых процессов, необходимость учета при моделировании ряда неопределенностей [7–11], различные варианты построения схем взаимодействия предприятий-изготовителей комплектующих и др. факторы требуют при обосновании разработки и серийного производства изделий создания специального программного обеспечения, служащего основой для построения единой автоматизированной системы поддержки принятия решений [6,13–21] и выбора на основе моделирования и проводимых расчетов конкретных (конкретного) проектов или вариантов развития БТС.

Заключение

Проведенные исследования и предлагаемые в работе частные и обобщенная экономическая модель развития технических систем по этапам их жизненного цикла направлены на снижение степени неопределенности основных экономических факторов и повышение точности расчетов, на основе которых должны приниматься управленческие решения в области технической политики развития систем.

Практическая реализация экономических моделей заключается в разработке рабочего инструмента — имитационного комплекса, направленного, в конечном итоге, по результатам проводимого моделирования на выработку конкретных экономических решений как по отдельным техническим системам, так и для всего комплекса БТС в целом.

Общие затраты на образец, на одну систему и на комплекс систем складываются из потребных затрат на тактах времени. При моделировании учитываются коэффициенты снижения затрат на соответствующих этапах ЖЦ, характеризующие особенности серийного производства и другие особенности, понижающие затраты на однотипную группу (серию) образцов. Моделируются различные режимы финансирования развития одной системы или их совокупности, учитываются особенности и могут уточняться получаемые значения для каждого этапа. Особый интерес при моделировании представляет возможность оценки различных вариантов финансирования (деления ассигнований как по долям между системами и в самой системе, так и по времени), а также структурной перестройки взаимосвязи предприятий участвующих в процессе производства и изменения очередности возможных технологических процессов производства образцов.

Литература

1. *Заграновская А. В.* Теория систем и системный анализ в экономике. М.: Юрайт, 2019. 266 с.

2. *Власов М. П., Шимко П. Д.* Моделирование экономических процессов. М.: Феникс, 2016. 409 с.

3. *Дрогобыцкий И. Н.* Системный анализ в экономике. М.: ЮНИТИ-ДАНА, 2017. 423 с.

4. *Бураков П. В.* Информационные системы в управлении предприятием. СПб: ИТМО, 2016. 96 с.

5. *Доросинский Л. Г., Зверева О. М.* Информационные технологии поддержки жизненного цикла изделия. Ульяновск: Зебра, 2016. 243 с.

6. *Исмагулова Ф. Е.* Технология моделирования и мониторинга состояния сложных систем — ТОФИ // Новый университет. Серия: Технические науки. 2016. № 3 (13). С. 35–37.

7. *Васильев О. И., Ахмадеев Б. А., Бойченко И. А.* Анализ ресурсопотребления в системе серийного производства стандартного посадочного материала на основе имитационного моделирования // Сборник научных трудов XXI международной научно-практической конференции: в 2-х томах. 2017. С. 403–411.

8. *Габалин А. В.* Вопросы выбора системы имитационного моделирования при исследовании сложных систем // В сборнике: Системы проектирования, технологической подготовки производства и управления этапами жизненного цикла промышленного продукта (CAD/CAM/PDM — 2017). Труды XVII международной научно-практической конференции. Под ред. А. В. Толока. Институт проблем упр. им. В. А. Трапезникова. 2017. С. 107–108.

9. *Боев В. Д., Волков Д. В., Кондрашов Ю. В.* Применение мультиагентного имитационного моделирования для построения системы поддержки принятия решений в системе связи военного назначения // Материалы всероссийской научно-практической конференции «Прошлое, настоящее и будущее российской цивилизации». 2016. С. 331–334.

10. *Чувиков Д. А.* Методика объединения экспертной системы и системы имитационного моделирования // Промышленные АСУ и контроллеры. 2017. № 3. С. 11–18.

11. *Boero R.* Agent-based models of the Economy: from theories to applications. London: Palgrave Macmillan, 2016. 232 p.

12. *Palm W. J.* System Dynamics. 3-rd ed. NY: McGraw-Hill, 2016. 926 p.

13. *Lim E. W. C.* Discrete event simulations: development and applications. Rijeka: InTech, 2017. 206 p.

14. *Sauter V. L.* Decision Support Systems for Business Intelligence. New Jersey: Wiley, 2016. 453 p.

15. *Burstein F., Holsapple C. W.* Handbook on decision support systems 1. Basic themes. Leipzig: Springer, 2017. 902 p.

16. *Karnopp Dean C.* System Dynamics: Modeling, Simulation, and Control of Mechatronic Systems (5-nd ed.). New Jersey: John Wiley & Sons, Inc., 2016. 645 p.

17. *Evon M. O.A.T., Asim A. E. S.* Handbook of research on discrete event simulation environments: technologies and applications. New York: Information Science Reference, 2017. 582 p.

18. *Robinson S.* Conceptual modeling for discrete-event simulation. New York: CRC Press, 2016. 530 p.

19. *Jao C. S.* Efficient Decision Support Systems — Practice and challenges from current to future. Rijeka: InTech, 2017. 566 p.

20. *Devlin G.* Advances in decision support systems. Rijeka: InTech, 2016. 352 p.

21. *Hokamp S., Gulyás L., Koehler M., Wijesinghe S.* Agent-based modeling of tax evasion: theoretical aspects and computational Simulations. Chennai: Wiley, 2018. 364 p.

22. *Маянский В.Д.* Обеспечение качества продукции оборонного значения на различных этапах ее жизненного цикла // Система добровольной сертификации «Военный регистр». URL: https://www.sds-vr.ru/assets/docs/MVK/2017/2_1.pdf (дата обращения 12.03.2021)

23. *Вадзинский Р.Н.* Справочник по вероятностным распределениям. СПб.: Наука, 2016. 295 с.

DEVELOPMENT OF ECONOMIC MODELS OF THE DEVELOPMENT OF TECHNICAL SYSTEMS BY STAGES OF THE LIFE CYCLE

IGOR I. MIKITENKO,
Moscow, Russia, iimiki@bk.ru

KEYWORDS: economic model, large technical system; life cycle stages; system cost estimate; simulation modeling of system development; serial production of products; effective management decisions.

ABSTRACT

Introduction: when forecasting and simultaneously supporting the planned development of a set of large technical systems, it is required to develop an integrated mechanism for accounting for the state and costs of resources. The approaches known for these purposes are characterized by high complexity of implementation and fragmentation over the stages of the system cycle. **Purpose**: to develop a complex of interrelated economic models for the analysis and development of large technical systems. **Results**: on the basis of scientific directions of economic analysis, aspects of the development of economic models were considered and on their basis, modeling of complex composite technical systems by stages of their life cycle was carried out in the interests of the subsequent formation and adoption of managerial decisions both in the selection and implementation of such systems, and in support of their development. The set of economic models is combined into a generalized model of systems development in the interests of a comprehensive assessment of current and future costs, further forecasting and, based on the results of simulation and calculations, making effective decisions. Since a feature of the production of large technical systems is their serial nature of products, then the integrated model evaluates the costs both for individual systems and for the stages of their life cycle, and for the entire complex of systems as a whole. Modeling of various modes of financing the development of one system or their complex is predicted, the features for each stage of development are taken into account. The possibility of assessing the structural restructuring of the relationship of enterprises participating in the production process and changing the possible sequence of technological processes for the production of samples has been obtained. **Practical relevance**: the proposed model (s) for batch or piece production of products is recommended in specific science-intensive areas of mechanical engineering and is applicable, for example, in aviation, space, shipbuilding, mining, and other industries. Its application is especially important in the design, development and manufacture of products for government orders. **Discussion**: the presented solution is proposed to be formalized with a software package and to create an automated control system for the development, production, operation and disposal of large technical systems with the functions of a decision support system.

REFERENCES

1. Zagranovskaja A.V. *Teorija sistem i sistemnyj analiz v jekonomike* [Systems theory and systems analysis in economics]. Moscow: Jurajt, 2019. 266 p. (In Rus)

2. Vlasov M.P., Shimko P.D. *Modelirovanie jekonomicheskih processov* [Modeling economic processes]. Moscow: Feniks, 2016. 409 p. (In Rus)

3. Drogobyckij I.N. *Sistemnyj analiz v jekonomike* [Systems analysis in economics]. Moscow: JUNITI-DANA, 2017. 423 p. (In Rus)

4. Burakov P.V. I*nformacionnye sistemy v upravlenii predprijatiem* [Information systems in enterprise management]. St-Petersburg: ITMO, 2016. 96 p. (In Rus)

5. Dorosinskij L.G., Zvereva O.M. *Informacionnye tehnologii podderzhki zhiznennogo cikla izdelija* [Information technology to support the product life cycle]. Ul'janovsk: Zebra, 2016. 243 p. (In Rus)
6. Ismagulova F.E. Technology for modeling and monitoring the state of complex systems – TOFI. *Novyj universitet. Serija: Tehnicheskie nauki* [New University. Series: Engineering]. 2016. № 3 (13). Pp. 35–37. (In Rus)
7. Vasil'ev O.I., Ahmadeev B.A., Bojchenko I.A. Analiz resursopotreblenija v sisteme serijnogo proizvodstva standartnogo posadochnogo materiala na osnove imitacionnogo modelirovanija [Analysis of resource consumption in the system of serial production of standard planting material based on simulation]. *Sbornik nauchnyh trudov XXI mezhdunarodnoj nauchno-prakticheskoj konferencii* [Collection of scientific papers of the XXI international scientific and practical conference]. In 2 vol. 2017. Pp. 403–411. (In Rus)
8. Gabalin A.V. Voprosy vybora sistemy imitacionnogo modelirovanija pri issledovanii slozhnyh system [Issues of choosing a simulation system in the study of complex systems]. *Trudy XVII mezhdunarodnoj nauchno-prakticheskoj konferencii "Sistemy proektirovanija, tehnologicheskoj podgotovki proizvodstva i upravlenija jetapami zhiznennogo cikla promyshlennogo produkta (CAD/CAM/PDM – 2017)"* [Proceedings of the XVII International Scientific and Practical Conference Systems for design, technological preparation of production and management of stages of the life cycle of an industrial product (CAD / CAM / PDM – 2017). Ed. A.V. Toloka, Institute for Problems of Ex. them. V.A. Trapeznikov]. 2017. Pp. 107–108. (In Rus)
9. Boev V.D., Volkov D.V., Kondrashov Ju.V. Primenenie mul'tiagentnogo imitacionnogo modelirovanija dlja postroenija sistemy podderzhki prinjatija reshenij v sisteme svjazi voennogo naznachenija [Application of multi-agent simulation to build a decision support system in a military communications system]. Materialy vserossijskoj nauchno-prakticheskoj konferencii "Proshloe, nastojashhee i budushhee rossijskoj civilizacii" [Materials of the All-Russian Scientific and Practical Conference "Past, present and future of Russian civilization"]. 2016. Pp. 331–334. (In Rus)
10. Chuvikov D.A. Technique for combining expert system and simulation system. *Promyshlennye ASU i kontrollery* [Industrial ACS and controllers]. 2017. No. 3. Pp. 11–18. (In Rus)
11. Boero R. *Agent-based models of the Economy: from theories to applications*. London: Palgrave Macmillan, 2016. 232 p.
12. Palm W.J. *System Dynamics*. 3-rd ed. NY: McGraw-Hill, 2016. 926 p.
13. Lim E.W.C. *Discrete event simulations: development and applications.* Rijeka: InTech, 2017. 206 p.
14. Sauter V.L. *Decision Support Systems for Business Intelligence*. New Jersey: Wiley, 2016. 453 p.
15. Burstein F., Holsapple C.W. Handbook on decision support systems 1. Basic themes. Leipzig: Springer, 2017. 902 p.
16. Karnopp Dean C. *System Dynamics: Modeling, Simulation, and Control of Mechatronic Systems*. 5-nd ed. New Jersey: John Wiley & Sons, Inc., 2016. 645 p.
17. Evon M.O.A.T., Asim A.E.S. *Handbook of research on discrete event simulation environments: technologies and applications.* New York: Information Science Reference, 2017. 582 p.
18. Robinson S. *Conceptual modeling for discrete-event simulation*. New York: CRC Press, 2016. 530 p.
19. Jao C.S. *Efficient Decision Support Systems – Practice and challenges from current to future*. Rijeka: InTech, 2017. 566 p.
20. Devlin G. *Advances in decision support systems*. Rijeka: InTech, 2016. 352 p.
21. Hokamp S., Gulyás L., Koehler M., Wijesinghe S. A*gent-based modeling of tax evasion: theoretical aspects and computational Simulations.* Chennai: Wiley, 2018. 364 p.
22. Majanskij V.D. Obespechenie kachestva produkcii oboronnogo znachenija na razlichnyh jetapah ee zhiznennogo cikla [Ensuring the quality of defense products at various stages of their life cycle]. *Sistema dobrovol'noj sertifikacii "Voennyj registr"* [Voluntary certification system "Military Register"]. URL: https://www.sds-vr.ru/assets/docs/MVK/2017/2_1.pdf (date of access 12.03.2021). (In Rus)
23. Vadzinskij R.N. *Spravochnik po verojatnostnym raspredelenijam* [Probability Distributions Handbook]. St-Petersburg: Nauka, 2016. 295 p. (In Rus)

INFORMATION ABOUT AUTHOR:

Mikitenko I.I., PhD, Senior Research Officer, Senior lecturer of the National Research Technological University Moscow Institute of Steel and Alloys.

For citation: Mikitenko I.I. Development of economic models of the development of technical systems by stages of the life cycle. *H&ES Research*. 2021. Vol. 13. No. 2. Pp. 44–51. Doi: 10.36724/2409-5419-2021-13-2-44-51 (In Rus)

Doi: 10.36724/2409-5419-2021-13-2-52-65

МОДЕЛИРОВАНИЕ УГРОЗ БЕЗОПАСНОСТИ ИНФОРМАЦИИ В ТЕРРИТОРИАЛЬНО-РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

МИНЯЕВ
Андрей Анатольевич

АННОТАЦИЯ

Введение: при проектировании систем защиты информации важным этапом является моделирование угроз безопасности, что подразумевает за собой определение перечня актуальных угроз для конкретной информационной системы, на основании которого принимаются решения по нейтрализации актуальных угроз. В настоящее время подавляющее большинство систем являются территориально-распределенными, что, потенциально, увеличивает количество актуальных угроз безопасности информации, ввиду сложности инфраструктуры, особенностей технологии обработки информации, незащищенных каналов связи. В этой связи, **целью исследования** является моделирование угроз безопасности информации в территориально-распределенных информационных системах. Проведенный анализ зарубежных и российских методологий моделирования угроз безопасности информации информационных систем показал, что основными проблемами являются большой объем данных для моделирования, а также недостатки экспертных методов. Установлено, что для решения поставленной задачи необходимо использовать **методы** машинного обучения, теорию адаптивных нечетких нейронных продукционных систем с алгоритмами нечеткого вывода и применение технологий Data Science при обработке большого объема данных. В работе были определены необходимые и достаточные показатели для моделирования угроз безопасности информации, на основе нечетких нейронных продукционных системах была предложена методика определения актуальных угроз безопасности информации. Методика автоматизирована и, гипотетически, исключает ошибки экспертов, увеличивает количество определяемых актуальных угроз безопасности информации, снижает финансовые затраты на закупку средств защиты информации, отличается от существующих тем, что процесс автоматизирован, обладает низкой вычислительной сложностью, отсутствует необходимость привлечения высококвалифицированных специалистов, позволяет определять перечень актуальных угроз в системах различных типов и классов, может быть адаптирована для работы с международными базами данных. **Практическая значимость** заключается в автоматизации процесса – разработке программы для ЭВМ, реализующей предложенную методику. **Обсуждение:** дальнейшие исследования целесообразно продолжить в определении наилучших параметров адаптивных нечетких нейронных продукционных систем и алгоритмов нечеткого вывода.

Сведения об авторе:
старший преподаватель Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, г. Санкт-Петербург, Россия, minyaev.a@gmail.com

КЛЮЧЕВЫЕ СЛОВА: угрозы безопасности информации; территориально-распределенные информационные системы; методологии моделирования; Data Science; ANFIS.

Для цитирования: *Миняев А.А.* Моделирование угроз безопасности информации в территориально-распределенных информационных системах // Наукоемкие технологии в космических исследованиях Земли. 2021. Т. 13. № 2. С. 52–65. Doi: 10.36724/2409-5419-2021-13-2-52-65

Введение

Моделирование угроз безопасности информации в информационных системах (ИС) является одним из основных этапов создания систем (подсистем) защиты информации (СЗИ) ИС. С точки зрения законодательства Российской Федерации (РФ) в области обеспечения безопасности информации этот этап жизненного цикла создания СЗИ является обязательным в соответствии с частью 2 статьи 19 закона № 152-ФЗ «О персональных данных», в соответствии с пунктом 4 приказа ФСТЭК России от 18 февраля 2013 г. № 21, а также в соответствии с приказом ФСТЭК России от 11 февраля 2013 г. № 17: «определение угроз безопасности информации, реализация которых может привести к нарушению безопасности информации в информационной системе, и разработку на их основе модели угроз безопасности информации».

В соответствии с методическими документами[1] (МД) регуляторов РФ этап моделирования УБИ необходимо проводить на этапе создания СЗИ, т.е. в том случае, когда еще на создана СЗИ, а иногда и сама ИС. В этом случае провести инструментальный анализ защищенности, необходимый в соответствии с МД не представляется возможным [1, 2]. В этой связи моделирование угроз безопасности информации (УБИ), как правило, проводится экспертным путем с привлечением специалистов в области информационной безопасности [3–9]. Такой подход является трудоемким и не исключает известные недостатки экспертного метода [10].

[1]Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных (Выписка) (утв. ФСТЭК России 15.02.2008).
Методика определения актуальных угроз безопасности персональных данных при их обработке в информационных системах персональных данных (утв. ФСТЭК России 14.02.2008).
Методические рекомендации по разработке нормативных правовых актов, определяющих угрозы безопасности персональных данных, актуальные при обработке персональных данных в информационных системах персональных данных, эксплуатируемых при осуществлении соответствующих видов деятельности (утв. ФСБ России 31.03.2015).
Методика моделирования угроз безопасности информации (проект ФСТЭК России, 2020 г.).

Методологии моделирования угроз безопасности информации

В настоящее время в мире существуют множество методологий моделирования УБИ. Основными их них являются следующие [11–16]:

1. STRIDE and Associated Derivations.
2. PASTA (The Process for Attack Simulation and Threat Analysis).
3. LINDDUN (Linkability, Identifiability, Non-Repudiation, Detectability, Disclosure of Information, Unawareness, Non-Compliance).
4. CVSS (Common Vulnerability Scoring System).
5. Attack Trees.
6. Persona non Grata.
7. Security Cards.
8. hTMM.
9. Quantitative Threat Modeling Method.
10. Trike.
11. VAST Modeling.
12. OCTAVE.
13. Mitre.
14. БДУ ФСТЭК России.

В табл. 1 приведен анализ существующих методологий моделирования УБИ.

Таблица 1

Анализ существующих методологий моделирования УБИ

Метод моделирования УБИ	Преимущества	Недостатки
STRIDE	Является ли наиболее зрелым. Прост в использовании	Низкий уровень ложных срабатываний и умеренно высокий уровень ложных отрицательных результатов. Длительный по временным характеристикам
PASTA	Непосредственный вклад в управление рисками. Содержит встроенные приоритеты по смягчению угроз. Наличие исчерпывающей документации	Требует ввода данных по безопасности от операционного уровня, уровня управления, архитектуры и развития ИС. Содержит встроенные приоритеты по смягчению угроз. Трудоемкий

Продолжение таблицы 1

Метод моделирования УБИ	**Преимущества**	**Недостатки**
LINDDUN	Содержит встроенные приоритеты для минимизации УБИ	Количество УБИ может быстро возрастать по мере модификации, усовершенствования и усложнения ИС. На эффективность и результативность данного метода отрицательно влияет угрозы общего характера. Трудоемкий. Длительный по временным характеристикам
CVSS	Наличие автоматизированных компонентов. Содержит встроенную приоритизацию мер по минимизации УБИ. Имеет постоянные результаты при повторении	Непрозрачные расчеты баллов и возможные несоответствия, вызванные экспертными оценками. Сложен в использовании. Нет руководства по оценке подцелей, атак или рисков
Attack Trees	Имеет постоянные результаты при повторении. Прост в использовании, если есть полное представление об ИС	Полезен только тогда, когда ИС и проблемы безопасности в ней хорошо известны и поняты. Эксперты должны имеют высокий опыт в области кибербезопасности. Не приводится руководство по оценке подцелей, атак или рисков
Persona non Grata	Прост в применения. Вносит непосредственный вклад в управление рисками. Имеет постоянные результаты при повторении	Имеет тенденцию обнаруживать только определенное подмножество типов УБИ
Security Cards	Определяет сложные угрозы	Приводит ко многим ложным срабатываниям
hTMM	Отсутствие ложных срабатываний. Отсутствие пропущенных УБИ. Стабильный результат независимо от того, кто занимается моделированием УБИ. Экономическая эффективность	Применение Security Cards на основе предложений разработчиков. Обязательное использование инструментальных средств. Трудоемкий. Длительный
Quantitative TMM	Содержит встроенные приоритеты для минимизации УБИ. Имеет автоматизированные компоненты	Трудоемкий. Длительный
Trike	Непосредственный вклад в управление рисками. Содержит встроенные приоритеты для минимизации УБИ. Имеет автоматизированные компоненты	Система весов Trike кажется слишком расплывчатой. Плохо поддерживается и нет никакой документации
VAST Modeling	Подход позволяет интегрировать VAST в развитие организации и DevOps. Непосредственный вклад в управление рисками. Содержит встроенную приоритизацию мер по минимизации рисков. Имеет согласованные результаты при повторении. Имеет автоматизированные компоненты	Предполагает наличие высококвалифицированных специалистов в области ИБ. Трудоемкий. Длительный

Окончание таблицы 1

Метод моделирования УБИ	Преимущества	Недостатки
OCTAVE	Метод глубокий, но гибкий	Процесс моделирования УБИ требует значительного времени, большой объем документации
Mitre Att&ck	Своевременное реагирование на инциденты. Расследование инцидентов. Анализ деятельности нарушителей. Имеет автоматизированные компоненты	Процесс моделирования УБИ требует значительного времени. Большой объем документации. Требует высокой квалификации специалистов в области ИБ
БДУ ФСТЭК России	Достаточно полный объем перечня известных уязвимостей, УБИ, потенциальных нарушителей и возможных методов нейтрализации УБИ. Структурирован. Имеется документация регулятора для процесса моделирования УБИ	Нет автоматизированных средств для моделирования УБИ Перечень УБИ и уязвимостей не является исчерпывающим. УБИ не являются элементами иерархической классификационной системы угроз. Предполагает наличие высококвалифицированных специалистов в области ИБ. Трудоемкий. Длительный

По результатам проведенного исследования можно сказать о том, что существующие методологии в большинстве своем имеют существенные недостатки, а именно: большой объем данных, отсутствие документации, отсутствие автоматизированных средств моделирования УБИ, необходимость в высокой квалификации специалистов по информационной безопасности.

В связи с вышесказанным в настоящей работе были поставлены следующие задачи:

1. Подготовка набора данных для моделирования УБИ на основании известных баз данных УБИ и уязвимостей, а также разработанных ранее моделей угроз для территориально-распределенных ИС.
2. Анализ сформированного набора данных.
3. Форматирование набора данных для последующей автоматизированной обработки.
4. Выбор и сравнение качества работы нескольких моделей, определение наилучшей.
5. Оптимизация параметров в наилучшей модели.
6. Проверка модели.
7. Разработка программы для ЭВМ для моделирования УБИ «Модель угроз и нарушителя» (МУиН).
8. Итоговое представление результатов работы.

Подготовка набора данных для моделирования УБИ

Для преставления набора данных для автоматизированной обработки и разработки программного обеспечения использовался язык программирования Python 3 и технологии Data Science [16, 18].

В качестве объекта исследования была выбрана территориально-распределенная ИС, которая является информационной системой обработки персональных данных и государственной информационной системой одновременно [1].

ИС обеспечивают принцип централизованного хранения, накопления и многократного использования данных. Для экономии ресурсов и обеспечения информационной безопасности на автоматизированных рабочих местах (АРМ) пользователей хранение данных не осуществляется. Обработка осуществляется на стороне серверной части. Для этого могут использоваться технологии терминального доступа. Такая технология может быть реализована в IT-инфраструктуре путем развертывания терминальной фермы RDS (Romote Desktop Services). При такой технологии рабочие профили пользователей хранятся на серверах терминальной фермы. Это упрощает организацию доступа пользователей к информационным ресурсам ИС, а также наиболее эффективно обеспечивает процессы информационной безопасности. Также технология обеспечивает процессы при удаленной работе пользователей.

Для территориально-распределенных ИС характерно размещение серверных компонент, сетевого оборудования и АРМ пользователей на всей территории страны и, возможно, за ее пределами. В этом случае такие ИС имеют сложную архитектуру с точки зрения расположения своих компонентов и технологий обработки информации.

В состав типовой территориально-распределенной ИС входят следующие компоненты:

1. серверы, включающие:

– серверное оборудование;

– прикладные и специализированные программы, обеспечивающие обработку информации и ее представление в виде, необходимом для последующей автоматизированной обработки;

2. АРМ пользователей:

– типовые АРМ пользователя (стационарные персональные компьютеры);

– мобильное АРМ: планшеты, мобильные телефоны;

– планшеты.

Какивбольшинстветерриториально-распределенных системах присутствуют иные категории пользователей, не относящихся к работникам владельца ИС. К таким категориям пользователей относятся разработчики программного обеспечения и технических средств — лица, обеспечивающие разработку, поставку и внедрение программных, аппаратно-программных и аппаратных средств ИС. Обслуживающий персонал — лица, обеспечивающие штатное функционирование технических средств ИС (работники эксплуатационных подразделений владельца ИС, осуществляющие обслуживание и техническое сопровождение инженерных систем и помещений, в которых размещается оборудование ИС).

Сложностью моделирования УБИ для такой ИС является обработка большого объема данных [13], необходимого при моделировании УБИ, а именно: данные из БДУ ФСТЭК России[2] и/или зарубежных баз данных и знаний, сложность использования методических документов регуляторов РФ в области обеспечения безопасности информации, таких как:

Базовая модель угроз ПДн ФСТЭК России, 2008 г.

Методика определения актуальных угроз ПДн, 2008 г.

Методический документ ФСТЭК России. Меры защиты информации в государственных информационных системах (утв. ФСТЭК России 11.02.2014 г.)

Проект методического документа ФСТЭК России. Методика определения угроз безопасности информации в информационных системах (2020 г.)

Банк данных угроз ФСТЭК России.

«Методические рекомендации по разработке нормативных правовых актов, определяющих угрозы безопасности персональных данных, актуальные при обработке персональных данных в информационных системах персональных данных, эксплуатируемых при осуществлении соответствующих видов деятельности» (утв. ФСБ России 31.03.2015 N149/7/2/6–432).

В данной работе набор данных был сформирован на основании данных из БДУ ФСТЭК России и на основании ранее разработанных моделей угроз подобных ИС, описанной выше.

На рис. 1 и 2 представлены графики анализа данных БДУ ФСТЭК России.

Также был проведен анализ уязвимостей БДУ ФСТЭК России. На рис. 3 представлен график зависимости количества уязвимостей от производителя.

Как видно из графиков информация имеет большой объем, что является трудоемким в процессе моделирования УБИ [1, 17]. При экспертном подходе моделирования УБИ влечет за собой ошибки, связанные с таким подходом,

[2]https://bdu.fstec.ru/

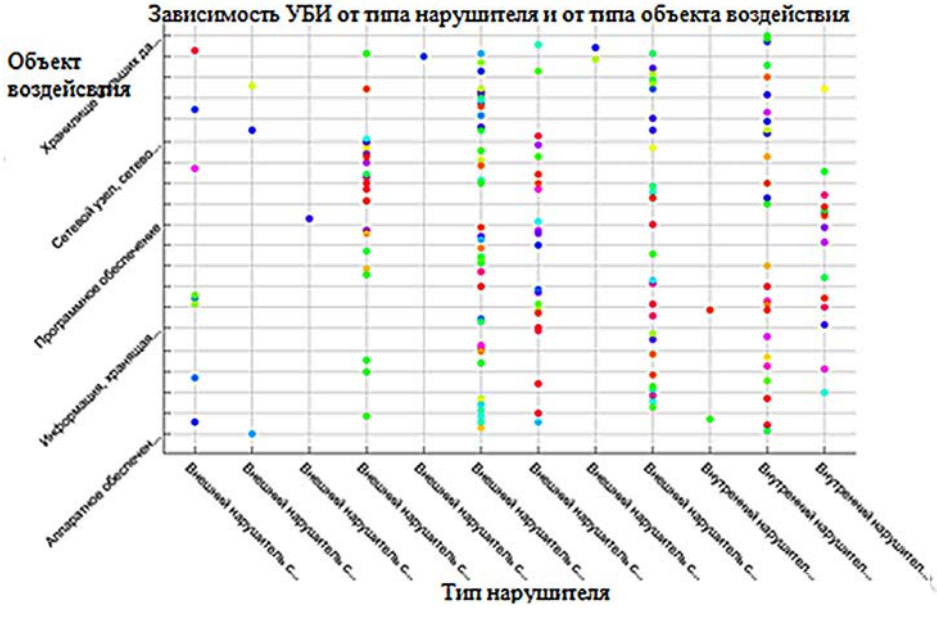

Рис. 1. Зависимость УБИ от типа нарушителя и от типа объекта воздействия

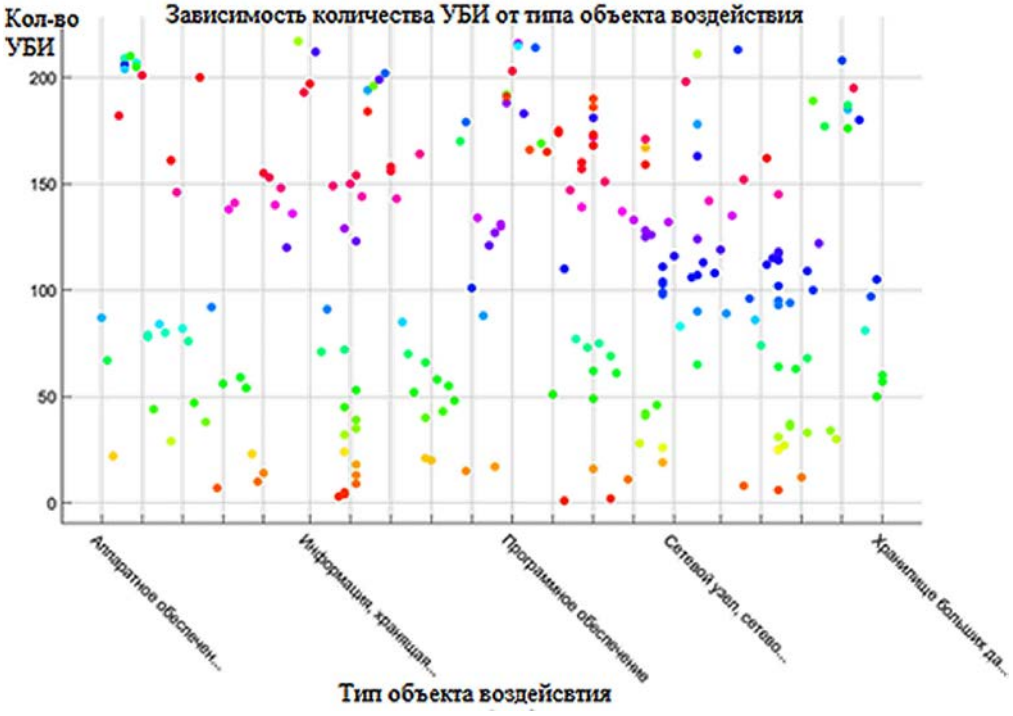

Рис. 2. Зависимость количества УБИ от типа объекта воздействия

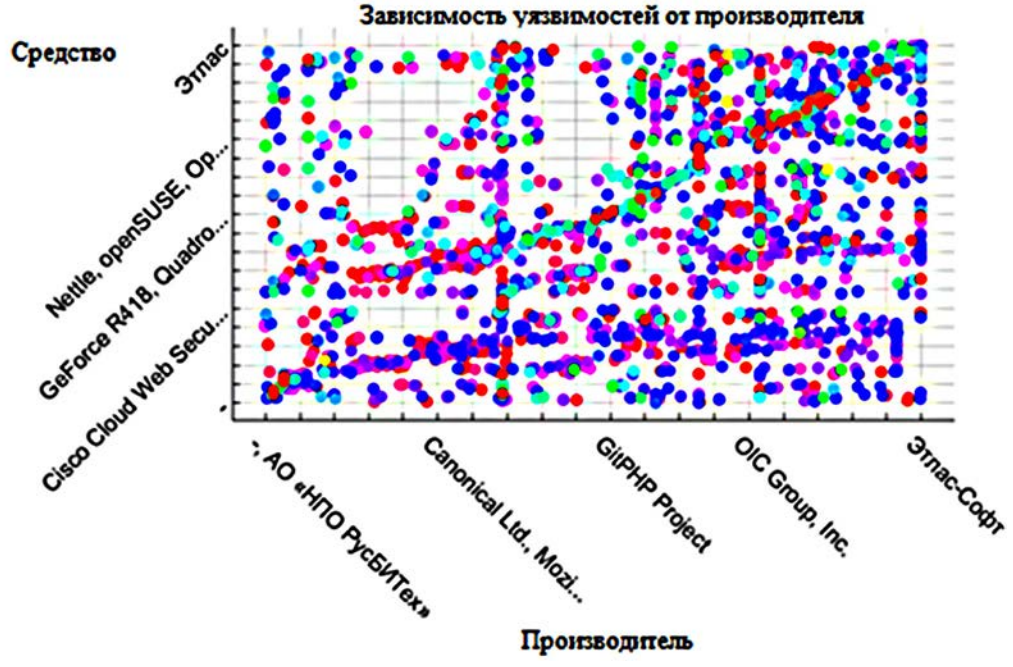

Рис. 3. График зависимости уязвимостей от производителя

такие как личное мнение эксперта, разрозненность и несогласованность мнений, трудоемкость. Методические документы регуляторов РФ определяют подход и этапы моделирования УБИ, при этом ошибки экспертов и трудоемкость не учитываются [1, 19–20].

В соответствии с МД регуляторов РФ актуальность УБИ определяется в зависимости от актуального нарушителя в ИС, перечнем потенциальных УБИ и уязвимостей в ИТ-инфраструктуре ИС, а также возможными последствиями от реализации УБИ. В этой связи набора данных

был сформирован из сведений базы данных угроз ФСТЭК России, моделей УБИ территориально-распределенных ИС и технических решений исследуемой ИС.

Форматирование набора данных

На рис. 4 и 5 изображены наборы данных dataframe БДУ ФСТЭК России до конвертации данных.

Для подготовки автоматизированной обработки данных необходимо выполнить конвертацию данных. Код на языке программирования Python 3 для преобразования данных dataframe (конвертирование данных) представлен ниже:

```
train = pd.read_csv('threats.csv', encoding='utf-8')
df = pd.get_dummies(train)
# Преобразование строковых данных
for col in list(df.columns):
    # Выбор колонок для преобразования
if ('ft²' in col or 'kBtu' in col or 'Metric Tons CO2e' in col or 'kWh' in
    col or 'therms' in col or 'gal' in col or 'Score' in col):
    # Конвертация
    df[col] = df[col].astype(float)
```

Выбор модели

Было проведено исследование нечетких нейронных продуктивных сетей ANFIS (adaptive neuro-fuzzy inference system) с применением алгоритмов Сугено-Такаги, Такаги-Сугено-Канга (TSK), Ванга-Менделя и Мамдани. Отмечено, что зависимость погрешности от количества правил при проверке на тестовой выборке меньше у сети TSK. В этой связи для моделирования УБИ была выбрана нечеткая продуктивная сеть ANFIS, основанная на нечеткой системе вывода Такаги-Сугено-Канга (TSK).

Алгоритм ее работы заключается в реализации нечеткой продукционной модели, основанной на правилах типа:

$$R_i : IFx_i ISA_{i1} AND...ANDx_j ISA_{ij} AND...ANDx_{im} ISA_{im}, THEN$$

$$y = c_{i0} + \sum_{j=1}^{m} c_{ij} x_j), j = 1,...,n$$

Для этого была сформирована база правил для моделирования УБИ. Пример заполнения базы правил для моделирования УБИ на основании сформированного в данной работе набора данных приведен в табл. 2.

```
In [5]: r_bdu = pd.read_excel('thrlist (8).xlsx')
        r_bdu
```

Out[5]:

	Общая информация	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Последствия	Unnamed: 6	Unnamed: 7	Дополните
0	Идентификатор УБИ	Наименование УБИ	Описание	Источник угрозы (характеристика и потенциал на...	Объект воздействия	Нарушение конфиденциальности	Нарушение целостности	Нарушение доступности	Дата вклю угрозы
1	1	Угроза автоматического распространения вредоно...	Угроза заключается в возможности внедрения и э...	Внешний нарушитель со средним потенциалом, Вну...	Ресурсные центры грид-системы	1	1	1	2015-00
		Угроза агрегирования	Угроза	Внешний					

Рис. 4. Dataframe УБИ БДУ ФСТЭК России до конвертации

```
In [9]: r_vul = pd.read_excel('vullist.xlsx')
        r_vul
```

Out[9]:

	Описание уязвимостей	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnam
0	Общая информация	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
1	Идентификатор	Наименование уязвимости	Описание уязвимости	Вендор ПО	Название ПО	Версия ПО	Тип ПО	Наименование ОС и тип аппаратной платформы	Уязви
2	BDU:2014-00001	Уязвимость микропрограммного обеспечения прогр...	Микропрограммное обеспечение модуля 140NOE7711...	Schneider Electric	Микропрограммное обеспечение программируемого ...	4.6 (Микропрограммное обеспечение программируе...	ПО программно-аппаратного средства АСУ ТП	Schneider Electric Микропрограммное обеспечени...	Уязви архите
3	BDU:2014-00002	Уязвимость микропрограммного обеспечения маршр...	Скрипт «/sogi-bin/platform.cgi» микропрограмми...	D-Link Corp.	Микропрограммное обеспечение маршрутизатора D-...	1.02b11 (Микропрограммное обеспечение маршрути...	ПО сетевого программно-аппаратного средства	D-Link Corp. Микропрограммное обеспечение марш...	Уязви

Рис. 5. Dataframe уязвимостей БДУ ФСТЭК России до конвертации

Таблица 2

База правил метода моделирования УБИ

№	ЕСЛИ (IF)			ТО (THEN)
	Тип нарушителя (источник воздействия)	ИТ – инфраструктура (объект воздействия)	Версия ПО	
1	Внешний нарушитель с низким потенциалом, Внутренний нарушитель с низким потенциалом	Виртуальная машина VMWare	6.5 (VMWare Workstation), от 7.0.0 до 7.1.4 включительно (VMWare Workstation)	Угроза несанкционированного доступа к защищаемым виртуальным машинам со стороны других виртуальных машин
2	Внешний нарушитель с высоким потенциалом	Мобильное устройство (аппаратное устройство) на базе iOS	(Android), до 10.3.3 включительно (iOS)	Угроза контроля вредоносной программой списка приложений, запущенных на мобильном устройстве
…				
n	Внешний нарушитель со средним потенциалом, Внутренний нарушитель со средним потенциалом	Средство защиты информации	12.4 (Cisco IOS), 12.4 (Cisco IOS), 15.0 (Cisco IOS), 15.0 (Cisco IOS), 15.1 (Cisco IOS), 15.1 (Cisco IOS), 12.2 (Cisco IOS), 12.2 (Cisco IOS), 15.2 (Cisco IOS), 15.2 (Cisco IOS)	Угроза несанкционированного воздействия на средство защиты информации

Количество правил *n* для случая моделирования УБИ на основе БДУ ФСТЭК России равно 222. Для определения уязвимостей на основе БДУ ФСТЭК России оно будет составлять *n* = 30321. Правила представлены в табл. 3 как единое, фактически оно представляет множество правил, состоящих отдельно по типу нарушителя, типу СрЗИ (например, SecretNet, Dallas Lock и т.д.) и по воздействию. Набор данных был сформирован уже с учетом этих нюансов. Объект воздействия — совокупность данных из БДУ ФСТЭК России и данных об ИТ-инфраструктуре ИС, взятых из моделей УБИ и проектных решений по СЗИ.

В предлагаемой методике определения актуальных УБИ ANFIS базируется на следующих положениях[1]:

– входные переменные являются четкими;

– функции принадлежности (ФП) определены функцией Гаусса:

$$\mu_{A_{ij}}(x_j) = \exp(-\frac{1}{2}(\frac{x_j - a_{ij}}{b_{ij}})^2)$$

где x — входные сети a_{ij}, b_{ij} — настраиваемые параметры ФП.

— нечеткая импликация Ларсена — нечеткое произведение;

– Т-норма — нечеткое произведение;

– композиция не производится;

– метод дефаззификации — метод центроида.

Функциональная зависимость после дефаззификации для получения выходной переменно принимает следующий вид [18]:

$$y' = \frac{\Sigma_i^n((c_{i0} + \Sigma_{j=1}^m c_{ij}x_j)\Pi_j^m \mu_{A_{ij}}(x_j'))}{\Sigma_{i=1}^n \Pi_j^m \mu_{A_{ij}}(x_j')} = \frac{\Sigma_i^n((c_{i0} + \Sigma_{j=1}^m c_{ij}x_j)\Pi_j^m \exp\left[-(\frac{x_j' - a_{ij}}{b_{ij}})^2\right]}{\Sigma_{i=1}^n \Pi_j^m \exp\left[-(\frac{x_j' - a_{ij}}{b_{ij}})^2\right]} \quad (1)$$

Выражение (1) лежит в основе сети ANFIS с применением алгоритма TSK, которая включает в себя пять слоев:

1) Слой состоит из элементов, которые выполняют фаззификацию входных четких переменных:

$$x_j'(j = 1, ..., n).$$

Элементы данного слоя вычисляют значения степеней ФП $\mu_{A_{ij}}\left[x_j'\right]$, заданных функциями Гаусса с параметрами a_{ij}, b_{ij} .

2) Количество элементов этого слоя равно количеству правил в базе (база правил, приведенная в данной работе выше), выполняет нечеткую импликацию степеней принадлежности правил.

[1] Хижняков Ю.Н. Алгоритмы нечеткого, нейронного, нейро-нечеткого управления в системах реального времени. Пермь: ПНИПУ, 2013. 160 с.

3) Данный слой генерирует значения функций $(c_{j0} + \Sigma_{j=1}^{m} c_{ij} x_j^{'})$, которые перемножаются на результаты вычислений элементами второго слоя.

4) Первый элемент слоя 4 необходим для активизации заключений правил в соответствии со значениями, агрегированных в 3 слое, степеней принадлежности предпосылок правил. Второй элемент четвертого слоя производит дополнительные вычисления для последующей дефаззификации результата работы сети ANFIS.

5) Данный слой состоит из одного нормализующего элемента и производит дефаззификацию результатов работы сети ANFIS.

Исходя из вышесказанного следует, что сеть ANFIS TSK содержит два параметрических слоя (слой 1 и 3). Настраиваемыми в процессе обучения сети ANFIS параметрами являются:

– в 1 слое — нелинейные параметры a_{ij}, b_{ij} ФП фаззификатора;

– в 3 слое — параметры c_{i0} и c_{ij} линейных функций $(c_{j0} + \Sigma_{j=1}^{m} c_{ij} x_j^{'})$ из заключений базы правил.

При наличии n правил и m-входных переменных число параметров 1 слоя равно $2nm$, а 2 — $n(m+1)$. Суммарное общее число настраиваемых параметров равно $n(3m+1)$.

На следующем шаге рассчитываются параметры c_{i0} и c_{ij} с линейных функций при условии фиксированных значений параметров a_{ij}, b_{ij}. Параметры c_{i0} и c_{ij} находятся путем решения системы линейных уравнений.

Выходную переменную из выражения (1) представляем в следующем виде:

$$y^{'} = \Sigma \sum_{i=1}^{n} w_i^{'} (c_{i0} + \Sigma \sum_{j=1}^{m} c_{ij} x_j),$$

где

$$w_i^{'} = \frac{\Pi_{j=1}^{m} \mu_{A_{ij}} (x_j^{'})}{\sum_{i=1}^{n} \Pi_j^m \mu_{A_{ij}} (x_j^{'})} = \frac{\Pi_j^m \exp\left[-(\frac{x_{ij}^{'} - a_{ij}}{b_{ij}})^2 \right]}{\sum_{i=1}^{n} \Pi_j^m \exp\left[-(\frac{x_{ij}^{'} - a_{ij}}{b_{ij}})^2 \right]} = const$$

Алгоритм обучения сети ANFIS с применением алгоритма TSK.

При k обучающих примерах $x_1^{(k)}, x_2^{(k)}, ..., x_m^{(k)}, y^{(k)}$, где $k = 1, \ldots, K$ и замене значений выходных переменных $y^{'(k)}$ значениями эталонных переменных $y^{(k)}$, получим систему из k линейных уравнений вида:

$$\begin{bmatrix} w_1^{'(1)} & w_1^{'(1)} x_1^{(1)} & \ldots & w_1^{'(1)} x_m^{(1)} & \ldots & w_n^{'(1)} & w_n^{'(1)} x_1^{(1)} & \ldots & w_n^{'(1)} x_m^{(1)} \\ w_1^{'(2)} & w_1^{'(2)} x_1^{(2)} & \ldots & w_1^{'(2)} x_m^{(2)} & \ldots & w_n^{'(2)} & w_n^{'(2)} x_1^{(2)} & \ldots & w_n^{'(2)} x_m^{(2)} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ w_1^{'(k)} & w_1^{'(k)} x_1^{(k)} & \ldots & w_1^{'(k)} x_m^{(k)} & \ldots & w_n^{'(k)} & w_n^{'(k)} x_1^{(k)} & \ldots & w_n^{'(k)} x_m^{(k)} \end{bmatrix} x \begin{bmatrix} c_{10} \\ \ldots \\ c_{1m} \\ \ldots \\ c_{n0} \\ \ldots \\ c_{nm} \end{bmatrix} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \ldots \\ y^{(k)} \end{bmatrix} \quad (3)$$

где $w_i^{'(k)}$ агрегированная степень истинности предпосылок по i-му правилу при предъявлении k-го входного вектора $(x_1^{(k)}, x_2^{(k)}, ..., x_m^{(k)})$.

Таким образом, (3) в сокращенном виде:

$$Wc = y$$

Размерность матрицы W равна $k(m+1)n$, при этом, как правило, количество строк k значительно больше количества столбцов: $k(m+1)n$. Решение этой системы уравнений можно провести за один шаг при помощи псевдоинверсии матрицы W:

$$c = W^{+} y = (W^T W)^{-1} W^T y$$

После определения линейных параметров ij фиксируем и рассчитываем фактические выходные сигналы сети для всех примеров, для чего используем линейную зависимость:

$$y^{'} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \ldots \\ y^{(k)} \end{bmatrix} = Wc$$

определяем вектор ошибок:

$$e = y^{'} - y$$

производим уточнение параметры:

$$a_{ij}^{(k)}(t+1) = a_{ij}^{(k)}(t) - C \frac{dE^{(k)}(t)}{da_{ij}^{(k)}}$$

$$b_{ij}^{(k)}(t+1) = b_{ij}^{(k)}(t) - C \frac{dE^{(k)}(t)}{db_{ij}^{(k)}}$$

После уточнения нелинейных параметров процесс адаптации параметров запускаем до тех пор, пока не наступит повторяемость результатов. Данный алгоритм является гибридным. Его особенность заключается в разделении этапов процесса обучения. Данный алгоритм более эффективен, чем аналогичные методы, например, метод Уидроу-Хоффа, у которого уточнение всех параметров выполняется параллельно и одновременно.

Структура нечеткой нейронной продукционной сети ANFIS с применением алгоритма TSK представлена на рис. 6.

Для проведения вычислений и определения актуальных УБИ в данной работе была разработана программа для ЭВМ «МУиН» на языке программирования Python 3, а также расчеты были проведены в среде MATLAB для сравнения и иллюстрации исследований.

Оптимизация параметров в наилучшей модели, проверка модели

При первоначальных исходных данных и параметров сети ANFIS ошибка обучения сети составляла 3,6–3,7. В ходе проведения экспериментов было установлено, что при определенных параметрах сети ANSIS и оптимизированного набора исходных данных ошибка обучения уменьшается.

На рис. 7–9 представлены настройки сети ANFIS в среде MATLAB.

В результате проведения обучения сети ANFIS с параметрами, указанными на рис. 8–10, при сформированном наборе данных ошибка обучения сети была равной 0,014, что является достаточно хорошим результатом работы.

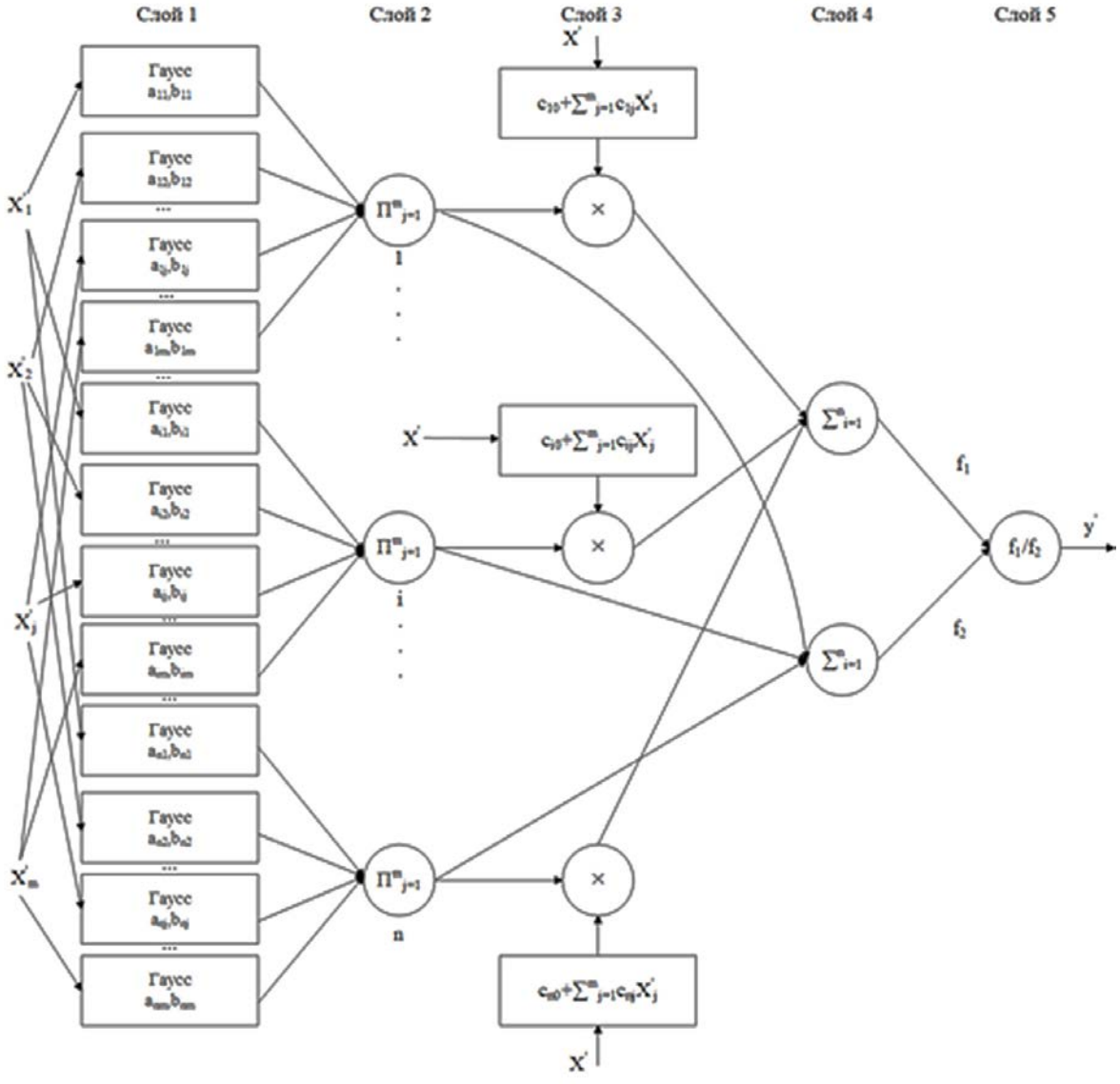

Рис. 6. Сеть ANFIS с применением алгоритма TSK

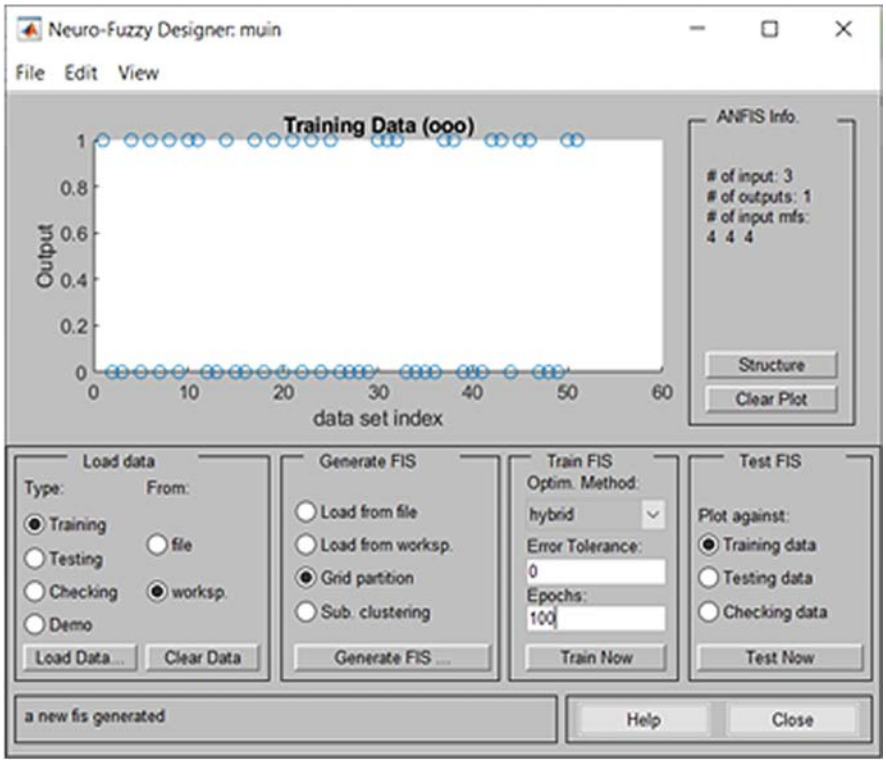

Рис. 7. Настройки ANFIS в среде MATLAB и распределение обучающих данных

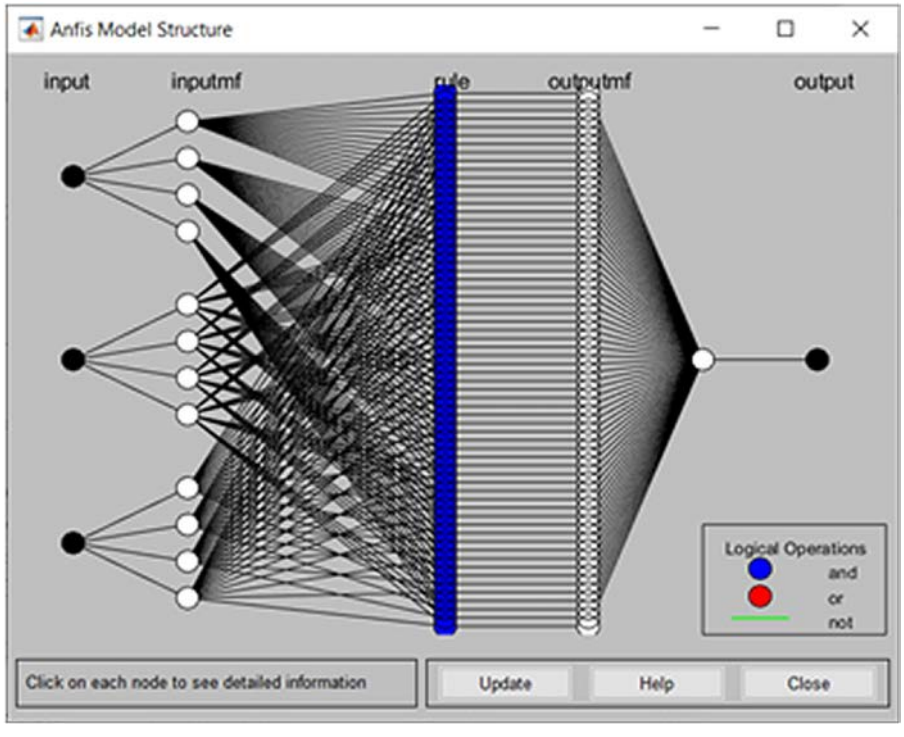

Рис. 8. Структура сети ANFIS

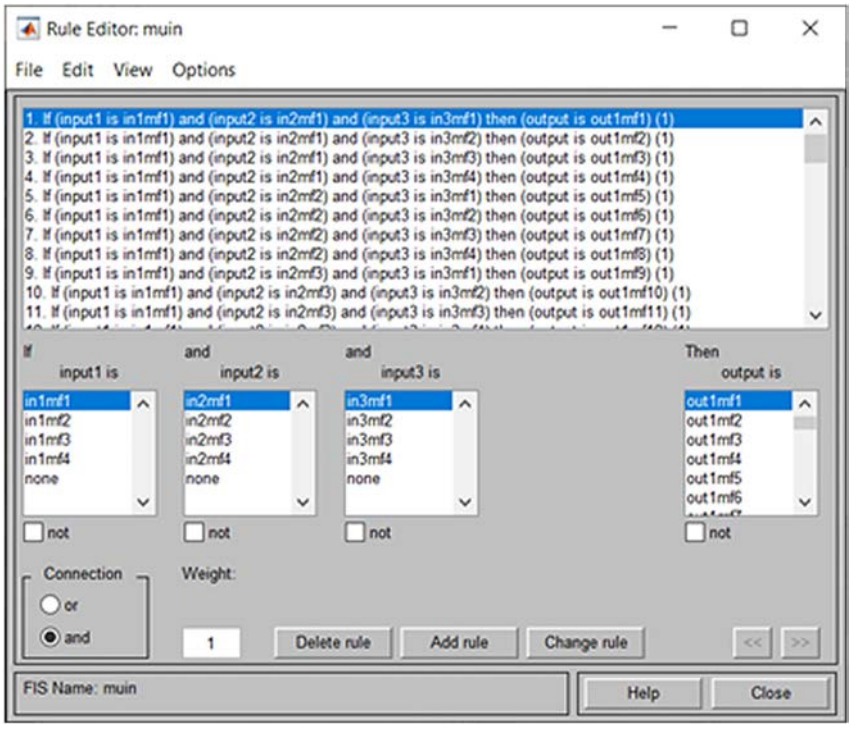

Рис. 9. Правила сети ANFIS

Заключение

Была решена задача подготовки данных для реализации метода моделирования УБИ. Проведен анализ известных зарубежных и российских баз данных и знаний угроз безопасности информации, проведен анализ моделей угроз ряда территориально-распределенных ИС. На основе проведенного анализа выявлены недостатки и преимущества существующих методологий. Установлены ключевые особенности архитектуры территориально-распределенных ИС, определены сложности при моделировании УБИ для таких ИС. Подготовлен набор данных и предложена методика определения актуальных УБИ, проведены необходимые работы для автоматизированной обработки набора данных, связанные с конвертацией данных. Были проанализированы нечеткие нейронные сети ANFIS, алгоритмы их работы. Выбрана наилучшая, основанная на алгоритме нечеткого вывода Такаги-Сугено-Канга. Проведены эксперименты, по результатам которых определены наилучшие параметры системы ANFIS, при которых RMSE достигает значения в диапазоне 0,012–0,023, что является локальным минимумом на заданном интервале и позволяет доказать выполнение поставленной в настоящей работе задачи. Разработана программа для ЭВМ, реализующую предложенную методику определения актуальных УБИ на языке программирования Python 3. Полученные результаты можно использовать для определения актуальных УБИ для территориально-распределенных ИС. В качестве продолжения работы по данной тематике могут являться разработка автоматизированной методики определения актуального нарушителя безопасности информации, а также уменьшения ошибки обучения сети.

Литература

1. *Миняев А.А., Будько М.Ю.* Метод оценки эффективности системы защиты информации территориально распределенных информационных систем // Информатизация и связь. 2017. № 3. С. 119–121.

2. *Будько М.Ю., Миняев А.А.* Методика оценки эффективности системы защиты персональных данных информационной системы // Проблема комплексного обеспечения информационной безопасности и совершенствование образовательных технологий подготовки специалистов силовых структур: Межвузовский сборник трудов VI Всероссийской научно-технической конференции (ИКВО НИУ ИТМО, 10 декабря 2015 г.). 2016. С. 43–45.

3. *Osako T.* Proactive Defense model based on Cyber threat analysis // FUJITSU. Sci. Tech. J. 2016. No. 52 (3). Pp. 72–77.

4. *Noor U.* A machine learning-based FinTech cyber threat attribution framework using high-level indicators of compromise // Future Generation Computer Systems. 2019. Pp. 227–242.

5. *Trifonov R.* Artificial intelligence methods for cyber threats intelligence // Int. J. Comput. 2017. Pp. 129–135.

6. *Huang L.* Adaptive strategic cyber defense for advanced persistent threats in critical infrastructure networks // Perform. Eval. Rev. 2018. No. 46 (2). Pp. 52–56.

7. *Ушаков И.А., Котенко И.В., Овраменко А.Ю., Преображенский А.И., Пелёвин Д.В.* Комбинированный подход к обнаружению инсайдеров в компьютерных сетях // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. 2020. № 4. С. 66–71.

8. *Ушаков И. А.* Обнаружение инсайдеров в корпоративной компьютерной сети на основе технологий анализа больших данных // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. 2019. № 4. С. 38–43.

9. *Савинов Н. В., Токарева К. А., Ушаков И. А., Красов А. В., Сахаров Д. В.* Исследование модели сети ЦОД на основе политик Cisco ACI // Защита информации. Инсайд. 2019. № 4 (88). С. 32–43.

10. *Livshitz I. I., Yurkin D. V., Minyaev A. A.* Formation of the Instantaneous Information Security Audit Concept // Communications in Computer and Information Science. 2016. Vol. 678. Pp. 314–324.

11. *Agrawal A., Ahmed C. M., Chang E. Poster.* Physics-Based Attack Detection for an Insider Threat Model in a Cyber-Physical System // In Proceedings of the 2018 on Asia Conference on Computer and Communications Security, 2018. Pp. 821–823.

12. *Mead N., Shull F., Vemuru K., Villadsen O. A.* Hybrid Threat Modeling Method // Software Engineering Institute, Carnegie Mellon University. 2018. URL: http://resources.sei.cmu.edu/library/asset-view.cfm? AssetID=516617 (дата обращения 15.06.2020).

13. *Khan R., McLaughlin K., Laverty D., Sezer Sakir.* STRIDE-based Threat Modeling for Cyber-Physical Systems // In Proceedings of the 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe.

14. *Shevchenko N., Chick T. A., O'Riordan P., Scanlon T. P., Woody C.* Threat modelling: A summary of available methods // Carnegie Mellon University Software Engineering Institute. 2018. Pp. 1–24.

15. *Yue Li, Teng Zhang, Xue Li, Ting Li*. A Model of APT attack Defense Based On Cyber Threat Detection // Communications in Computer and Information Science, Cyber Security,15th International Annual Conference, CNCERT 2018. Pp. 122–134.

16. *Cao X., Gong N. Z.* Mitigating Evasion Attacks to Deep Neural Networks via Region-based Classification // Proceedings of the 2017 Annual Computer Security Applications Conference (ACSAC). ACM, 2017. Pp. 278–287.

17. *Котенко И. В., Ушаков И. А., Пелёвин Д. В., Преображенский А. И., Овраменко А. Ю.* Выявление инсайдеров в корпоративной сети: подход на базе UBA и UEBA // Защита информации. Инсайд. 2019. № 5 (89). С. 26–35.

18. *Mohassel P., Zhang Y.* SecureML: A System for Scalable PrivacyPreserving Machine Learning // Proceedings of the 2017 IEEE Symposium on Security and Privacy (S&P). IEEE, 2017. Pp. 19–38.

19. *Кузнецов И. А., Липатников В. А., Сахаров Д. В.* Управление АСМК организации интегрированной структуры с прогнозированием состояния информационной безопасности // Электросвязь. 2016. № 3. С. 28–36.

20. *Проноза А. А., Виткова Л. А., Чечулин А. А., Котенко И. В., Сахаров Д. В.* Методика выявления каналов распространения информации в социальных сетях // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2018. Т. 14. Вып. 4. С. 362–377.

MODELING INFORMATION SECURITY THREATS IN TERRITORIAL-DISTRIBUTED INFORMATION SYSTEMS

ANDREY A. MINYAEV,
St. Petersburg, Russia, minyaev.a@gmail.com

KEYWORDS: information security threats; distributed information systems; modeling methodology; data science; ANFIS.

ABSTRACT

Introduction: in the design of information security systems importance is the modeling of security threats, which implies the definition of a list of software threats to the information system, on the basis of which decisions are made to neutralize actual threats. Today the number of urgent threats to information security is increasing, due to the complexity of the infrastructure, information processing technologies, and unprotected communication channels. In this regard, the goal is to simulate security threats in geographically distributed information systems. The analysis of foreign and Russian methodologies for threat modeling showed problems associated with a large amount of data for modeling, as well as expert methods. It has been established that to solve the problem posed use machine learning methods, the theory of adaptive fuzzy neural production systems with fuzzy inference algorithms and the use of Data Science technologies when processing large amounts of data. The paper uses such data protection methods as data protection tools, based on fuzzy neural security systems, it is proposed to determine the actual threats to information security. The proposed methodology is automated and hypothetically eliminates expert errors, increases the number of frequently used topical threats to information security, reduces financial costs for the purchase of information security tools, differs in that the process is automated, has low computational complexity, there is no need attracting highly qualified specialists, allows you to determine the list of actual threats in systems of various types and classes, can be adapted to work with databases. The practical significance lies in the automation of the process – the development

of a computer program that implements the proposed methodology. **Discussion:** further research is advisable to continue determining the best parameters of adaptive fuzzy neural production systems and fuzzy inference algorithms.

REFERENCES

1. Minyaev A.A., Budko M. Yu. Method for assessing the effectiveness of the information protection system of geographically distributed information systems. I*nformatization and communication.* 2017. No. 3. Pp. 119–121.

2. Budko M. Yu., Minyaev A.A. Methodology for assessing the effectiveness of the personal data protection system of the information system. *Problema kompleksnogo obespecheniya informacionnoj bezopasnosti i sovershenstvovanie obrazovatel'nyh tehnologij podgotovki specialistov silovyh struktur: Mezhvuzovskij sbornik trudov VI Vserossijskoj nauchno-tehnicheskoj konferencii* [The problem of comprehensive information security and the improvement of educational technologies for training specialists from law enforcement agencies: Interuniversity collection of works of the VI All-Russian scientific and technical conference, IKVO NRU ITMO, December 10, 2015]. 2016. Pp. 43–45.

3. Osako T. Proactive Defense model based on Cyber threat analysis. *FUJITSU. Sci. Tech. J.* 2016. No. 52 (3). Pp. 72–77.

4. Noor U. A machine learning-based FinTech cyber threat attribution framework using high-level indicators of compromise. *Future Generation Computer Systems*. 2019. Pp. 227–242.

5. Trifonov R. Artificial intelligence methods for cyber threats intelligence. *J. Comput*. 2017. Pp. 129–135.

6. Huang L. Adaptive strategic cyber defense for advanced persistent threats in critical infrastructure networks. *Perform. Eval. Rev.* 2018. No. 46 (2). Pp. 52–56.

7. Ushakov I.A., Kotenko I.V., Ovramenko A. Yu., Preobrazhensky A.I., Pelevin D.V. Combined approach to detecting insiders in computer networks. *Bulletin of the St. Petersburg State University of Technology and Design. Series 1: Natural and technical sciences.* 2020. No. 4. Pp. 66–71.

8. Ushakov I.A. Detection of insiders in a corporate computer network based on big data analysis technologies. *Bulletin of the Saint Petersburg State University of Technology and Design. Series 1: Natural and technical sciences*. 2019. No. 4. Pp. 38–43.

9. Savinov N.V., Tokareva K.A., Ushakov I.A., Krasov A.V., Sakharov D.V. Investigation of a data center network model based on Cisco ACI policies. *Information Security. Inside*. 2019. No. 4 (88). Pp. 32–43.

10. Livshitz I.I., Yurkin D.V., Minyaev A.A. Formation of the Instantaneous Information Security Audit Concept. *Communications in Computer and Information Science*. 2016. Vol. 678. Pp. 314–324.

11. Agrawal A., Ahmed C.M., Chang E. Poster. Physics-Based Attack Detection for an Insider Threat Model in a Cyber-Physical System. *In Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, 2018. Pp. 821–823.

12. Mead N., Shull F., Vemuru K., Villadsen O.A. *Hybrid Threat Modeling Method.* Software Engineering Institute, Carnegie Mellon University. 2018. URL: http://resources.sei.cmu.edu/library/asset-view.cfm? AssetID=516617 (date of access 15.06.2020).

13. Khan R., McLaughlin K., Laverty D., Sezer Sakir. STRIDE-based Threat Modeling for Cyber-Physical Systems. *In Proceedings of the 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe*.

14. Shevchenko N., Chick T.A., O'Riordan P., Scanlon T.P., Woody C.. Threat modelling: A summary of available methods. *Carnegie Mellon University Software Engineering Institute*. 2018. Pp. 1–24.

15. Yue Li, Teng Zhang, Xue Li, Ting Li. A Model of APT attack Defense Based On Cyber Threat Detection. *Communications in Computer and Information Science, Cyber Security,15th International Annual Conference*, CNCERT 2018. Pp. 122–134.

16. Cao X., Gong N.Z. Mitigating Evasion Attacks to Deep Neural Networks via Region-based Classification. *Proceedings of the 2017 Annual Computer Security Applications Conference (ACSAC)*. ACM, 2017. Pp. 278–287.

17. Kotenko I.V., Ushakov I.A., Pelevin D.V., Preobrazhensky A.I., Ovramenko A. Yu. Identifying Insiders on the Corporate Network: AUBA and UEBA Approach. *Protection of information. Inside*. 2019. No. 5 (89). Pp. 26–35.

18. Mohassel P., Zhang Y. SecureML: A System for Scalable Privacy Preserving Machine Learning. *Proceedings of the 2017 IEEE Symposium on Security and Privacy (S&P).* IEEE, 2017. Pp. 19–38.

19. Kuznetsov I.A., Lipatnikov V.A., Sakharov D.V. Management of ASMK organization of an integrated structure with forecasting the state of information security. *Electrosvyaz.* 2016. No. 3. Pp. 28–36.

20. Pronoza A.A., Vitkova L.A., Chechulin A.A., Kotenko I.V., Sakharov D.V. Methods for identifying channels of information dissemination in social networks. *Bulletin of St. Petersburg University. Applied math. Informatics. Management processes.* 2018. Vol. 14. Iss. 4. Pp. 362–377.

INFORMATION ABOUT AUTHOR:

Minyaev A.A., Senior Lecturer, The Bonch-Bruevich Saint-Petersburg State University of Telecommunications.

For citation: Minyaev A.A. Modeling information security threats in territorial-distributed information systems. *H&ES Research*. 2021. Vol. 13. No. 2. Pp. 52–65. Doi: 10.36724/2409-5419-2021-13-2-52-65 (In Rus)

Doi: 10.36724/2409-5419-2021-13-2-66-73

ВОПРОСЫ КИБЕРГИГИЕНЫ ПОЛЬЗОВАТЕЛЕЙ И ОПЕРАТОРОВ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ УПРАВЛЕНИЯ ЭЛЕКТРОННОЙ БИБЛИОТЕКОЙ

КРЮКОВА
Елена Сергеевна[1]

МАЛОФЕЕВ
Валерий Александрович[2]

ПАРАЩУК
Игорь Борисович[3]

АННОТАЦИЯ

Введение: проведен обзор в области современных подходов к созданию и совершенствованию информационных систем типа «электронная библиотека». Обобщены и систематизированы цели и задачи электронных библиотек, особенности условий их функционирования. Исследованы роль и место автоматизированной системы управления электронной библиотекой, ее цели, состав и функции. Обоснованы актуальность обеспечения кибербезопасности электронных библиотек и автоматизированных систем управления ими. Сформулированы ключевые понятия в области кибербезопасности электронных библиотек. Показано, что повышение эффективности мероприятий кибербезопасности возможно на пути применения (в комплексе с традиционными средствами защиты) правил и процедур кибергигиены пользователей и операторов автоматизированной системы управления электронной библиотекой. Предметом исследования является кибергигиена, как набор практик кибербезопасности, совокупность методов и мер предосторожности, а также привычек, знаний и навыков защиты, которые позволяют существенно снизить риски киберугроз. **Цель исследования:** целью исследования является анализ существующих и выработка новых направлений теоретических исследований и организационно-практических разработок в области кибергигиены пользователей и системных администраторов электронной библиотеки для обеспечения кибербезопасности как самой библиотеки, так и ее автоматизированной системы управления. **Результаты:** представлен подход к анализу понятий физической сущности кибергигиены, как взаимосвязанного комплекса методов и практических шагов, которые могут и должны быть предприняты для поддержания работоспособности системы и повышения кибербезопасности. Предложены формулировки частных задач кибергигиены и возможные пути их решения, направленные на создание современных средств и методов защиты электронных библиотек. **Практическая значимость:** представленный подход позволяет сформулировать и обосновать направления повышения эффективности кибергигиены, как многоуровневой и многоаспектной практики выполнения повседневных (иногда рутинных) операций по предотвращению киберугроз. Они позволят не только предотвратить серьезные последствия утраты или модификации собираемой, хранимой и обрабатываемой информации, но и осуществить прогнозирование рисков кибербезопасности электронных библиотек.

КЛЮЧЕВЫЕ СЛОВА: электронная библиотека; кибербезопасность; автоматизированная система управления; кибергигиена; ресурс; пользователь; оператор.

Сведения об авторах:
[1]адъюнкт Военной академии связи имени Маршала Советского Союза С.М. Буденного, г. Санкт-Петербург, Россия, e.kkrukovaa69@yandex.ru

[2]курсант Военной академии связи имени Маршала Советского Союза С.М. Буденного, г. Санкт-Петербург, Россия, valeron12.1366@gmail.com

[3]д.т.н., профессор, Заслуженный изобретатель Российской Федерации, профессор Военной академии связи имени Маршала Советского Союза С.М. Буденного, г. Санкт-Петербург, Россия, shchuk@rambler.ru

Для цитирования: Крюкова Е.С., Малофеев В.А., Паращук И.Б. Вопросы кибергигиены пользователей и операторов автоматизированной системы управления электронной библиотекой // Наукоемкие технологии в космических исследованиях Земли. 2021. Т. 13. № 2. С. 66–73. Doi: 10.36724/2409-5419-2021-13-2-66-73

Введение

В современном IT-пространстве страны в целом, и информационной среде системы образования, в частности, все большую роль играют электронные (цифровые) библиотеки.

Электронная библиотека (ЭБ) представляет собой взаимоувязанную по целям и задачам, емкую и мощную информационную систему, предназначенную для организации и хранения упорядоченного фонда электронных объектов и обеспечения доступа к ним с помощью единых средств навигации и поиска [1–3]. Электронные библиотеки предоставляют доступ к огромному объему информационных ресурсов (контенту), работают в широком диапазоне информационно-справочных и информационно-поисковых услуг, а их развертывание не требует значительных временных и материальных затрат. Электронная библиотека способна оказать большую и реальную помощь в информационном обеспечении, как в области подготовки и повышения квалификации профессиональных кадров, так и в области практической разработки технологий, технических систем, подсистем управления и автоматизации различных процессов в промышленности, финансовой сфере, обороне [4]. Они способны обеспечить поиск и управляемый доступ (по информационно-телекоммуникационным сетям) к электронным документам, базам данных, справочным и поисковым системам, а также к иным информационным ресурсам [5].

Важная роль в рамках современных исследований отведена развитию и совершенствованию автоматизированных систем управления (АСУ) ЭБ. Целью развития АСУ ЭБ является как повышение качества ее функционирования, так и достижение иных технико-экономических, социальных и других эффектов, таких, как повышение своевременности и обоснованности решений оператора (администратора) ЭБ, снижению численности административного персонала, повышению качества управления системой в целом [6]. Элементами АСУ ЭБ являются комплексы и средства автоматизации технического функционирования и управления структурой, параметрами и режимами работы ЭБ, комплексы и средства автоматизации комплектования библиотечного фонда, резервного копирования данных ЭБ и восстановления работы ЭБ при сбоях, хранения контрольных экземпляров информационных ресурсов и иные.

Автоматизированная система управления ЭБ нарабатывают на достижение целей создания, развития и функционирования сложных управляемых цифровых систем такого класса. При этом в АСУ ЭБ обеспечивается совместимость между частями библиотеки, а также совместимость со всеми автоматизированными системами, взаимосвязанными с данной ЭБ. Если АСУ ЭБ создана на базе компьютерной сети, то для обеспечения совместимости между элементами такой сети применяют системы протоколов многоуровневого взаимодействия. Важным требованием является возможность (приспособленность) автоматизированной системы в целом и всех видов ее обеспечения к масштабированию (наращиванию), модернизации и развитию. Надежность и адаптивность АСУ ЭБ должны быть достаточными для достижения установленных целей функционирования ЭБ в заданном диапазоне изменений условий применения. Кроме того, в АСУ ЭБ должны быть предусмотрены контроль правильности выполнения автоматизируемых функций и диагностирование с указанием места, вида и причины возникновения нарушений правильности функционирования ЭБ и автоматизированной системы управления ею [7].

Достаточно много современных общих требований к АСУ ЭБ касаются информации: любая поступающая в АСУ ЭБ информация должна быть надежна и достоверна; информация, содержащаяся в базах данных АСУ ЭБ, должна быть актуализирована в соответствии с периодичностью ее использования при выполнении функций системы; АСУ ЭБ должна быть защищена от утечки информации. Система управления ЭБ в необходимых объемах должна в автоматизированном режиме выполнять сбор, обработку и анализ информации (сигналов, сообщений, документов и т.п.) о состоянии объекта управления, выработку управляющих воздействий (программ, планов и т.п.), передачу управляющих воздействий (сигналов, указаний, документов) на исполнение и ее контроль, реализацию и контроль выполнения управляющих воздействий, а также обмен информацией с взаимосвязанными системами [6, 7].

Кибербезопасность и киберустойчивость электронной библиотеки и автоматизированной системы управления электронной библиотекой

Электронная библиотека представляет собой сложную управляемую динамическую систему, которая может быть подвержена воздействию различных негативных, деструктивных факторов. В первую очередь, речь идет о кибербезопасности ЭБ и АСУ ЭБ. Решение задач кибербезопасности занимает ключевое место в комплексе проблем создания и совершенствования электронной библиотеки, это позволит повысить безопасность информационного обеспечения, защищенность объектов (элементов и форда, контента) и субъектов (пользователей и операторов) ЭБ, в конечном итоге, экономя время и деньги.

Электронная библиотека и автоматизированная система управления электронной библиотекой может быть подвергнута попытке внедрения вредоносного программного обеспечения, вводу ложной информации (попытка размещения контента с недостоверными сведениями), попытке влияния на функционирование ЭБ, выводу из строя как элементов цифровой библиотеки, так и системы в целом, применению электромагнитного оружия с целью

уничтожения микроэлектроники аппаратной части ЭБ, а также влиянию иных внешних факторов. Анализ факторов, потенциально влияющих на функционирование такой сложной информационно-поисковой и информационно-справочной системы, как ЭБ говорит о том, что необходимо эффективно структурировать систему, максимально ограничив негативные воздействия на нее, применяя всевозможные средства защиты, как самой системы, так и содержащейся в ней информации.

Все эти процедуры и механизмы составляют основу кибербезопасности ЭБ и АСУ ЭБ.

При этом под кибербезопасностью понимается отрасль технических знаний, отвечающая за эффективное применение взаимоувязанной совокупности способов и средств поддержания устойчивой работы ЭБ и АСУ ЭБ в условиях целенаправленного деструктивного воздействия. Это не только процесс реализации мер по защите подсистем и программных средств (приложений) ЭБ и АСУ ЭБ от компьютерных (цифровых) атак, но и набор инструментов, идей, принципов (концепций и политик) и мер обеспечения безопасности [8, 9].

С учетом того факта, что современные компьютерные (цифровые) атаки, или кибератаки, на системы хранения данных (СХД), иные информационно-справочные и/или информационно-поисковые системы, обычно нацелены на получение доступа к конфиденциальной информации, ее модификации, уничтожению, или на нарушение нормальной работы АСУ ЭБ и ЭБ в целом, создаются условия, разрабатываются политики и меры безопасности, руководства и организационно-технологические подходы к управлению рисками кибербезопасности для защиты ЭБ, АСУ ЭБ и активов пользователей ЭБ. Активы ЭБ, АСУ ЭБ и пользователя ЭБ включают в себя подключенные устройства, персонал, инфраструктуру, приложения, услуги, телекоммуникационные системы и совокупность передаваемой и/или хранимой информации в кибер-среде электронной библиотеки.

Таким образом, кибербезопасность можно также охарактеризовать как совокупность условий, обеспечивающих гарантированную защищенность всех компонент ЭБ, АСУ ЭБ от как можно большего количества киберугроз и негативных, вредоносных воздействий, а также от негативных последствий в случае совершения ошибок, при аварийных ситуациях, повреждениях либо другого вреда в киберпространстве ЭБ, считающегося нежелательным ущербом [10].

Информационная безопасность при этом отличается от кибербезопасности тем, что ИБ предназначена для комплексной безопасности защищаемых информационных ресурсов и данных в любой форме, а кибербезопасность предназначена исключительно для защиты обрабатываемых цифровых данных.

Существенным является то, что кибербезопасность ЭБ и информационная безопасность ЭБ — не аналоги, поскольку кибербезопасность нацелена не на конфиденциальность, целостность и доступность данных в ЭБ и в АСУ ЭБ (как информационная безопасность), а на непрерывность процесса управления этими взаимоувязанными физическими системами.

Но, тем не менее, основную задачу обеспечения кибербезопасности обычно трактуют как применение методов и средств по реализации информационной безопасности для обеспечения устойчивости функционирования ЭБ и АСУ ЭБ в условиях действия целенаправленных киберугроз. А под киберугрозой понимается либо совокупность факторов и условий, создающих опасность нарушения кибербезопасности ЭБ, либо угроза потери данных или нарушения работы ЭБ и АСУ ЭБ в результате кибератаки — компьютерного (цифрового) несанкционированного воздействия на ЭБ и АСУ ЭБ специальными программными средствами с целью нарушения их работы, получения конфиденциальной информации или с иными негативными целями.

Общепризнано, что кибератаки — наиболее эффективный способ воздействия на ЭБ и АСУ ЭБ, поскольку кибератаки могут воздействовать на любом расстоянии, воздействие осуществляется скрытно, изменение функционирования библиотеки и автоматизированной системы управления ею происходит без их разрушения [11, 12].

Киберустойчивость — способность ЭБ и АСУ ЭБ сохранять свое нормальное функционирование в условиях кибератак (компьютерных цифровых атак) [11]. При этом:

– ключевое отличие киберустойчивости от информационной безопасности (как обеспечения конфиденциальности, целостности и доступности данных) состоит в том, что целью является сохранение управления ЭБ в условиях целенаправленных киберугроз, воздействующих на ее систему управления (АСУ ЭБ);

– киберустойчивость характеризует собой интегрированный, обобщенный показатель структуры ЭБ и АСУ ЭБ, избыточности их компонентов и параметров их функционирования;

– для ЭБ различного типа и АСУ ЭБ границы области киберустойчивости могут различаться.

Эффективная кибербезопасность ЭБ и АСУ ЭБ обеспечивается выполнением комплекса мероприятий, представляющего собой многоуровневую защиту, охватывающую сети обмена данными ЭБ, ее программное обеспечение, сами данные (контент), а также компьютеры операторов АСУ ЭБ и автоматизированные рабочие места пользователей, которые необходимо обезопасить. Причем рабочие процессы, технологии, операторы АСУ ЭБ (сотрудники) и пользователи ЭБ должны дополнять друг друга, чтобы обеспечить ее эффективную защиту от кибератак.

Под эффективными рабочими процессами для обеспечения кибербезопасности ЭБ и АСУ ЭБ подразумевается реализация заранее разработанного и принятого в ЭБ набора

основных (базовых) мер по противодействию предпринимаемым и успешно осуществленным атакам. Обычно руководствуются одним надежным набором мер (определяется политикой кибербезопасности), который содержит простые инструкции по определению (идентификации) атак, защите ЭБ и АСУ ЭБ, выявлению киберугроз и противодействию им, а также восстановлению работоспособности ЭБ после осуществленных кибератак [12].

Под эффективными технологиями для обеспечения кибербезопасности ЭБ и АСУ ЭБ подразумевается аппаратно-программные инструменты, необходимые для защиты компонентов ЭБ и АСУ ЭБ (оконечные устройства, например, компьютеры, интеллектуальные устройства и маршрутизаторы, сети и облачная среда ЭБ) от кибератак. К наиболее распространенным технологиям, используемым для защиты перечисленных компонентов, относятся криптографические методы (включая квантовую криптографию), аутентификация пользователей (включая биометрику), межсетевые экраны нового поколения, фильтрация DNS, защита от вредоносного программного обеспечения (ПО), антивирусное ПО, интеллектуальные методы доступа и решения для защиты подсистемы служебной электронной почты для ЭБ [13].

Но особую и важную роль для обеспечения кибербезопасности ЭБ и АСУ ЭБ играют эффективные действия сотрудников и абонентов библиотеки — пользователей и операторов автоматизированной системы управления ею. Понимание и соблюдение сотрудниками и «посетителями» ЭБ основных принципов кибербезопасности принято называть кибергигиеной.

Сущность и направления обеспечения кибергигиены пользователей и операторов автоматизированной системы управления электронной библиотекой

Обеспечение противодействия киберугрозам с точки зрения эффективной деятельности пользователей ЭБ и операторов АСУ ЭБ подразумевает не только их желание, но и обязанность понимать и соблюдать основные принципы кибербезопасности, например, такие как выбор надежных паролей, внимательное отношение к вложениям в электронную корреспонденцию, резервное копирование данных.

Как и любой иной организации, у электронной библиотеки существует ряд определенных правил для пользователей и операторов АСУ ЭБ для использования техники с доступом к информационным ресурсам ЭБ через общедоступную, публичную сеть. К основным правилам относятся:

– запрет на скачивание, открытие и запуск неизвестных файлов;

– использование лицензионного ПО, запрет на скачивание и установку нового ПО (с этой целью часто ставят пароль администратора на рабочие места пользователей ЭБ и операторов АСУ ЭБ);

– запрет на несанкционированное использование съемных цифровых носителей (флеш-накопители, CD-диски, дискеты);

– запрет на открытие электронной корреспонденции с подозрительных и неизвестных адресов;

– закрытый доступ к внутренним файлам ЭБ, запрет использования исключительно системы файлового обмена ЭБ для передачи доступа к ним;

– использование сложных паролей с регулярным их изменением;

– ограниченный доступ пользователей ЭБ к системообразующим информационным ресурсам (основному фонду, контенту), а операторов АСУ ЭБ — к ключевым, критическим процедурам управления. Каждый пользователь и оператор АСУ должен иметь доступ исключительно к информации ЭБ и процедурам управления ею, которые нужны им для работы.

Понятие «кибергигиена» в различных контекстах трактуется по-разному. Например, в приложении к задачам обеспечения кибербезопасности со стороны пользователей ЭБ и операторов АСУ ЭБ можно рассмотреть такие подходы:

– кибергигиена — набор практик кибербезопасности, которыми должны заниматься потребители услуг и администраторы ЭБ, чтобы защитить безопасность и целостность информации на своих устройствах с поддержкой доступа к ЭБ через Интернет от компрометации в результате кибератаки [14];

– кибергигиена — методы и меры предосторожности, которые пользователи ЭБ и операторы АСУ ЭБ принимают с целью обеспечения безопасности и защиты данных от краж и внешних атак [15];

– кибергигиена — справочник по методам и шагам, которые пользователи компьютеров и других устройств на рабочих местах абонентов ЭБ и операторов АСУ ЭБ предпринимают для поддержания работоспособности системы и повышения онлайн-безопасности [16];

– кибергигиена — шаги, которые пользователи компьютеров и других устройств на рабочих местах абонентов ЭБ и операторов АСУ ЭБ могут предпринять, чтобы улучшить свою кибербезопасность и лучше защитить себя при доступе к ресурсам ЭБ через Интернет [17];

– кибергигиена — набор ежедневных привычек, знаний и навыков, которые позволяют существенно снизить риски работы в глобальной сети Интернет.

В ряде современных работ, посвященном анализу практик кибергигиены в различных странах [18–20], отмечается, что кибигигиена должна рассматриваться таким же образом, как и личная гигиена, и после ее правильной интеграции в организацию будет просто организовать рас-

порядок дня, правомерное поведение и периодические осмотры, чтобы убедиться в оптимальном состоянии «информационного здоровья» организации, имеющей доступ в Интернет.

Многогранность аспектов кибергигиены, ее руководящие принципы для лиц, работающих в сфере предоставления информационных услуг, распространяются на защиту от взлома как информации, так и аппаратно-программных элементов ЭБ, доверенных им и находящихся под их контролем. Как и в вопросах обычной личной гигиены, когда исследования и вмешательства в области общественного здравоохранения обычно фокусируются на повышении осведомленности людей, и именно уровень осведомленности общественности является основной мерой вмешательства и оценок, так и вопросах измерения кибергигиены делается акцент на уровне осведомленности о различных действиях людей в области кибербезопасности.

Кибергигиена может и должна решать ряд проблем, например:

– проблему некорректных (неверных) данных — плохая кибергигиена приводит к потере данных, даже если информация не повреждена или не утеряна навсегда, большое число мест для хранения данных, размещение файлов во многих местах становится все более распространенным явлением;

– проблему потери данных — не поддерживающие резервное копирование и сохранение данных современные облачные хранилища, СХД и жесткие диски на рабочих местах пользователей ЭБ и операторов АСУ ЭБ, уязвимы для взлома, сбоев, повреждений и других коллизий, которые могут привести к потере информации;

– проблему устаревшего ПО — устаревшее ПО более уязвимо для кибератак и воздействия вредоносных программ, поэтому программные приложения нуждаются в регулярном обновлении для гарантии того, что последние версии для системы безопасности ЭБ и самые последние версии ПО используются в рамках ЭБ и АСУ ЭБ для всех приложений;

– и наконец, собственно проблему нарушения кибербезопасности — перманентные и непосредственные угрозы (кибератаки, фишинг, хакеры, вредоносные программы, спам, вирусы и другие) всем данным, хранящимся и обрабатываемым в рамках ЭБ, не просто существуют в современной среде угроз, но и постоянно изменяются и совершенствуются.

Очевидно, что кибергигиена не является абсолютной и гарантированной защитой на рабочих местах пользователей ЭБ и операторов АСУ ЭБ, но соблюдение «привычек» кибергигиены поможет снизить риски, позволит закрыть «бреши в обороне», закрыть доступ нарушителям через незапатентованные и устаревшие решения и непредвиденные пробелы в безопасности электронных библиотек. Помимо ежедневных регулярных «привычек» кибергигиены для пользователей и специалистов ЭБ всех уровней, профессионалы советуют осуществить десять простых процедур:

1. Определение раз и навсегда принятой и узаконенной для ЭБ последовательности установки ПО пользователями, с учетом ограничений установки доверенного ПО или вовсе запрета и блокировки установки ПО без предварительного разрешения со стороны администратора АСУ ЭБ;

2. Инвентаризация, учет и текущий аудит всех аппаратно-программных средств и ПО в ЭБ и АСУ ЭБ (включая сети для обмена данными);

3. Обучение субъектов ЭБ (пользователей ЭБ, сетевых администраторов и операторов АСУ ЭБ) правильному поведению в сети (информационной пространстве) ЭБ, включая знание того, какие и как аппаратно-программные средства подключаться к ЭБ, как выявлять признаки кибератаки, как управлять паролями, как идентифицировать потенциальные попытки фишинга и др.;

4. Сведение к минимуму числа субъектов ЭБ с правами администратора;

5. Установка на всех уровнях архитектуры ЭБ стандартных конфигураций (политик) безопасности (например, NIST или CIS Benchmark), которые могут помочь определить такие элементы кибербезопасности, как длина и вид пароля, алгоритмы шифрования, порядок доступа к портам и особенности аутентификации субъектов ЭБ (пользователей, сетевых администраторов и операторов АСУ ЭБ);

6. Периодическое обновление уязвимых приложений и удаление неиспользуемых, с учетом простого правила: «чем больше приложений — тем больше уязвимостей). Регулярный анализ списка приложений в рамках информационного пространства ЭБ и удаление всех любых программ, которые больше не используются, позволяет снизить риски кибератак, сузить диапазон их потенциального негативного воздействия с одновременным «обеззараживанием» устройств электронной библиотеки;

7. Установка регулярного резервного копирования данных (возможен автоматический режим резервного копирования) и надежное сохранение копий в нескольких местах, например, в облачном хранилище ЭБ и на внешнем жестком диске каждого из рабочих мест субъектов ЭБ (пользователей, сетевых администраторов и операторов АСУ ЭБ);

8. Регулярная периодическая проверка статуса и правильности настройки программ безопасности, анализ их способности блокировать актуальные угрозы;

9. Установка запретов (ограничений) в приложениях для обмена данными, разрешение только наиболее необходимого, обязательного обмена данными в настройках аппаратно-программных средств пользователей ЭБ и операторов АСУ ЭБ;

10. Генерация (создание) новых надежных паролей (подключение двухфакторной (2FA) или многофакторной

аутентификации (MFA), а также процедуры регулярного переключения паролей), поскольку управление паролями является важнейшим звеном эффективной стратегии кибергигиены.

Эти простые процедуры и направления реализации кибергигиены обязательны, но не отменяют стратегических задач — создание общей политики ЭБ в области кибергигиены и разработку комплексных процедур кибергигиены пользователей ЭБ и операторов АСУ ЭБ. При грамотной постоянной реализации этих процедур в сочетании с надежными, общеорганизационными практиками кибербезопасности, правильные практики кибергигиены помогают поддерживать «здоровое» состояние кибербезопасности для современных электронных библиотек.

Заключение

Таким образом, кибергигиена является важнейшей составной частью политики кибербезопасности ЭБ, это многоуровневая и многоаспектная практика выполнения повседневных (иногда рутинных) операций, которые должны быть, безусловно, реализованы потребителями услуг и администраторами ЭБ, чтобы защитить безопасность и целостность информации на своих устройствах с поддержкой доступа к ресурсам (контенту) ЭБ через глобальную сеть Интернет. Это набор обязательных к исполнению пользователями ЭБ и операторами АСУ ЭБ способов, методов и мер предосторожности, применение которых поможет существенно повысить кибербезопасность электронных библиотек.

«Рутина» процедур кибергигиены обладает рядом безусловных преимуществ, не только с точки зрения кибербезопасности, но и с точки зрения технического обслуживания ЭБ. Например, с точки зрения рутинных процедур технического обслуживания рабочих мест пользователей ЭБ и операторов АСУ ЭБ, могут быть выявлены многие из проблем безопасности на ранних этапах и появляется возможность предотвратить возникновение серьезных проблем, поскольку хорошо обслуживаемая система менее подвержена рискам кибербезопасности.

Процедуры и повседневная практика кибергигиены не только нарабатывают на прогнозирование угроз, но и позволяют избавиться от фрагментированных файлов и устаревшего ПО, а это снижает риск уязвимостей, уменьшают шансы хакеров, похитителей личных данных, современных вирусов и интеллектуальных вредоносных программ.

Создание современных методов и средств обеспечения кибербезопасности современных сложных информационно-справочных и информационно-поисковых систем подразумевает включение практик кибергигиены, приучение субъектов ЭБ к ее рутинным операциям. Обучение субъектов ЭБ регулярному мониторингу кибербезопасности существенно увеличивает шансы избежать киберугроз, но, как и любая привычка, которую человек хочет сохранить, кибергигиена требует цикличности, рутины и повторения. Например, все пользователи ЭБ и операторы АСУ ЭБ приучены изменять пароли каждые две недели, сканировать на наличие вирусов с помощью антивирусного ПО каждый день, обновлять операционную систему (с разрешения администратора) раз в полгода, очищать жесткий диск и проверять обновления не реже одного раза в неделю и так далее. Повседневная привычка к выполнению рутинных процедур кибергигиены должна и будет постепенно становится естественной ежедневной задачей для пользователей ЭБ и операторов АСУ ЭБ, позволит предотвратить серьезные последствия утраты или модификации информации, собираемой, хранимой и обрабатываемой в электронной библиотеке.

Литература

1. *Богданова И.Ф., Богданова Н.Ф.* Электронные библиотеки: история и современность // Информационное общество: образование, наука, культура и технологии будущего. 2017. Вып. 1. С. 133–154.

2. *Авдеева Н.В., Сусь И.В.* Национальные электронные библиотеки разных стран: реальность и перспективы // Информационные ресурсы России. 2016. № 2. С. 15–19.

3. Шварцман М.Е. Шесть шагов в продвижении электронных библиотек // Российская ассоциация электронных библиотек. URL: http://www.aselibrary.ru/blogs/archives/1664/ (дата обращения: 21.05.2020).

4. *Паращук И.Б., Ренсков А.А.* Особенности технического обеспечения и этапов процесса создания единой электронной библиотеки в рамках системы «электронного вуза» // Труды III Межвузовской научно-практической конференции «Проблемы технического обеспечения войск в современных условиях» (Санкт-Петербург, 16 февраля 2018). Санкт-Петербург, 2018. Т. 1. С. 347–351.

5. *Yang S. Q., Li L.* Emerging Technologies for Librarians: A Practical Approach to Innovation. Elsevier, Chandos Publishing, 2016. URL: https://www.elsevier.com/books/emerging-technologies-for-librarians/yang/978–1–84334–788–0 (дата обращения: 21.05.2020).

6. *Ермолаева В.В., Калашников Д.А.* Автоматизированные системы управления // Молодой ученый. 2016. № 11. С. 166–168.

7. *Hernandez-Guzman V.M., Silva-Ortigoza R.* Automatic Control with Experiments. Cham: (Switzerland). Springer Nature Switzerland AG. 2019. 996 p.

8. *Гайфулина Д.А., Котенко И.В.* Применение методов глубокого обучения в задачах кибербезопасности. Часть 1 // Вопросы кибербезопасности. 2020. № 3. С. 76–86.

9. *Almeida V., Doneda D., de Souza A.* Cyberwarfare and digital governance // IEEE Internet Computing. 2017. No. 21 (2). Pp. 68–71.

10. *Brooks C.J., Grow C., Craig P.R., Short D.* Cybersecurity Essentials. Indianapolis: John Wiley & Sons Inc. 2018. 767 p.

11. *Авраменко В.С., Маликов А.В.* Подход к аутентификации пользователей инфокоммуникационных систем с применением машинного обучения // Сборник научных статей IX Международной научно-технической и научно-методической

конференции «Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2020)». Санкт-Петербург, 2020. С. 13–17.

12. *Михайличенко Н.В.* Проблемы и перспективы обеспечения безопасности центров обработки данных // Региональная информатика и информационная безопасность. 2017. Вып. 4. С. 137–138.

13. *Савенкова Д.Д.* Кибербезопасность финансово-кредитных организаций в условиях новых вызовов и угроз в цифровом пространстве // Право и государство: теория и практика. 2018. № 4. С. 126–131.

14. *Козырева А.А.* Определение термина «кибергигиена» и возможность его применения в правовом поле // Пробелы в российском законодательстве. 2018. № 7. С. 92–94.

15. *Karpenko O.* Formation and realization of participative culture in the conditions of the digital society development: communicative governance and cyber hygiene // National Academy of Managerial Staff of Culture & Arts Herald. 2019. Iss. 2. Pp. 74–76.

16. *Trevors M., Wallen C.M.* Cyber Hygiene: A Baseline Set of Practices // Carnegie Mellon University. Software Engineering Institute. URL: https://resources.sei.cmu.edu/asset_files/Presentation/2017_017_001_508771.pdf (дата обращения: 22.05.2020).

17. *Maennel K., Mäses S., Maennel O.* Cyber Hygiene: The Big Picture // Nordic Conference of Secure IT Systems (NordSec 2018). Springer Link, 2018. Pp. 291–305. URL: https://link.springer.com/chapter/10.1007/978–3–030–03638–6_18 (дата обращения: 22.05.2020).

18. *Jomin G., Aroma E.* Cyber Hygiene in Health Care Data Breaches // International Journal of Privacy and Health Information Mana-gement (IJPHIM). 2018. No. 6(1). URL: https://www.igi-global.com/article/cyber-hygiene-in-health-care-data-breaches/202466 (дата обращения: 22.05.2020).

19. *Turpin J.R.* Practicing Good Cyber Hygiene // The NEWS. 2019. URL: https://www.achrnews.com/ articles/140446-practicing-good-cyber-hygiene (дата обращения: 22.05.2020).

20. *Brent D., Washington D.* Cyber Hygiene for Physical Security // DOCPLAYER. URL: https://docplayer.net/18566512-Cyber-hygiene-for-physical-security.html (дата обращения: 22.05.2020).

QUESTIONS OF CYBER HYGIENE FOR USERS AND OPERATORS OF THE AUTOMATED MANAGEMENT SYSTEM OF THE ELECTRONIC LIBRARY

ELENA S. KRYUKOVA

St. Petersburg. Russia, e.kkrukovaa69@yandex.ru

VALERY A. MALOFEEV

St. Petersburg. Russia, valeron12.1366@gmail.com

IGOR B. PARASHCHUK

St. Petersburg. Russia, shchuk@rambler.ru

KEYWORDS: electronic library; cybersecurity; automated management system; cyber hygiene; resource; user; operator.

ABSTRACT

Introduction: a review of modern approaches to the creation and improvement of information systems such as "electronic library" is conducted. The goals and objectives of electronic libraries, the peculiarities of their functioning conditions are summarized and systematized. The role and place of the automated electronic library management system, its goals, composition and functions are investigated. The relevance of ensuring the cybersecurity of electronic libraries and automated management systems is justified. The key concepts in the field of cybersecurity of electronic libraries are formulated. It is shown that it is possible to increase the effectiveness of cybersecurity measures by applying (in combination with traditional means of protection) the rules and procedures of cyber-hygiene of users and operators of an automated electronic library management system. The subject of the study is cyber hygiene, as a set of cybersecurity practices, a set of methods and precautions, as well as habits, knowledge and skills of protection that can significantly reduce the risks of cyber threats. **Purpose:** the purpose of the research is to analyze the existing and develop new areas of theoretical research and organizational and practical developments in the field of cyber hygiene of users and system administrators of the electronic library to ensure the cybersecurity of both the library itself and its automated management system. **Results:** the paper presents an approach to the analysis of the concepts of the physical essence of cyber hygiene, as an interconnected set of methods and practical steps that can and should be taken to maintain the system's performance and improve cybersecurity. The formulations of particular tasks of cyber-hygiene and possible ways of their solution, aimed at creating modern means and methods of protection of electronic libraries, are proposed. **Practical relevance:** the presented approach allows us to formulate and justify the directions for improving the effectiveness of cyber hygiene, as a multi-level

and multidimensional practice of performing everyday (sometimes routine) operations to prevent cyber threats. They will allow not only to prevent serious consequences of loss or modification of the collected, stored and processed information, but also to predict the risks of cybersecurity of electronic libraries.

REFERENCES

1. Bogdanova I.F., Bogdanova N.F. Digital Libraries: History and Present. *Informacionnoe obshhestvo: obrazovanie, nauka, kul`tura i texnologii budushhego* [Information Society: Education, Science, Culture and Future Technologies]. 2017. Issue 1. Pp. 133–154. (In Rus)
2. Avdeeva N.V., Sus I.V. National electronic libraries of different countries: reality and prospects. *Informacionny`e resursy` Rossii* [Information resources of Russia]. 2016. No. 2. Pp.15–19. (In Rus)
3. Schwartzman M.E. Shest` shagov v prodvizhenii e`lektronny`x bibliotek [Six steps in promoting electronic libraries]. *Rossijskaya associaciya e`lektronny`x bibliotek* [Russian Association of Digital Libraries] URL: http://www.aselibrary.ru/blogs/archives/1664/ (date of access 21.05.2020). (In Rus)
4. Parashchuk I.B., Renskov A.A. Osobennosti texnicheskogo obespecheniya i e`tapov processa sozdaniya edinoj e`lektronnoj biblioteki v ramkax sistemy` "e`lektronnogo vuza" [Features of technical support and stages of the process of creating a unified electronic library within the framework of the "electronic university"]. *III Mezhvuzovskaya nauchno-prakticheskaya konferenciya "Problemy` texnicheskogo obespecheniya vojsk v sovremenny`x usloviyax"*[Proceedings of the III Interuniversity scientific-practical conference "Problems of technical support of troops in modern conditions." Conference proceedings, St. Petersburg, February 16, 2018]. St. Petersburg, 2018. Vol. 1. Pp. 347–351. (In Rus)
5. Yang S. Q., Li L. *Emerging Technologies for Librarians: A Practical Approach to Innovation.* Elsevier, Chandos Publishing, 2016. URL: https://www.elsevier.com/books/emerging-technologies-for-librarians/yang/978–1–84334–788–0 (date of access 21.05.2020).
6. Ermolaeva V.V., Kalashnikov D.A. Automated control systems. *Molodoj ucheny`j* [Young scientist]. 2016. No. 11. Pp. 166–168. (In Rus)
7. Hernandez-Guzman V.M., Silva-Ortigoza R. *Automatic Control with Experiments.* Cham: (Switzerland). Springer Nature Switzerland AG. 2019. 996 p.
8. Gaifulina D.A., Kotenko I.V. Application of deep learning methods in cybersecurity tasks. Part 1. *Voprosy` kiberbezopasnosti* [Cybersecurity issues]. 2020. No. 3. Pp. 76–86 (in Rus)
9. Almeida V., Doneda D., de Souza A. Cyberwarfare and digital governance. *IEEE Internet Computing*. 2017. No. 21 (2). Pp. 68–71.
10. Brooks C.J., Grow C., Craig P.R., Short D. *Cybersecurity Essentials*. Indianapolis: John Wiley & Sons Inc. 2018. 767 p.
11. Avramenko V.S., Malikov A.V. Approach to authentication of users of infocommunication systems using machine learning. *Aktual'nye problemy infotelekommunikacij v nauke i obrazovanii (APINO 2020). Sbornik nauchnyh statej IX Mezhdunarodnoj nauchno-tekhnicheskoj i nauchno-metodicheskoj konferencii* [In the collection: Actual problems of infotelecommunications in science and education (APINO 2020). Collection of scientific articles of the IX International Scientific-technical and scientific-methodological Conference]. St. Petersburg: 2020. Pp. 13–17. (In Rus)
12. Mikhaylichenko N.V. Problems and prospects for ensuring the security of data centers. *Regional`naya informatika i informacionnaya bezopasnost`* [Regional informatics and information security]. 2017. Iss. 4. Pp. 137–138. (In Rus)
13. Savenkova D.D. Cybersecurity of financial and credit organizations in the face of new challenges and threats in the digital space. *Pravo i gosudarstvo: teoriya i praktika* [Law and State: Theory and Practice]. 2018. No. 4. Pp. 126–131. (In Rus)
14. Kozyreva A.A. Definition of the term "cyber hygiene" and the possibility of its application in the legal field. *Probely` v rossijskom zakonodatel`stve* [Gaps in Russian law]. 2018. No. 7. Pp. 92–94. (In Rus)
15. Karpenko O. Formation and implementation of participative culture in the conditions of the digital society development: communicative governance and cyber hygiene. *National Academy of Managerial Staff of Culture & Arts Herald*. 2019. Iss. 2. Pp. 74–76.
16. Trevors M., Wallen C.M. Cyber Hygiene: A Baseline Set of Practices. *Carnegie Mellon University. Software Engineering Institute*. URL: https://resources.sei.cmu.edu/asset_files/Presentation/2017_017_001_508771.pdf (date of access 22.05.2020).
17. Maennel K., Mäses S., Maennel O. Cyber Hygiene: The Big Picture. *Nordic Conference of Secure IT Systems (NordSec 2018)*. Springer Link. 2018. Pp. 291–305. URL: https://link.springer.com/chapter/10.1007/978–3–030–03638–6_18 (date of access 22.05.2020).
18. Jomin G., Aroma E. Cyber Hygiene in Health Care Data Breaches. I*nternational Journal of Privacy and Health Information Management (IJPHIM)*. 2018. No. 6(1). URL: https://www.igi-global.com/article/cyber-hygiene-in-health-care-data-breaches/202466 (date of access 22.05.2020).
19. Turpin J.R. Practicing Good Cyber Hygiene. *The NEWS*. 2019. URL: https://www.achrnews.com/articles/140446-practicing-good-cyber-hygiene (date of access 22.05.2020).
20. Brent D., Washington D. Cyber Hygiene for Physical Security. *DOCPLAYER*. URL: https://docplayer.net /18566512-Cyber-hygiene-for-physical-security.html (date of access 22.05.2020).

INFORMATION ABOUT AUTHORS:

Kryukova E.S., Postgraduate student of the Military Telecommunication Academy;
Malofeev V.A., Cadet of the Military Telecommunication Academy;
Parashchuk I.B., PhD, Full Professor, Professor of the Military Telecommunication Academy.

For citation: Kryukova E.S., Malofeev V.A., Parashchuk I.B. Questions of cyber hygiene for users and operators of the automated management system of the electronic library. *H&ES Research*. 2021. Vol. 13. No. 2. Pp. 66–73. Doi: 10.36724/2409-5419-2021-13-2-66-73 (In Rus)

Doi: 10.36724/2409-5419-2021-13-2-74-84

БИНАРНАЯ КЛАССИФИКАЦИЯ МНОГОАТРИБУТНЫХ РАЗМЕЧЕННЫХ АНОМАЛЬНЫХ СОБЫТИЙ КОМПЬЮТЕРНЫХ СИСТЕМ С ПОМОЩЬЮ АЛГОРИТМА SVDD

ШЕЛУХИН
Олег Иванович[1]

РАКОВСКИЙ
Дмитрий Игоревич[2]

АННОТАЦИЯ

Введение: в настоящее время объем системных журналов компьютерных систем, объединенных в распределенную сетевую инфраструктуру, делает невозможным их ручную проверку в режиме реального времени. Как правило, структура каждой записи журнала содержит численное значение наблюдаемого атрибута и соответствующую пометку (маркер), помечающее запись как нормальное или аномальное. Алгоритм описания данных опорными векторами демонстрирует высокую точность классификации уже при малых объемах обучающей выборки. Особенностью алгоритма является работа с многоатрибутным набором данных, где каждое наблюдение содержит общую классифицирующую маркировку. Следовательно, возникает задача о сведении маркировок атрибутов исходных данных к единой маркировке всего наблюдения. **Цель исследования:** исследование точности бинарной классификации экспериментальных данных алгоритмом описания данных опорными векторами при малом объеме обучающей выборки для случая поатрибутно маркированных экспериментальных данных. **Методы исследования:** предложен метод для решения задачи о сведении маркировок атрибутов исходных данных к единой маркировке посредством подходов «полностью нормальное наблюдение» и голосования по мажоритарному принципу. Рассмотрены два вида данных: упорядоченные во времени и равномерно перемешанные. Точность классификации оценена при помощи вычисления площади под ROC-кривыми с проведением кросс-валидации при разном количестве атрибутов. **Результаты:** сравнительный анализ способов маркировки наблюдений показал преимущество подхода «полностью нормальное наблюдение» перед подходом «мажоритарное голосование» без «взвешивания». Показано, что точность классификации на перемешанных данных выше на 7% по сравнению с вариантом упорядочивания данных во времени. Исследована точность алгоритма при различном количестве атрибутов с использованием подхода «полностью нормальное наблюдение». Максимально достигнутая точность классификации составила порядка 96% при работе с 6 атрибутами, при равномерном перемешивании входного набора данных. Дальнейшее увеличение количества атрибутов приводит к снижению средней точности классификации по причине роста доли аномальных наблюдений. Показано, что при использовании равномерного перемешивания входных данных выигрыш по точности может быть увеличен на 15–20%. **Практическая значимость:** алгоритм демонстрирует экспоненциальный рост потребления вычислительных ресурсов при увеличении объема входных данных. **Обсуждение:** для достижения максимальной точности классификации при приемлемом потреблении ресурсов необходимо сформировать компактный набор входных данных, наиболее полно отражающий функционирование компьютерной системы в нормальном режиме.

КЛЮЧЕВЫЕ СЛОВА: разметка данных; перемешивание данных; полностью нормальное наблюдение; голосование по мажоритарному принципу; малая обучающая выборка; малые данные.

Сведения об авторах:
[1]д.т.н., профессор, заведующий кафедрой Московского технического университета связи и информатики, г. Москва, Россия, sheluhin@mail.ru

магистрант кафедры «Информационная безопасность» Московского технического Университета связи и информатики, г. Москва, Россия, dimitor1998@mail.ru

Для цитирования: *Шелухин О.И., Раковский Д.И.* Бинарная классификация многоатрибутных размеченных аномальных событий компьютерных систем с помощью алгоритма SVDD // Наукоемкие технологии в космических исследованиях Земли. 2021. Т. 13. № 2. С. 74–84. Doi: 10.36724/2409-5419-2021-13-2-74-84

Введение

При управлении крупномасштабной сетевой инфраструктурой важную роль играет обнаружение аномалий посредством анализа системных журналов, в которые заносится информация о времени работы системы и о событиях, происходящих в ней. Как правило, обслуживающий персонал (разработчики системы или операторы) проверяют системные журналы вручную. Поиск аномалий может быть осуществлен при помощи поиска: либо по ключевым словам, либо с применением соответствующих правил. Тем не менее, объем записей в системных журналах растет пропорционально масштабу и сложности современных сетевых инфраструктур, что затрудняет или делает невозможным их ручную проверку.

Исследования проводились на сетевой инфраструктуре, состоящей из 6 хостов, образующих кластер под управлением Rancher (рис. 1). В результате исследований были получены поатрибутно размеченные экспериментальные данные, снятые с локальной компьютерной системы (КС), состоящей из нескольких серверных и клиентских хостов, в период с 2019–09–24 05:35:06 по 2019–09–26 23:39:00 с шагом 1 с. Каждому атрибуту присвоена одна из двух маркировок — «нормальное значение атрибута» и «аномальное значение атрибута». Структура каждой записи (наблюдения) в наборе экспериментальных данных представляет собой 68 чисел, разделенных на 2 группы: 34 численных значения по каждому из атрибутов и столько же бинарных маркировок, относящих соответствующий атрибут либо к нормальному, либо к аномальному состоянию [1]. Набор экспериментальных данных сведен в csv-таблицу.

В качестве иллюстрации, на рис. 2 приведен пример записи из таблицы экспериментальных данных. Для большей наглядности бинарные маркировки выделены курсивом. Численные значения, находящиеся в аномальном состоянии и соответствующие им бинарные маркировки выделены полужирным начертанием. Из рисунка следует, что в рассматриваемый момент времени 7 атрибутов находились в состоянии «аномальное значение атрибута».

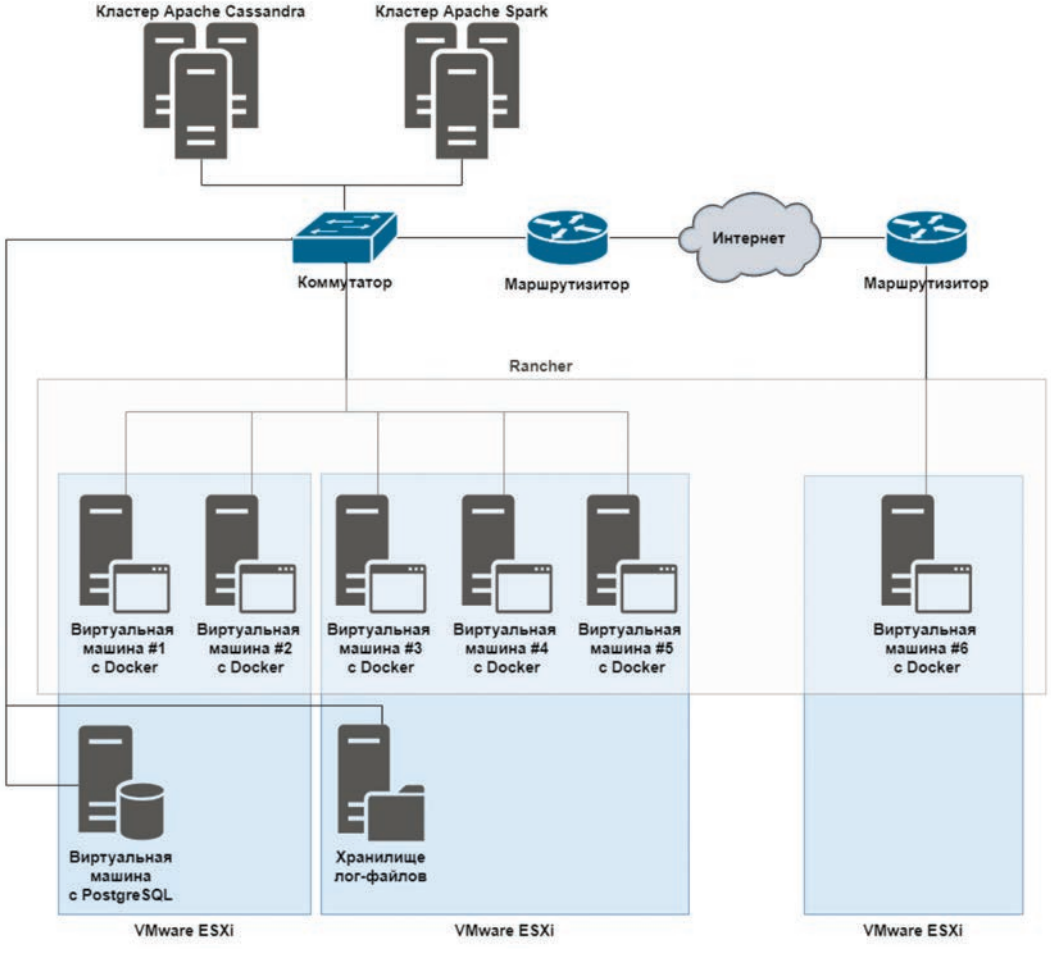

Рис. 1. Схема исследуемой сетевой инфраструктуры

0.0109,0.0006666666666666666,0.016833333333333332,0.3548,**1.4250833333333333**,42732146688.0,8364023808.0,

0.0,0.4,0.4,0.0,0.0,4257.2,0.4,0.0,11.0,6235489.2,1.65,1.74,0.69,**139734635834.0,513290302.0,139734635834.0**,0.0,0.0,

58.0,120.0,137.0,**4.803999999999999,7.572**,2.036,0.004,1.2670000000000001,1.268,0.0,0.0,0.0,0.0,0.0,0.0,**1.0,1.0**,0.0,0.0,

0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,**1.0,1.0,1.0**,0.0,0.0,0.0,0.0,0.0,0.0,**1.0,1.0**,0.0,0.0,0.0,0.0,0.0

Рис. 2. Пример записи экспериментальных данных вместе с маркировками

Как правило, крупномасштабные сетевые структуры являются системами реального времени, на которые накладываются соответствующие ограничения [2–4]: быстродействие (обеспечение времени выполнения задач), предсказуемость (поведение системы должно быть прогнозируемо), учет максимально допустимого времени отклика на события и пр. Учитывая эти ограничения, целесообразно подойти к анализу поведения подобных последовательностей как к задаче многоатрибутой бинарной классификации при помощи методов интеллектуального анализа данных.

Широкое распространение в задачах бинарной классификации получило семейство алгоритмов Support Vector Machine (SVM), демонстрирующее хорошие результаты при малых объемах обучающей выборки [5, 6]. Достигается такое соотношение за счет представления набора данных как точек в многомерном пространстве, где каждая координата точки представляет собой числовое значение соответствующего атрибута. Тогда задача классификации сводится к нахождению уравнения разделяющей гиперплоскости, разделяющей пространство оптимальным образом. Темпоральность данных, как правило, не учитывается. Соответственно, дополнительный выигрыш в точности классификации может быть достигнут за счет равномерного перемешивания набора данных.

Алгоритм описания данных опорными векторами[1] (SVDD, Support Vector Data Description) позволяет достичь еще меньшего отношения обучающей выборки к набору данных при сохранении приемлемых результатов. Больший выигрыш в точности по сравнению с алгоритмами семейства SVM достигается за счет использования разделяющей гиперсферы вместо гиперплоскости. Использование гиперсферы позволяет работать с данными, имеющими сложную структуру в пространстве атрибутов [7–8]. Алгоритм демонстрирует лучшие результаты по точности в некоторых случаях слабо разделимых данных [9]. Данный алгоритм широко используется при обнаружении аномалий во многих прикладных областях: в медицине [10], в тяжелой промышленности [11], при биометрической аутентификации [12]. В большинстве случаев алгоритм SVDD демонстрирует лучшие результаты по достигнутой точности бинарной классификации по сравнению с другими методами классификации.

Целью работы является исследование точности бинарной классификации экспериментальных данных алгоритмом SVDD при малом объеме обучающей выборки для случая поатрибутно маркированных экспериментальных данных.

Структура алгоритма SVDD

Общая постановка задачи классификации заключается в следующем[2]. Пусть дан набор из m неклассифицированных данных $S_m = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_m\}, \boldsymbol{x}_i \in \mathbb{R}^p$, где $\boldsymbol{x}_i$ — i-й элемент в наборе данных (наблюдение), являющийся p-мерной величиной, отражающей это наблюдение. Размерность определяется количеством атрибутов, по которым ведется наблюдение.

В данный набор могут входить нормальные и аномальные элементы. Тогда задачей классификации $f : \boldsymbol{x} \to \{-1; +1\}$ будет являться задача классификации i-того элемента последовательности S_m по двум классам — нормальному и аномальному. После выполнения операции классификации, каждый $\boldsymbol{x}_i$ — i-й элемент будет иметь бинарную маркировку, где $f(\boldsymbol{x}) = +1$ означает, что наблюдение $\boldsymbol{x}_i$ является элементом нормального класса, а $f(\boldsymbol{x}) = -1$ — является элементом аномального класса. Соответствующий маркированный набор из m данных, полученный в результате действия классификатора f над набором немаркированных данных S_m возможно представить, как:

$$S_m = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^m, \boldsymbol{x}_i \in \mathbb{R}^p, y_i \in \{-1; +1\}, \quad (1)$$

где $(\boldsymbol{x}_i, y_i)$ — i-тый элемент набора данных, в котором $\boldsymbol{x}_i$ является i-тым p-мерным наблюдением в наборе данных, y_i — это бинарная маркировка данного наблюдения, относящая соответствующее наблюдение к нормальному ($y_i \in +1$) или аномальному ($y_i \in -1$) классу.

Алгоритму классификации SVDD для работы с неклассифицированными данными необходимо проанализировать некоторый объем маркированных данных, при

[1] Tax D.M.J. and Duin R.P.W. Support Vector Data Description // Machine Learning, vol. 54. 2004. pp. 45–66. doi:10.1023/B: MACH.0000008084.60811.49

[2] Hastie T., Tibshirani R., Friedman J. The elements of statistical learning: data mining, inference and prediction. Springer, 2nd edition. 2009. 745 p.

этом маркировка должна быть произведена в соответствии с (1). Каждое наблюдение в таком наборе данных можно представить, как точку в *p*-мерном пространстве. Тогда вокруг данных нормального класса может быть описана гиперсфера с центром $\boldsymbol{a}$ и радиусом $R>0$:

$$F(R,\boldsymbol{a}) = R^2 \tag{2}$$

При этом не все объекты нормального класса должны находиться внутри гиперсферы. Допущение о возможности выбросов данных нормального класса за границу гиперсферы усложняет нахождение радиуса гиперсферы. С этой целью в формулу (2) добавляются дополнительные переменные ξ_i, что превращает отыскание радиуса гиперсферы в задачу двойственной оптимизации, которая может быть решена с помощью нахождения соответствующих множителей Лагранжа α.

Ненулевыми множители Лагранжа становятся только для тех объектов нормального класса $\boldsymbol{x}_i$, для которых выполняется равенство $\|\boldsymbol{x}_i-\boldsymbol{a}\|^2 = R^2+\xi_i$. Совокупность параметров в виде центра сферы $\boldsymbol{a}$, набора объектов (наблюдений) $\boldsymbol{x}_i$, ненулевых множителей Лагранжа α и образуют гиперсферу. Объекты с ненулевыми множителями Лагранжа называются опорными векторами. Для бинарно маркированных данных решением двойственной оптимизационной задачи является Лагранжиан:

$$\begin{aligned} L = &\sum_i \alpha_i(\boldsymbol{x}_i\cdot\boldsymbol{x}_i) - \sum_l \alpha_l(\boldsymbol{x}_l\cdot\boldsymbol{x}_l) - \sum_{i,l}\alpha_i\alpha_l(\boldsymbol{x}_i\cdot\boldsymbol{x}_l) + \\ &+2\sum_{i,j}\alpha_i\alpha_j(\boldsymbol{x}_i\cdot\boldsymbol{x}_j) - \sum_{l,k}\alpha_l\alpha_k(\boldsymbol{x}_l\cdot\boldsymbol{x}_k) \end{aligned} \tag{3}$$

где $\boldsymbol{x}_i$, $\boldsymbol{x}_j$ — элементы набора данных (данных наблюдения), отнесенные к нормальному классу, а $\boldsymbol{x}_l$, $\boldsymbol{x}_k$ — элементы набора, отнесенные к аномальному классу; α_i, $\alpha_j \geq 0$ — множители Лагранжа, соответствующий нормальным наблюдениям $\boldsymbol{x}_i$, $\boldsymbol{x}_j$; α_l, $\alpha_k \geq 0$ — множители Лагранжа, соответствующие аномальным наблюдениям $\boldsymbol{x}_l$, $\boldsymbol{x}_k$.

Для определения к какому классу отнесен немаркированный объект $\boldsymbol{z}$, необходимо вычислить расстояние от данного объекта до центра сферы:

$$\begin{aligned} \|\boldsymbol{z}-\boldsymbol{a}\|^2 = &(\boldsymbol{z}_i\cdot\boldsymbol{z}_i) - \sum_i \alpha_i(\boldsymbol{x}_i\cdot\boldsymbol{x}_i) - \\ &-2\sum_l \alpha_l(\boldsymbol{z}_l\cdot\boldsymbol{x}_l) + \sum_{i,j}\alpha_i\alpha_j(\boldsymbol{x}_i\cdot\boldsymbol{x}_j) \leq R^2 \end{aligned}$$

Для повышения точности классификации осуществляется преобразование набора исходных данных из евклидова в иные пространства посредством замены скалярного произведения $(\boldsymbol{x}_i\cdot\boldsymbol{x}_j)$ на ядерные функции $\boldsymbol{K}(\boldsymbol{x}_i\cdot\boldsymbol{x}_j) = (\Phi(\boldsymbol{x}_i)\cdot\Phi(\boldsymbol{x}_j))$, например на гауссово ядро [13, 14] или другие функции [15–17].

Иллюстрация решения оптимизационной задачи и нахождение множителей Лагранжа на примере двумерной выборки приведена на рис. 3. На рис. 3а представлена поверхность, заданная двумерным пространством атрибутов и набором радиусов, соответствующих допустимым выбросам данных нормального класса за радиус гиперсферы. Цветом закодирован штраф ξ_i для данных нормального класса, расстояние которых от центра гиперсферы превышает соответствующий радиус. На рис. 3б изображена проекция трехмерной поверхности с нанесенными на нее численными значениями штрафа ξ_i. Цветом закодирован штраф ξ_i для данных нормального класса, расстояние которых от центра гиперсферы превышает соответствующий радиус. На рис. 3в изображен набор данных (кружками — обучающая выборка нормального класса; квадратами — обучающая выборка аномального класса; треугольниками — опорные векторы) и соответствующая им двумерная гиперсфера.

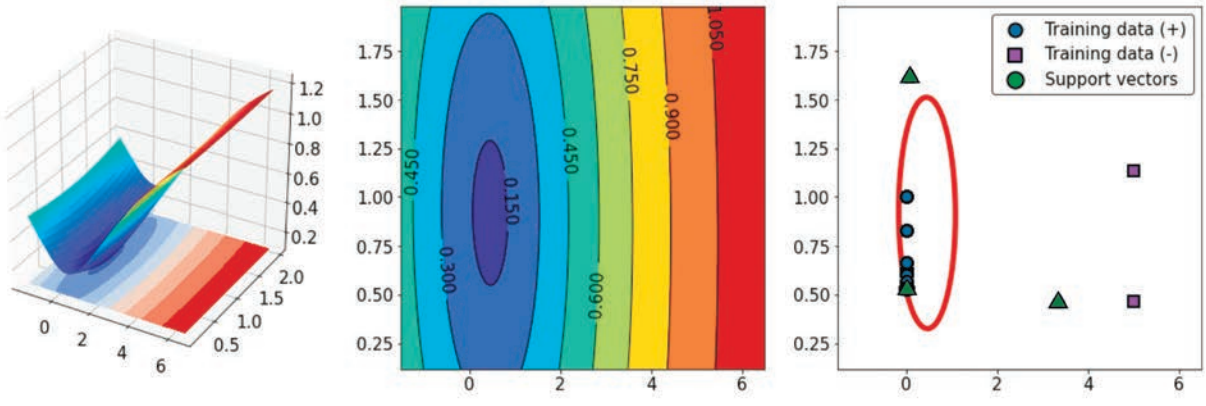

Рис. 3. Решение оптимизационной задачи и нахождение множителей Лагранжа на примере двумерной выборки:
а) поверхность, заданная двумерным пространством данных и допустимыми расстояниями от центра гиперсферы;
б) двумерная проекция поверхности с численными значениями штрафов за расстояния от центра гиперсферы;
в) набор данных и соответствующая им двумерная сфера

Реализация алгоритма SVDD

Рассмотрим наиболее популярный вариант практической реализаций алгоритма SVDD, выполненный на языке Python[3] с использованием в качестве ядерной функции гауссовой ядерной функции.

Результаты потребления временных и процессорных ресурсов, потребляемых алгоритмом для трех атрибутов и двух режимов работы: обучения и тестирования приведены в табл. 1.

Таблица 1

Затраты времени и памяти при одинаковых входных данных

Обучающая выборка		
Размер выборки	Затраты времени, с	Затраты памяти, Гб
1000	0,5	~2,4
2000	3,4	~2,4
3000	11	~2,5
4000	30	~2,6
5000	60	~2,7
Тестовая выборка		
Размер выборки	Затраты времени, с	Затраты памяти, Гб
15 000	4	<1
20 000	10	4
30 000	20	8
40 000	67	18
44 000	170	22-28
50 000	212	26-31

Из представленных данных видно, что наблюдается экспоненциальный рост потребления процессорной мощности для размера обучающей выборки, большей 2000 наблюдений, при этом затраты оперативной памяти остаются на примерно одном уровне. При увеличении размера тестовой выборки наблюдается экспоненциальный рост как по затрачиваемому времени, так и затрату оперативной памяти на обработку и хранение выборки.

Как следует из (1), алгоритм SVDD оперирует мультиатрибутным набором данных, бинарно маркированным по каждому наблюдению. Соответственно, каждому мультиатрибутному наблюдению должна быть сопоставлена одна оценка данного наблюдения, которая выражается в отнесении данного наблюдения либо к нормальному, либо к аномальному классу.

Исходные данные, как правило не обладают маркировкой, присвоенной всему наблюдению, а имеется набор маркировок значений каждого атрибута в каждый момент наблюдения. В результате возникает задача о сведении 34-х маркировок атрибутов исходных данных к единой маркировке всего наблюдения.

[3]Support Vector Data Description (SVDD) // github URL: https://github.com/iqiukp/SVDD (дата обращения: 22.11.2020).

Наиболее простым к определению аномальности всего наблюдения в целом является подход, при котором наблюдение считается «нормальным», если никакой из атрибутов не находится в «аномальном» состоянии. Этот способ назовем «*полностью нормальным наблюдением*».

Альтернативным способом является проведение голосования [18]. Голосование может быть проведено, например, по мажоритарному принципу, при котором наблюдение признается аномальным, если большинство атрибутов находится в «аномальном» состоянии.

Минусом голосования по мажоритарному принципу является одинаковый вес (значимость) каждого атрибута при определении аномальности всего наблюдения. Решением может стать ввод весов или весовых функций [19, 20], соответствующих каждому атрибуту, и их учет при подсчете голосов путем «взвешенного» голосования. В свою очередь, нахождение весовых функций может представлять собой трудоемкую задачу.

В любом случае при работе с алгоритмом SVDD экспериментальные данные необходимо привести к виду (3). Ниже рассматривается два случая: «полностью нормальное наблюдение» и мажоритарное голосование (без «взвешивания»).

Классификация экспериментальных данных при маркировке «полностью нормальное наблюдение»

Структура записи экспериментальных данных подразумевает, что каждому *p*-мерному наблюдению сопоставлена *p*-мерная бинарная маркировка соответствующего наблюдения:

$$S_m = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i=1}^{m}, \boldsymbol{x}_i \in \mathbb{R}^p, \boldsymbol{y}_i \in \{1;0\}^p \qquad (4)$$

где ($\boldsymbol{x}_i$, $\boldsymbol{y}_i$) — *i*-й элемент набора данных, в котором $\boldsymbol{x}_i$ является *i*-м *p*-мерным наблюдением в наборе данных, $\boldsymbol{y}_i$ — *p*-мерная маркировка данного наблюдения (состоящая из последовательности бинарных маркеров $y_{i,1} \in \{1;0\}, y_{i,2} \in \{1;0\}, ..., y_{i,p} \in \{1;0\}$). В соответствии с (4), маркер вида $y_{i,j} \in 0$ относит j-тый атрибут наблюдения $\boldsymbol{x}_i$ к нормальному классу, а маркер вида $y_{i,j} \in 1$ относит *j*-й атрибут наблюдения $\boldsymbol{x}_i$ к аномальному классу.

Для работы алгоритма SVDD необходимо преобразовать *p*-мерную бинарную маркировку $\boldsymbol{y}_i$ в одномерную бинарную y_i, относящую соответствующее наблюдение $\boldsymbol{x}_i$ к нормальному ($y_i \in +1$) или аномальному ($y_i \in -1$) классу, то есть необходимо выполнить преобразование $f : \boldsymbol{y} \to \{-1;+1\}$.

В соответствии с (4), наблюдение должно быть отнесено к нормальному классу при выполнении следующего правила:

$$f(\mathbf{y}_i) = \begin{cases} +1, \sum_{j=1}^{p} y_{i,j} = 0 \\ -1, \sum_{j=1}^{p} y_{i,j} \neq 0 \end{cases}$$

Наблюдение считается принадлежащим нормальному классу ($y_i \in +1$) если сумма бинарных маркеров $\mathbf{x}_i$ наблюдения $\sum_{j=1}^{p} y_{i,j} = 0$. В любом другом случае наблюдение маркируется как аномальное.

Рассмотрим результаты исследования точности классификации алгоритмом SVDD для первых 24000 наблюдений экспериментальных данных по четырем атрибутам КС, представленным в табл 2. Атрибуты выбирались таким образом, чтобы в исследуемом наборе данных содержались записи, относящиеся как к нормальному, так и к аномальному классам. Преобразование экспериментальных данных выполнялось при помощи соотношения (5).

Таблица 2

Описание атрибутов, используемых при исследовании

«cpu_iowait»	Процессорное время, затрачиваемое хостом в ожидании устройства ввода-вывода, выраженное в процентах
«cpu_softirq»	Процессорное время, затрачиваемое хостом на обработку программных прерываний, выраженное в процентах.
«dns_answerscount»	Количество ответов на dns-запрос
«ping_max»	Максимальное время ответа сервера на запрос (при отправке трёх запросов)

Рассматривались два варианта обработки и анализа данных.

При первом варианте данные упорядочивались во времени (вариант 1). При втором варианте данные равномерно перемешивались (вариант 2). Перемешивание данных допустимо, поскольку построение гиперсферы в алгоритме SVDD происходит по отдельным наблюдениям с игнорированием корреляционных связей между ними.

Для упорядоченных во времени данных (вариант 1) точность классификации алгоритмом SVDD для четырех атрибутов по параметру AUC составила ~75%. ROC-кривая для четырех атрибутов при соотношении обучающей выборки к тестовой ~17% изображена на рис. 4а. На рис. 4б изображен график зависимости удаленности точек от центра гиперсферы для каждого из 20 тысяч наблюдений тестовой выборки. По оси абсцисс отложены номера наблюдений, по оси ординат — расстояния каждого наблюдения от центра гиперсферы. Граница гиперсферы изображена в виде горизонтальной линии, нанесенной поверх точек.

Видно, что основная часть наблюдений тестовой выборки сконцентрирована вблизи границы гиперсферы.

Низкая точность классификации обусловлена малым объемом обучающей выборки и временной зависимостью появления аномалий в разных атрибутах. Повышение точности наблюдалось либо за счет увеличения доли обучающей выборки в наборе данных, либо за счет перемешивания набора данных.

Рассмотрим второй вариант обработки данных. С этой целью, перед разделением данных на обучающую и тестовую выборки произведем равномерное перемешивание исследуемого набора.

Эффективность второго варианта перемешивания данных может быть оценена по ROC-кривой алгоритма классификации SVDD для четырех атрибутов при соотношении обучающей выборки к тестовой ~17% приведенной на рис. 5а. Точность классификации по параметру AUC составила до 82%. На рис. 5б изображен график зависимости удаленности точек наблюдений тестовой выборки от центра гиперсферы для каждого из 20 тысяч наблюдений.

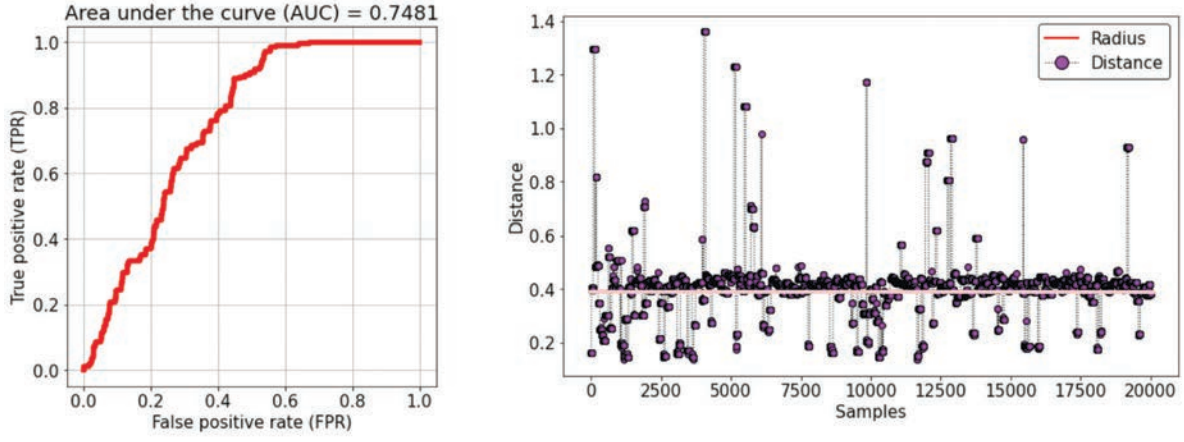

Рис. 4. Эффективность SVDD при первом варианте упорядочивания данных:
а) ROC-кривая; б) зависимость удаленности точек от центра гиперсферы

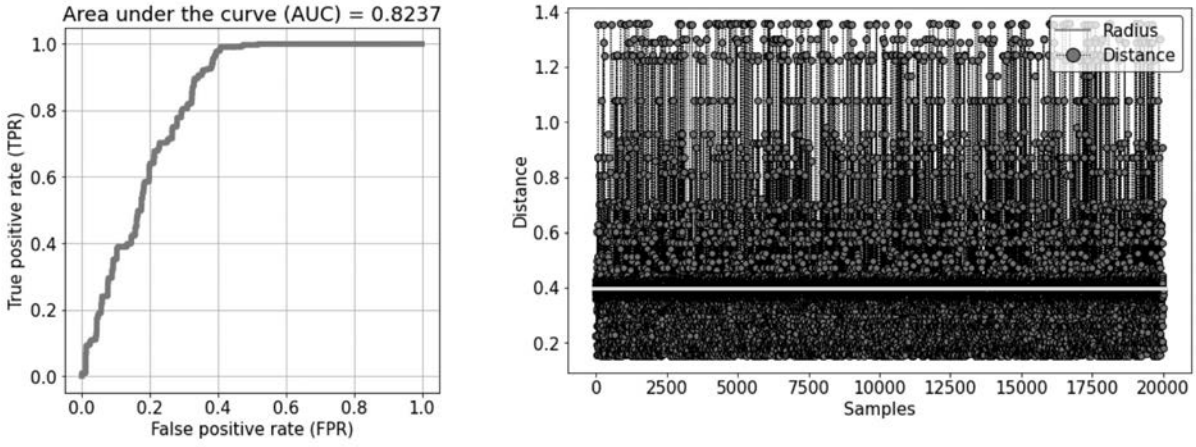

Рис. 5. Эффективность SVDD при втором варианте упорядочивания данных:
а) ROC-кривая; б) зависимость удаленности точек от центра гиперсферы

Граница гиперсферы изображена в виде горизонтальной линии, нанесенной поверх точек.

По сравнению с первым вариантом упорядочивания данных представленным на рис. 4б, основная часть наблюдений, отнесенных к тестовой выборке, находится на удалении от центра гиперсферы в промежутке расстояний (0,2; 0,6), при этом концентрируясь около границы гиперсферы. Кроме того, точность классификации на перемешанных данных выросла на 7% по сравнению с вариантом упорядочивания данных. Дальнейшее повышение точности классификации возможно за счет увеличения доли обучающей выборки в наборе данных.

Классификация экспериментальных данных при маркировке наблюдений вида «мажоритарное голосование»

Рассмотрим преобразование многомерной бинарной маркировки (4) к одномерной величине «голосованием». Голосование может быть проведено по мажоритарному принципу, при котором наблюдение признается аномальным, когда подавляющее большинство атрибутов (>50%) находятся в «аномальном» состоянии:

$$f(\boldsymbol{y}_i) = \begin{cases} +1, \sum_{j=1}^{p} y_{i,j} > 0,5 \\ -1, \sum_{j=1}^{p} y_{i,j} \le 0,5 \end{cases}$$

Проведем исследование точности классификации алгоритмом SVDD для первых 24000 наблюдений экспериментальных данных для четырех атрибутов, представленных в таблице 2. По правилу мажоритарного голосования, выполненного для наблюдений по четырем атрибутам, аномальным признается наблюдение, в котором количество атрибутов, находящихся в состоянии «аномальное значение атрибута» превышает 2.

Рассматривались два описанных выше варианта преобразования экспериментальных данных. На рис.6 пред-

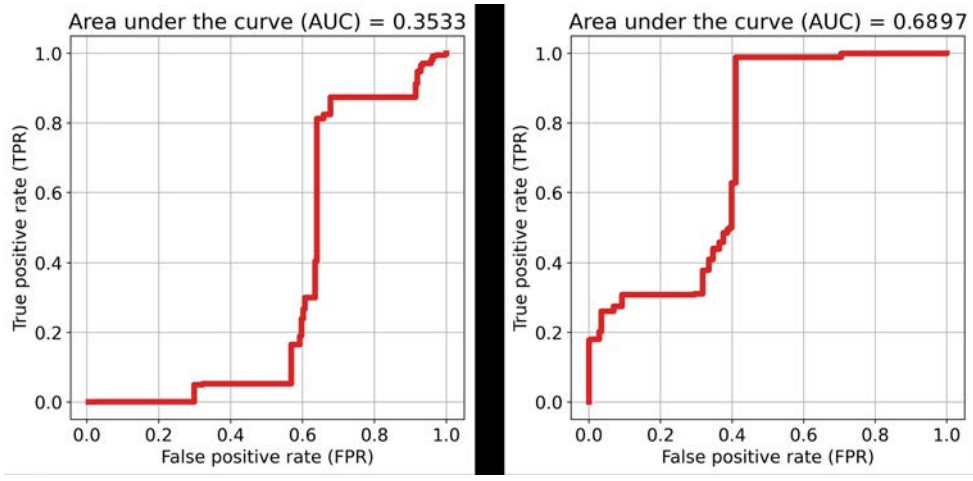

Рис. 6. ROC-кривая алгоритма классификации SVDD при различных вариантах упорядочивания данных:
а) 1-й вариант; б) 2-й вариант

ставлены ROC-кривые алгоритма классификации SVDD при использовании подхода «мажоритарное голосование» к маркировке наблюдений и соотношении обучающей выборки к тестовой ~16%.

Как видно из рисунков, при использовании подхода «мажоритарное голосование» точность классификации по параметру AUC в случае упорядоченных во времени данных составила 35%, а в случае равномерно перемешанных данных — 69%.

Точность классификации при изменении объема экспериментальных данных

Проведем исследование зависимости средней точности классификации анализируемых экспериментальных последовательностей от количества исследуемых атрибутов при фиксированных параметрах: количество атрибутов — 4, 5, 6, 7, 8; размер обучающей выборки — 4000; размер тестовой выборки — 20000 наблюдений. Как и прежде, рассмотрим два варианта представления данных, а классификацию экспериментальных данных будем анализировать при маркировке «полностью нормальное наблюдение».

Результаты точности классификации для различного количества атрибутов с учетом проведенной кросс-валидации представлены в таблице 3. Здесь же приведено время, затрачиваемое на обработку всей выборки во время одной итерации и количество аномальных наблюдений, выраженное в процентах.

Как видно из представленных данных, наблюдается тенденция снижения средней точности классификации алгоритма SVDD при увеличении количества атрибутов. Также с увеличением количества атрибутов процент аномальных наблюдений в выборке растет, что обусловлено выбранным подходом «полностью нормальное наблюдение».

Заключение

Исследование алгоритма SVDD показывает его эффективность при малом объеме обучающей выборки — около 17%. Точность алгоритма не превышает 96% при работе с 6 атрибутами, при равномерном перемешивании входного набора данных.

Сравнительный анализ способов маркировки наблюдений показал преимущество подхода «полностью нормальное наблюдение» перед подходом «мажоритарное голосование». При использовании подхода «полностью нормальное наблюдение» точность классификации в случае упорядоченных во времени данных составила ~75%, а в случае равномерно перемешанных данных — 82%. При использовании подхода «мажоритарное голосование» точность классификации в случае упорядоченных во времени данных составила 35%, а в случае равномерно перемешанных данных — 69%.

Найдено, что при увеличении количества атрибутов наблюдается тенденция снижения средней точности классификации как перемешанных данных, так и упорядоченных по времени. С увеличением количества атрибутов процент аномальных наблюдений в выборке растет, что приводит к уменьшению точности.

Проведенная кросс-валидация подтверждает сильную зависимость точности классификации от распределения данных, представленных в обучающей выборке, что, однако может быть устранено равномерным перемешиванием входных данных. Выигрыш по точности при равномерном перемешивании составляет 15–20%.

Максимальная точность классификации при приемлемом потреблении ресурсов достигается при формировании компактного набора входных данных, наиболее полно описывающего нормальный режим работы компьютерной системы. Объем набора входных данных зависит, в первую очередь, от конкретной реализации алгоритма SVDD и может быть увеличен либо за счет оптимизации его программной реализации, либо за счет наращивания вычислительной мощности аппаратуры, на которой выполняется обработка данных.

Литература

1. *Шелухин О. И., Костин Д. В., Резник И. Ю.* Мониторинг и структура аномальных паттернов системных журналов компьютерных систем // REDS: Телекоммуникационные устройства и системы. 2020. № 2. С. 3–8.

2. *Водяхо А. И., Никифоров В. В.* Онтологические модели для систем реального времени // Онтология проектирования. 2018. № 2. С. 240–252. Doi:10.18287/2223–9537–2018–8–2–240–252

Таблица 3

Точность классификации

Количество атрибутов		**4**	**5**	**6**	**7**	**8**
Точность классификации по результатам проведенной кросс-валидации, %	**Данные упорядочены**	80,2	82,2	84,5	82,4	75,6
	Данные перемешаны	82,2	81,3	96,5	92,3	87,2
Время, затрачиваемое на обработку всей выборки, одна итерация, с		40	46	50	52	60
Процент аномальных наблюдений, %		21	25	46	50	57

3. *Чернов Д.В., Сычугов А.А.* Современные подходы к обеспечению информационной безопасности АСУ ТП // Известия Тульского государственного университета. Технические науки. 2018. № 10. С. 58–64.

4. *Довгаль В.А., Довгаль Д.В.* Роль туманных вычислений в интернете вещей // Вестник Адыгейского государственного университета. 2018. № 4. С. 205–209.

5. *Utkin V. L.* An imprecise extension of SVM-based machine learning models // Neurocomputing. 2019. No. 331. Pp. 18–32. Doi:10.1016/j.neucom.2018.11.053

6. *Kranjčić N., Medak D., Župan R., Rezo M.* Support Vector Machine Accuracy Assessment for Extracting Green Urban Areas in Towns // Remote Sens. 2019. No. 11. 655 p. Doi:10.3390/rs11060655

7. *Liu Z., Kang J., Zuo M.J., Zhao X., Qin Y., Jia L.* Modeling of the safe region based on support vector data description for health assessment of wheelset bearings // Applied Mathematical Modelling. 2019. No. 73. Pp. 19–39. Doi:10.1016/j.apm.2019.03.040

8. *Lv Y., Zhang J., Qin W., Yang J.* Adjustment mode decision based on support vector data description and evidence theory for assembly lines // Industrial Management & Data Systems. 2018. No. 8. C. 1711–1726. Doi:10.1108/IMDS-01–2017–0014

9. *Dai S., Yan J., Wang X., Zhang L.* A deep one-class model for network anomaly detection // IOP Conference Series: Materials Science and Engineering. 2 Ser. "2019 International Conference on Advanced Electronic Materials, Computers and Materials Engineering, AEMCME2019 — Computer Programming and Industrial Design". Changsha: Institute of Physics Publishing, 2019. P. 042007. Doi:10.1088/1757–899X/563/4/042007

10. *Копылов А.В., Середин О.С., Кушнир О.А., Грачева И.А., Ларин А.О.* Устойчивое детектирование ладони на изображениях на основе комбинирования информации о цвете и форме // Известия Тульского государственного университета. Технические науки. 2016. № 11–1. С. 24–40.

11. *Tan J., Fu W., Wang K., Hu W., Xue X., Shan Y.* Fault diagnosis for rolling bearing based on semi-supervised clustering and support vector data description with adaptive parameter optimization and improved decision strategy // Industrial Management & Data Systems. 2019. No. 8. Pp. 1676. Doi: 10.3390/app9081676

12. *Zhang H., Liu J., Li K., Tan H., Wang G.* Gait learning based authentication for intelligent things // IEEE Transactions on Vehicular Technology. 2020. No. 4. Pp. 4450–4459. Doi: 10.1109/TVT.2020.2977418

13. *Xiao Y., Gao H., Yan Y.* Indirect Gaussian kernel parameter optimization for one-class SVM in fault detection // Proceedings of SPIE — The International Society for Optical Engineering. 3. Ser. "Third International Workshop on Pattern Recognition" 2018. Jinan: SPIE, 2018. P. 108280K. Doi:10.1117/12.2501776

14. *Roy A., Ghosh A.K.* Some tests of independence based on maximum mean discrepancy and ranks of nearest neighbors // Statistics & Probability Letters. 2020. No. 164. P. 108793. Doi:10.1016/j.spl.2020.108793

15. *Wang Q., Lindsay B.* Pseudo-kernel method in u-statistic variance estimation with large kernel size // Statistica Sinica. 2017. No. 3. C. 1155–1174.

16. *Schölkopf B., Smola A.J., Bach F.* Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond // London: MIT Press, 2018. 648 p.

17. *Lv F., Han M.* Hyperspectral image classification based on multiple reduced kernel extreme learning machine // International Journal of Machine Learning and Cybernetics. 2019. No. 12. Pp. 3397–3405. Doi:10.1007/s13042–019–00926–5

18. *Шелухин О.И., Осин А.В., Костин Д.В.* Диагностика «здоровья» компьютерной сети на основе секвенциального анализа последовательностных паттернов // T-Comm: Телекоммуникации и транспорт. 2020. Том 14. № 2. С. 9–16. Doi:10.36724/2072–8735–2020–14–2–9–16

19. *Dorenskaya E.A., Semenov Yu.A.* The improved algorithm for calculation of the contextual words meaning in the text // Modern Information Technologies and IT-Education. 2019. Vol. 15. No. 4. Pp. 954–960. doi:10.25559/SITITO.15.201904.954–960

20. *Савченко Л.В.* Распознавание изолированных слов на основе взвешенного голосования дикторозависимых нейросетевых моделей // International Journal of Machine Learning and Cybernetics. 2020. № 5. С. 290–296. Doi:10.17587/it.26.290–296

BINARY CLASSIFICATION OF MULTI-ATTRIBUTE TAGGED DATA ABOUT ANOMALOUS EVENTS IN COMPUTER SYSTEMS USING THE SVDD ALGORITHM

OLEG I. SHELUHIN

Moscow, Russia, sheluhin@mail.ru

DMITRIY I. RAKOVSKIY

Moscow, Russia, dimitor1998@mail.ru

KEYWORDS: data markup, shuffling data, completely normal observation, voting by majority principle, small training sample, small data.

ABSTRACT

Introduction: At present, the volume of system logs of computer systems integrated into a distributed network infrastructure makes it impossible to manually check them in real time. Typically, the structure of each log record contains the numeric value of the observed attribute and a corresponding flag to mark the record as normal or abnormal. The support vector data description algorithm demonstrates high classification accuracy even with small volumes of the training sample. A feature of the algorithm is the work with a multi-attribute dataset, where each observation contains a common classifying marking. Con-

sequently, the problem arises of reducing the set of markings of the attributes of the initial data to one marking of the entire observation. **Purpose:** to investigate the accuracy of the binary classification of experimental data of the Support Vector Data Description algorithm with a small volume of the training sample, provided that the data are labeled for each attribute separately. **Methods:** a method is proposed for solving the problem of reducing the set of markings of the attributes of the initial data to one single marking of the entire observation by means of two approaches: "normal observation" and voting by the majority principle. Two types of data are considered: ordered in time and uniformly mixed. The classification accuracy was assessed by calculating the area under the ROC curves with cross-validation for a different number of attributes. **Results:** a comparative analysis of observation labeling methods showed the advantage of the "completely normal observation" approach over the "majority vote" approach without "weighting". It is shown that the classification accuracy on mixed data is 7% higher compared to the variant of data ordering in time. The accuracy of the algorithm was investigated for a different number of attributes using the "completely normal observation" approach. The maximum achieved classification accuracy was about 96% when working with 6 attributes, with uniform mixing of the input dataset. A further increase in the number of attributes leads to a decrease in the average classification accuracy due to an increase in the proportion of anomalous observations. It is shown that when using uniform mixing of input data, the gain in accuracy can be increased by 15–20%. **Practical relevance:** the algorithm demonstrates an exponential growth in the consumption of computing resources with an increase in the amount of input data. **Discussion:** to achieve the maximum classification accuracy with acceptable resource consumption, it is necessary to form a compact set of input data, which most fully reflects the functioning of the computer system in normal mode.

REFERENCES

1. Sheluhin O.I., Kostin D.V., Reznik I. Yu. Monitoring i struktura anomal'nykh patternov sistemnykh zhurnalov komp'yuternykh sistem [Monitoring and structure of abnormal patterns of system logs of computer systems]. *REDS: Telecommunication devices and systems*. 2020. No. 2. Pp. 3–8. (In Rus)

2. Vodyakho A.I., Nikiforov V.V. Ontology models for real time systems. *Ontologiya proektirovaniya* [Design Ontology]. 2018. No. 2. Pp. 240–252. (In Rus). Doi:10.18287/2223–9537–2018–8–2–240–252

3. Chernov D.V., Sychugov A.A. A modern approaches to information security of automated process control systems. *Izvestiya Tul'skogo gosudarstvennogo universiteta* [Izvestiya Tula State University]. 2018. No. 10. Pp. 58–64. (in Rus)

4. Dovgal V.A., Dovgal D.V. Rol' tumannykh vychisleniy v internete veshchey [Role of fog computing in the internet of things]. *Vestnik Adygeyskogo gosudarstvennogo universiteta* [The Bulletin of the Adyghe State University: Internet Scientific Journal]. 2018. No 4. Pp. 205–209. (In Rus)

5. Utkin V.L. An imprecise extension of SVM-based machine learning models. *Neurocomputing*. 2019. No. 331. Pp. 18–32. Doi:10.1016/j.neucom.2018.11.053

6. Kranjčić N., Medak D., Župan R., Rezo M. Support Vector Machine Accuracy Assessment for Extracting Green Urban Areas in Towns. *Remote Sens*. 2019. No. 11. 655 p. Doi:10.3390/rs11060655

7. Liu Z., Kang J., Zuo M.J., Zhao X., Qin Y., Jia L. Modeling of the safe region based on support vector data description for health assessment of wheelset bearings. *Applied Mathematical Modelling*. 2019. No. 73. Pp. 19–39. Doi:10.1016/j.apm.2019.03.040

8. Lv Y., Zhang J., Qin W., Yang J. Adjustment mode decision based on support vector data description and evidence theory for assembly lines. *Industrial Management & Data Systems*. 2018. No 8. Pp. 1711–1726. Doi:10.1108/IMDS-01–2017–0014

9. Dai S., Yan J., Wang X., Zhang L. A deep one-class model for network anomaly detection. *IOP Conference Series: Materials Science and Engineering. 2. Ser. "2019 International Conference on Advanced Electronic Materials, Computers and Materials Engineering, AEMCME2019 – Computer Programming and Industrial Design"*. Changsha: Institute of Physics Publishing, 2019. Pp. 042007. Doi:10.1088/1757–899X/563/4/042007

10. Kopylov A.V., Seredin O.S., Kushnir O.A., Gracheva I.A., Larin A.O. Robust palm detection based on combining of color and shape information. *Izvestiya Tul'skogo gosudarstvennogo universiteta* [Izvestiya Tula State University]. 2016. No. 11–1. Pp. 24–40. (In Rus)

11. Tan J., Fu W., Wang K., Hu W., Xue X., Shan Y. Fault diagnosis for rolling bearing based on semi-supervised clustering and support vector data description with adaptive parameter optimization and improved decision strategy. *Industrial Management & Data Systems*. 2019. No. 8. Pp. 1676. Doi: 10.3390/app9081676

12. Zhang H., Liu J., Li K., Tan H., Wang G. Gait learning based authentication for intelligent things. *IEEE Transactions on Vehicular Technology*. 2020. No. 4. Pp. 4450–4459. Doi: 10.1109/TVT.2020.2977418

13. Xiao Y., Gao H., Yan Y. Indirect Gaussian kernel parameter optimization for one-class SVM in fault detection. *Proceedings of SPIE – The International Society for Optical Engineering. 3. Ser. "Third International Workshop on Pattern Recognition" 2018*. Jinan: SPIE, 2018. Pp. 108280K. Doi:10.1117/12.2501776

14. Roy A., Ghosh A.K. Some tests of independence based on maximum mean discrepancy and ranks of nearest neighbors. *Statistics & Probability Letters*. 2020. No 164. Pp. 108793. Doi:10.1016/j.spl.2020.108793

15. Wang Q., Lindsay B. Pseudo-kernel method in u-statistic variance estimation with large kernel size. *Statistica Sinica*. 2017. No. 3. Pp. 1155–1174.

16. Scholkopf B., Smola A. J., Bach F. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. *MIT Press*. 2018. 648 p.

17. Lv F., Han M. Hyperspectral image classification based on multiple reduced kernel extreme learning machine. *International Journal of Machine Learning and Cybernetics*. 2019. No 12. Pp. 3397–3405. Doi:10.1007/s13042–019–00926–5

18. Sheluhin O.I., Osin A.V., Kostin D.V. Diagnostika "zdorov'ya" komp'yuternoj seti na osnove sekvencial'nogo analiza posledova-

tel'nostnyh patternov [Health monitoring of a computer network based on sequential analysis of serial pattern]. *T-Comm*. 2020. Vol. 14. No. 2. Pp. 9–16. (In Rus). Doi:10.36724/2072-8735-2020-14-2-9-16

19. Dorenskaya E.A., Semenov Yu.A. The improved algorithm for calculation of the contextual words meaning in the text. *Modern Information Technologies and IT-Education*. 2019. No. 4. Pp. 954–960. Doi:10.25559/SITITO.15.201904.954-960

20. Savchenko L.V. Raspoznavanie izolirovannykh slov na osnove vzveshennogo golosovaniya diktorozavisimykh neyrosetevykh modeley [Isolated words recognition based on weighted voting of speaker-dependent neural network acoustic models]. *International Journal of Machine Learning and Cybernetics*. 2020. No. 5. Pp. 290–296. Doi:10.17587/it.26.290-296 (In Rus)

INFORMATION ABOUT AUTHORS:

Sheluhin O. I., PhD, Full Professor, Head of Department Information Security of the Moscow Technical University of Communications and Informatics;

Rakovskiy D.I., Moscow Technical University of Communication and Informatics.

For citation: Sheluhin O. I., Rakovskiy D.I. Binary classification of multi-attribute tagged data about anomalous events in computer systems using the SVDD algorithm. *H&ES Research*. 2021. Vol. 13. No. 2. Pp. 74–84. Doi: 10.36724/2409-5419-2021-13-2-74-84 (In Rus)