Intellectual Technologies on Transport No 3

Интеллектуальные технологии на транспорте № 3

Санкт-Петербург
St. Petersburg
2022

Интеллектуальные технологии на транспорте
№ 3, 2022

ISSN 2413-2527

Сетевой электронный научный журнал, свободно распространяемый через Интернет.
Публикуются статьи на русском и английском языках с результатами исследований
и практических достижений в области интеллектуальных технологий
и сопутствующих им научных исследований.

Журнал основан в 2015 году.

Учредитель и издатель
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Петербургский государственный университет путей сообщения Императора Александра I» (ФГБОУ ВО ПГУПС)

Адрес редакции:
190031, Санкт-Петербург, Московский пр., 9, ауд. 1–202
e-mail: itt-pgups@yandex.ru

Журнал зарегистрирован Федеральной службой по надзору в сфере связи и массовых коммуникаций,
свидетельство ЭЛ № ФС77-61707 от 07 мая 2015 г.

Журнал зарегистрирован в Российском индексе научного цитирования (РИНЦ).

Периодичность выхода – 4 номера в год. Выпуски журнала доступны на сайте http://itt-pgups.ru.

Копии архивов с выпусками журнала проходят государственную регистрацию как электронное издание
сетевого распространения в НТЦ "Информрегистр".

Информация предназначена для детей старше 12 лет.

Intellectual Technologies on Transport
Issue 3, 2022

ISSN 2413-2527

Network electronic scientific journal, open access. It publishes articles in Russian and English with the results of research and practical achievements in the field of intelligent technologies and associated research.

Founded in 2015.

Founder and Publisher

Federal State Educational Institution of Higher Education
«Emperor Alexander I Petersburg State Transport University»

Editorial adress:

190031, St. Petersburg, Moskovsky ave., 9, aud. 1–202
e-mail: itt-pgups@yandex.ru

The journal is registered by the Federal Service for Supervision of Communications and Mass Media, EL No. FS77-61707 testimony from May 7, 2015.

The journal is registered in the Russian Science Citation Index (RSCI).

Frequency of release - 4 issues per year. Issues of the magazine are available at http://itt-pgups.ru.

Copies of the archives with the issues of the journal are state-registered as an electronic publication of network distribution in the Scientific and Technical Center «Informregister».

The content is for children over the age of 12.

Содержание

Contents

DOI: 10.24412/2413-2527-2022-331-5-18

Методика обоснования тестовых информационно-технических воздействий при анализе защищенности объектов информатизации железнодорожного транспорта

Г. Е. Смирнов
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)
Санкт-Петербург, Россия
science.cybersec@yandex.ru

Аннотация. **В статье рассмотрены основные информационные системы и объекты информатизации железнодорожного транспорта. Показано, что они являются ключевыми объектами критической информационной инфраструктуры Российской Федерации, а анализ состояния их реальной защищенности — важной государственной задачей. Предложено проводить анализ защищенности за счет использования тестовых информационно-технических воздействий, аналогичных воздействиям, которые прогнозируются к применению злоумышленниками. Разработана методика обоснования тестовых воздействий при анализе защищенности объектов информатизации железнодорожного транспорта на основе алгоритма Дейкстры, позволяющая формировать множество путей, ранжированных по суммарной метрике пути, и состоящая из двух этапов: формирования упорядоченного множества путей тестирования и выбора путей тестирования, обеспечивающих рациональную полноту оценки уязвимостей, при ограничениях на ресурсы.**

Ключевые слова: **информационная безопасность, аудит, тестирование, алгоритм Дейкстры, информационно-техническое воздействие, критическая информационная инфраструктура, объект информатизации, железнодорожный транспорт.**

Введение

В 2017 году в России принят федеральный закон № 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации». Данный закон устанавливает перечень объектов, относящихся к критической информационной инфраструктуре (КИИ) РФ, а также обязует владельцев объектов КИИ разработать комплекс мер, направленных на обеспечение их информационной безопасности (ИБ). При этом к КИИ отнесен и железнодорожный транспорт (ЖТ), в связи с чем актуальным является формирование новых предложений по повышению полноты аудита ИБ объектов информатизации (ОИ) ЖТ как объекта КИИ.

Вопросы обеспечения ИБ и оценки защищенности различных ОИ исследованы в работах [1–9]. Вопросам состава, структуры и функционирования информационных систем (ИС) и ОИ ЖТ посвящены работы [10–12]. Работы [13–21] посвящены вопросам оценки ИБ ОИ и ИС ЖТ. Вместе с тем вопросы оценки защищенности ОИ ЖТ, именно за счет использования тестовых информационно-технических воздействий (ИТВ), исследованы в недостаточной степени.

Целью статьи является разработка методики обоснования тестовых ИТВ при анализе защищенности ОИ ЖТ. Такое тестирование, по замыслу автора, дополнит стандартные мероприятия анализа защищенности ОИ ЖТ и повысит полноту оценки их ИБ.

Данная работа продолжает и развивает направление исследований, проводимое научной школой С. И. Макаренко, посвященное развитию теории и практики тестирования на проникновения в рамках аудита ИБ, представленное работами [22–28].

Анализ объектов информатизации и задач обеспечения ее защищенности

Анализ работ [10–12] показал, что ИС ЖТ относится к классу больших корпоративных систем, содержит большое количество ОИ и предназначена для решения как информационных задач, так и задач управления ЖТ. Главная цель применения ИС ЖТ состоит в информационном обеспечении технологических процессов и автоматизации принятия решений в сфере ЖТ в интересах достижения максимальной эффективности его работы в условиях рыночной экономики.

ИС ЖТ представляется в виде двухуровневой структуры. Первый уровень — обеспечивающий — представлен информационной средой и инфраструктурой информатизации, второй уровень — прикладной — реализуется путем использования ОИ и информационных технологий (ИТ), объединенных в ИТ-комплексы, решающих конкретные задачи управления и автоматизации функций ЖТ.

Информационная среда — информация, реализованная в системе баз данных (БД), которая обеспечивает функционирование ОИ, органов управления и отдельных пользователей ЖТ. Информационная среда формирует единое информационное пространство (ЕИП), в котором все абоненты и пользователи ЖТ обеспечены необходимой им информацией.

Инфраструктура информатизации ЖТ включает в себя:

1. Главный вычислительный центр (ГВЦ) ЖТ, объединяющий и поддерживающий БД для проведения общесетевой маркетинговой, финансовой и экономической деятельности и управления перевозочным процессом.

2. Информационно-вычислительные центры (ИВЦ) ЖТ на дорогах, реализующие комплексы информационных услуг для управлений и отделений дорог.

3. Сети связи и телекоммуникаций, устройства автоматического съема информации с подвижного состава, вычислительное оборудование, обеспечивающее выполнение операций формирования, сбора, передачи, хранения, обработки и представления информации.

Обеспечение автоматизации основных функций ЖТ выполняют ИТ-комплексы:

- управления перевозочным процессом;
- управления маркетингом, экономикой и финансами;
- управления инфраструктурой ЖТ;
- управления непроизводственной сферой.

Рассмотрим эти комплексы более подробно.

ИТ-комплекс управления перевозочным процессом обеспечивает информационное сопровождение в области грузовых и пассажирских перевозок. Основными функциями по управлению грузовыми перевозками являются организация поездо- и грузопотоков на сети, диспетчерское управление поездной работой, управление локомотивными и вагонными парками, грузовой и коммерческой работой, обслуживание грузовой клиентуры, разработка графика движения поездов, норм эксплуатационной работы, планирование перевозок и прочее. Основными функциями по управлению пассажирскими перевозками являются организация обслуживания пассажиров и информационно-справочный сервис, планирование пассажирских перевозок в международном и внутридорожном сообщении, управление нормативами, тарифами внутренних и международных перевозок, организация эксплуатации и ремонта парка пассажирских вагонов, управление багажными и почтовыми перевозками, организация билетно-кассовых операций и др. В рамках этого ИТ-комплекса функционируют:

– автоматизированная система оперативного управления перевозками (АСОУП) — основной элемент ИТ-комплекса управления перевозочным процессом;

– система резервирования и продажи билетов («Экспресс-2»);

– единые центры диспетчерского управления (ЕЦДУ);

– система учета, контроля дислокации, анализа использования и регулирования вагонного парка (ДИСПАРК);

– автоматизированная система контроля за использованием и продвижением контейнеров (ДИСКОН);

– автоматизированная система фирменного транспортного обслуживания (АКС ФТО);

– автоматизированные системы управления сортировочными (АСУ СС), грузовыми (АСУ ГС) станциями и контейнерными пунктами (АСУ КП);

– автоматизированная система централизованной подготовки и оформления перевозочных документов (ЭТРАН);

– сетевая интегрированная Российская информационно-управляющая система (СИРИУС) и др.

ИТ-комплекс управления маркетингом, экономикой и финансами охватывает финансовую деятельность, бухгалтерский учет, маркетинговую деятельность и тарифную политику, управление развитием отрасли ЖТ, технической политикой и научно-исследовательскими и опытно-конструкторскими работами, нормативно-правовую работу, управление эксплуатационными расходами и др. ИТ этого комплекса ориентированы на формирование заказов, увеличение доходов, укрепление конъюнктурного положения за счет сохранения и увеличения доли ЖТ на транспортном рынке страны, на стабильное обеспечение денежных и платежных ресурсов, минимизацию затрат, на совершенствование экономической работы и инвестиционной политики. В рамках комплекса функционируют и внедряются ИТ управления финансовой деятельностью, ресурсами, способы расчетов за грузовые перевозки, взаиморасчетов за пользование вагонами и др. Основу этого ИТ-комплекса составляет единый комплекс автоматизированной системы управления финансовой деятельностью (ЕК АСУФР).

ИТ-комплекс управления инфраструктурой ЖТ представлен базовыми информационными технологиями, охватывающими управление эксплуатационной работой пассажирского хозяйства, хозяйств пути и сооружений, информатизации и связи, хозяйства энергоснабжения, локомотивного и вагонного хозяйств, управление проектированием и капитальным строительством объектов инфраструктуры, управление ремонтно-восстановительными работами и работами в чрезвычайных условиях, управление промышленностью ЖТ, материально-техническим снабжением и т. д. В составе этого ИТ-комплекса функционируют различные автоматизированные системы управления технологическими процессами (АСУ ТП): управления путевым хозяйством, устройствами энергоснабжения, сигнализации, средствами информатизации и связи.

ИТ-комплекс управления непроизводственной сферой железнодорожного транспорта представляет собой совокупность функций, обеспечивающих управление персоналом, учебными заведениями, жилищно-коммунальным хозяйством, рабочим снабжением, здравоохранением.

Основными факторами, актуализирующими значимость вопросов обеспечения ИБ, применительно к ИС ЖД являются следующие [19]:

– интеграция в единые ИТ-комплексы подавляющего числа критических функций, связанных с управлением движением поездов и жизнедеятельности ЖТ;

– постоянное усложнение программного обеспечения (ПО) и оборудования, используемых в ИТ-комплексах управления ЖТ;

– существующая практика удаленной настройки и технического обслуживания элементов ИС ЖТ, осуществляемая разработчиками и поставщиками оборудования, входящего в состав элементов информационной инфраструктуры железнодорожного транспорта;

– интенсивное совершенствование потенциальными злоумышленниками средств и способов ИТВ, методов социальной инженерии для нанесения ущерба, а также участившиеся попытки их применения в противоправных целях и конкурентной борьбе;

– риск сокрытия попыток или фактов нарушения штатного функционирования ИС ЖТ со стороны эксплуатирующих подразделений;

– временное вынужденное привлечение к созданию элементов ИТ-комплексов ЖТ, в том числе АСОУП и различных АСУ ТП, производителей и поставщиков программно-аппаратных средств обработки, хранения и передачи информации и применение неконтролируемых программно-аппаратных решений.

Помимо вышеуказанных факторов нужно отметить следующее. ЖТ является одним из ключевых элементов транспортной инфраструктуры РФ, обеспечивая до 88 % грузооборота страны (для сравнения: доля автомобильного транспорта составляет 4 %, а водного — 8%) [16]. В связи с этим ЖТ выступает одной из основных целей для злоумышленников и профессиональных нарушителей — сил информационных операций недружественных стран при ведении информационного противоборства. При обострении геополитической обстановки в мире информационная инфраструктура и ИС ЖТ РФ могут оказаться объектом воздействия не только злоумышленников, но и профессиональных нарушителей, поэтому оценка реальной защищенности ОИ и ИС ЖТ является важной задачей, имеющей государственное значение.

Постановка задачи на разработку методики

В предыдущей статье по этой тематике [25] автором была сформирована модель процесса тестирования ОИ ЖТ в виде многоуровневой топологической модели, которая взаимосвязанно учитывает: эффективность отдельных ИТВ i в части выявленного и потенциально предотвращенного ущерба $\{z\}$; ориентированность их на проверку конкретного множества уязвимостей $\{u\}$ элементов $\{e\}$ объекта информатизации; расход в процессе тестирования определенного количества ресурса r_i (в данном случае под абстрактным ресурсом может пониматься расход времени аудитора, оплата его труда, стоимость машинного времени, затраты на специализированное оборудование и т. д.). В данной работе будет показано, как с использованием модели [25] сформировать набор тестовых ИТВ, обеспечивающий рациональную полноту аудита защищенности ОИ ЖТ.

Задача на разработку методики m обоснования набора тестовых ИТВ для рациональной полноты оценки уязвимостей ОИ ЖТ формулируется следующим образом. Сформировать такой набор тестовых ИТВ $I = \{i\}$, который бы в условиях ограниченности ресурсов аудитора R максимизировал важность выявляемых уязвимостей $\{u\}$, с учетом того, что отдельным уязвимостям u и элементам ОИ ЖТ e сопоставляются уровни ущерба $z(e, u, i, \sigma)$, наносимого ОИ S по определенному свойству ИБ σ (конфиденциальность, целостность, доступность) при потенциальной эксплуатации уязвимости u элемента e злоумышленником путем применения i-го ИТВ. При этом абсолютным показателем рациональной полноты π является сумма «стоимости выявленного и потенциально предотвращенного ущерба» $z(e, u, i, \sigma)$ при использовании тестового набора $\{i\}$ для тестирования уязвимостей $\{u\}$ относительно тестируемых элементов ОИ $\{e\}$ и свойств ИБ $\{\sigma\}$:

$$\sum_{\{i\},\{u\},\{e\}} z(e,u,i,\sigma) = \pi.$$

Относительным значением рациональной полноты $\pi_{отн}$ является абсолютный показатель рациональной полноты π, отнесенный к сумме ущерба П по всем возможным комбинациям ИТВ $\{i\}$ потенциальных злоумышленников, уязвимостей $\{u\}$ элементов объекта $\{e\}$ и свойств ИБ $\{\sigma\}$:

$$\pi_{отн} = \frac{\pi}{\Pi}.$$

Фактически, требуется найти такие тестовые ИТВ, которые при ограниченных затратах ресурса R максимизировали бы стоимость выявленного и предотвращенного ущерба π.

Введение системы обозначений

Для формализации методики введем следующие обозначения:

$\pi/\pi_{отн}$ — абсолютное/относительное значение полноты выявленного и потенциально предотвращенного ущерба;

$\pi_m/\pi_{отн\ m}$ — абсолютное/относительное значение полноты выявленного и потенциально предотвращенного ущерба m-м ИТВ в тестовом наборе;

B — множество узлов потенциальных дополнительных путей тестирования;

C — множество весов ребер потенциальных дополнительных путей тестирования;

$E = \{e\}$ — множество элементов, составляющих ОИ;

e_j — j-й элемент ОИ;

$G(W, V)$ — граф модели тестирования защищенности ОИ;

$I = \{i\}$ — множество тестовых ИТВ;

i_j — j-е тестовое ИТВ;

j, l, m, n — переменные-счетчики;

L — множество смежных помеченных вершин графа G, то есть множество расстояний до помеченных вершин от начальной вершины;

N — количество узлов в графе G;

N_I — количество тестовых ИТВ, которое соответствует количеству элементов множества I;

N_U — количество уязвимостей, которое соответствует количеству элементов множества U;

P — множество помеченных вершин в графе G;

Q — множество дополнительных путей в узлы, которое содержит дополнительные пути в рассматриваемый узел, сформированные в результате проведения логических операций над входящими в него элементами и элементами множеств B и L;

R — исходный узел ресурсов в графе G модели тестирования защищенности ОИ;

$R_{гр}$ — ограничения на ресурс, расходуемый в процессе тестирования защищенности ОИ;

$R_{тест}$ — затраты ресурса, необходимые для тестирования защищенности ОИ тестовым набором T;

r_j — количество ресурса аудитора, расходуемое на организацию и проведение j-го тестового ИТВ;

S — множество весов дополнительных путей к узлам графа G;

$T = \{t\}$ — множество тестовых ИТВ, выбранных для проведения тестирования защищенности ОИ в результате применения методики;

t — тестовое ИТВ, включенное в тестовый набор T для проведения тестирования защищенности ОИ;

u — уязвимость ОИ;

$U = \{u\}$ — множество уязвимостей ОИ;

V — множество весов ребер в графе G модели тестирования защищенности ОИ;

$V(W_n, W_j)$ — вес ребра, соединяющего произвольные n-й и j-й узлы графа G;

W — множество узлов графа G модели тестирования защищенности ОИ независимо от уровней расположения ($W = R \vee I \vee U \vee E \vee Z$);

Z — конечный узел ущерба в графе G модели тестирования защищенности ОИ;

z — ущерб;

$z(e_j, \sigma_n)$ — ущерб от нарушения свойства ИБ σ_n у элемента e_j;

$Z = \{z\}$ — суммарный показатель ущерба, который может быть причинен ОИ;

σ_n – свойство ИБ: $n = 1$ — доступность; $n = 2$ — целостность; $n = 3$ — конфиденциальность;

П — сумма ущерба по всем возможным комбинациям ИТВ $\{i\}$ потенциальных злоумышленников, уязвимостей $\{u\}$ элементов объекта $\{e\}$ и свойств ИБ $\{\sigma\}$.

Исходные положения и посылки

Разработку методики обоснования набора тестовых ИТВ предполагается вести на основе приложения подходов к исследованию теории графов к модели тестирования защищенности объекта информатизации [17]. Введем понятие пути тестирования.

Путь тестирования — путь на графе модели тестирования защищенности объекта информатизации, проходящий через узлы и ребра, которые соответствуют единственной оригинальной комбинации ресурса r_i, тестового ИТВ i, уязвимости u элемента ОИ e и уровня ущерба $z(i, u, e, \sigma)$, наносимого ОИ по свойству ИБ σ.

В результате введения такого понятия задача обоснования набора тестовых ИТВ может быть сведена к задаче поиска множества кратчайших путей тестирования на графе модели тестирования защищенности ОИ.

В качестве графа, на котором будет вестись поиск путей тестирования, а также соответствующих им ИТВ, будем использовать преобразованную модель модели оценки защищенности ОИ, вариант которой представлен в работе [17] (рис. 1).

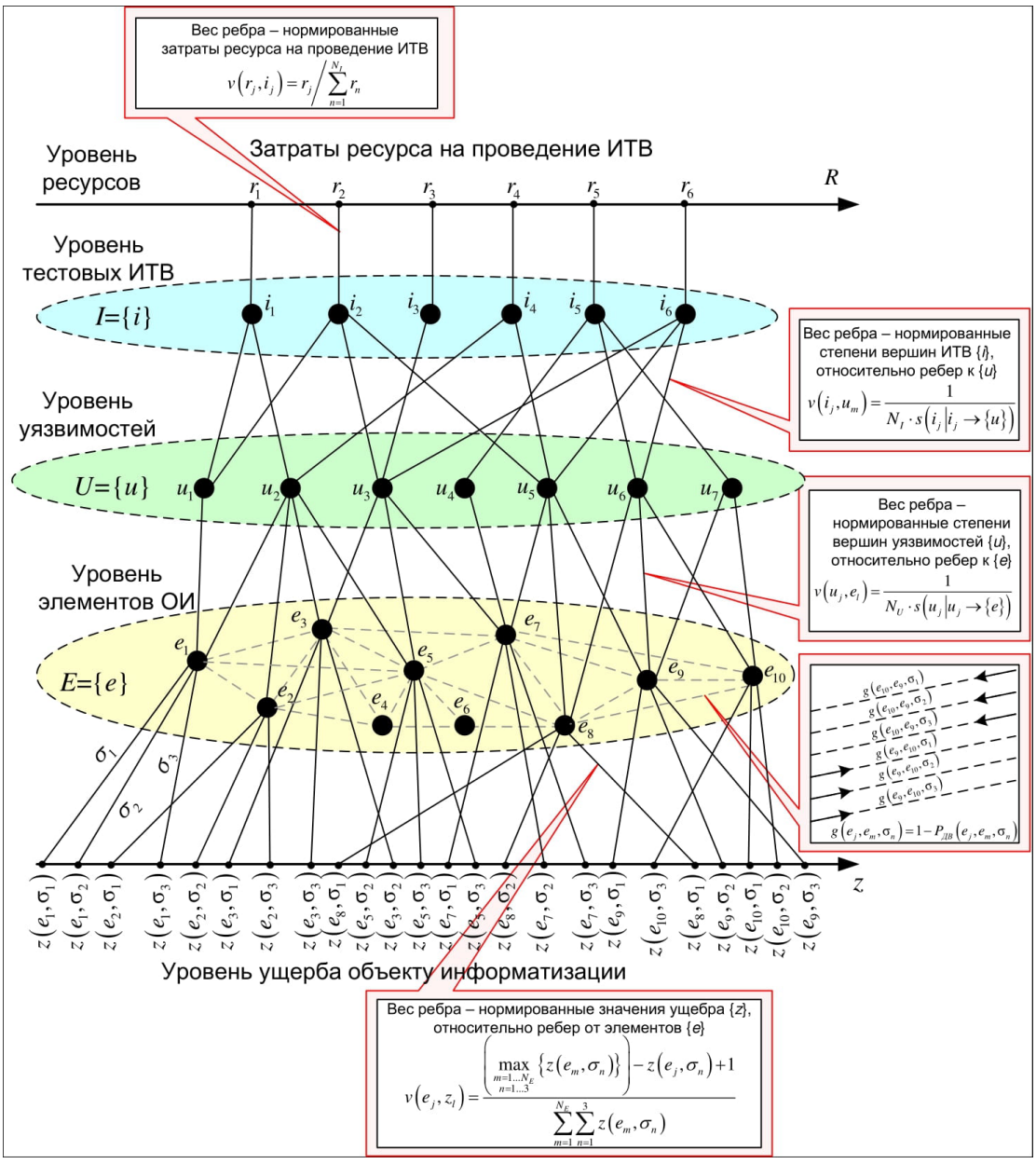

Рис. 1. Схема модели оценки защищенности ОИ тестовыми ИТВ

Особенностью этого графа является то, что «наилучшие» ребра, с точки зрения полноты и стоимости тестирования, обладают минимальным весом, а в целом веса ребер упорядочены по мере возрастания весов при переходе от «лучших» к «худшим» путям тестирования. Логика формирования набора тестовых ИТВ подразумевает наличие направленного графа. В связи с этим преобразуем ненаправленный граф модели оценки защищенности ОИ (рис. 1) в направленный граф, в котором направления ребер заданы сверху вниз (рис. 2).

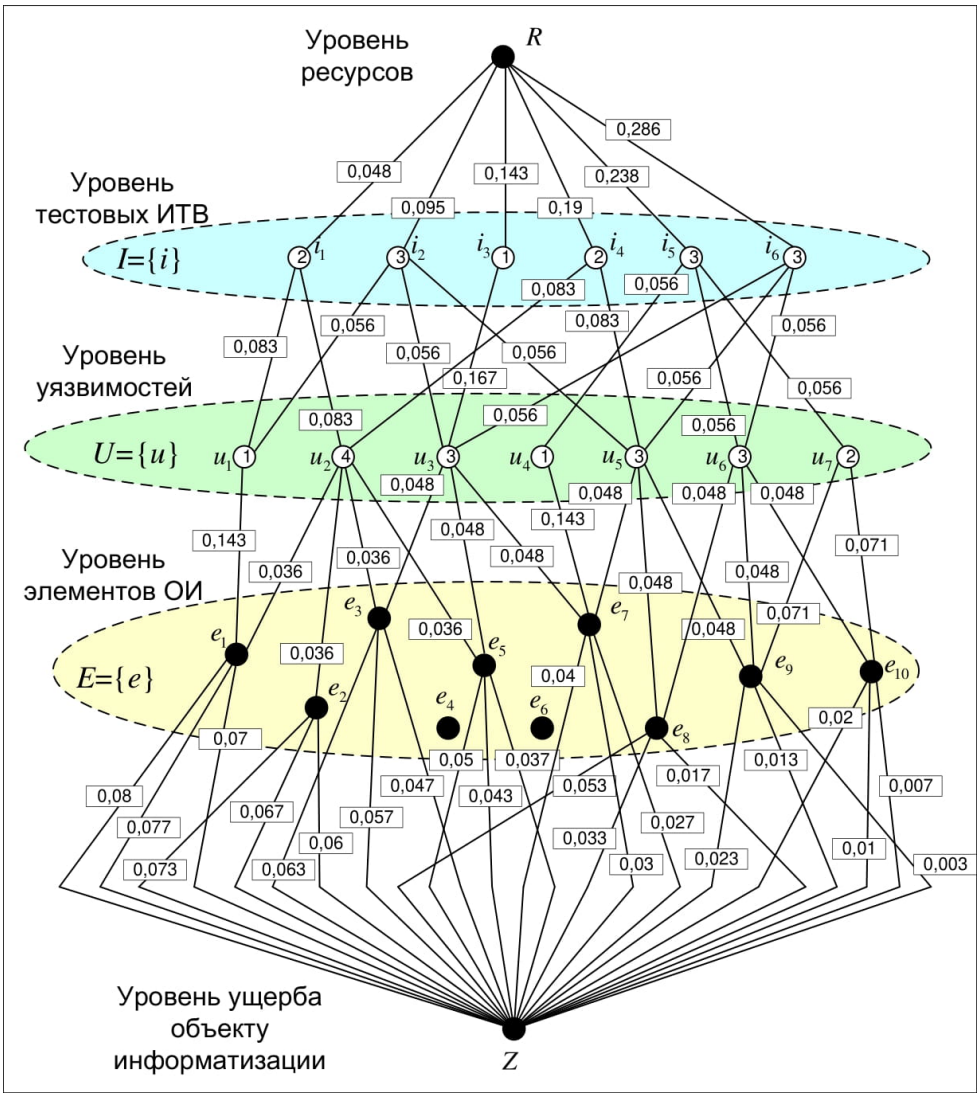

Рис. 2. Вариант модели оценки защищенности ОИ

Анализ фундаментальных работ в области теории графов [29, 30] показал, что для решения задачи вычисления кратчайших путей в графах применяются соответствующие математические алгоритмы поиска кратчайших путей. При этом наиболее широко используемым таким алгоритмом является алгоритм Дейкстры [31]. Однако особенностью этого алгоритма является то, что он является «поглощающим» и формирует из каждого узла графа к другому узлу только один путь, являющийся кратчайшим по сумме весов ребер в сети. Таком образом, можно обосновать только единственный оптимальный вариант одиночного ИТВ. Вместе с тем для обоснования набора нескольких ИТВ необходимо вычислять не только кратчайшие пути тестирования, но и другие комбинации путей, соответствующих другим ИТВ, после чего группировать их степени увеличения стоимости тестирования. Это требует формирования набора путей тестирования, которые были бы ранжированы, с одной стороны, по уровню вскрываемого ущерба, а с другой — по степени затрат ресурсов на тестирование. Решение этой задачи потребует создания нового математического алгоритма на основе алгоритма Дейкстры с целью разработки новой функциональности — способности формировать множество путей, ранжированных по суммарной метрике пути, из начального узла графа (R) в конечных узел (Z). Решение подобной задачи уже рассматривалось в работах [32–36], однако эти работы не имеют отношения к вопросам ИБ, а посвящены исключительно вопросам обоснования маршрутов передачи данных в компьютерных сетях. Предлагается, приняв работы [32–36] за теоретический базис, разработать методику обоснования набора тестовых ИТВ путем нахождения комбинаций путей тестирования в графе модели, представленной на рисунке 2, при этом в основу методики положить математический алгоритм, основанный на алгоритме поиска кратчайших путей Дейкстры [33].

ПЕРВЫЙ ЭТАП МЕТОДИКИ — ФОРМИРОВАНИЕ УПОРЯДОЧЕННОГО МНОЖЕСТВА ПУТЕЙ ТЕСТИРОВАНИЯ

В ходе модификации алгоритма Дейкстры в него дополнительно вносятся изменения, направленные на расширение его функциональности, связанной с возможностью формирования нескольких путей, ранжированных по степени повышения метрики. Основой предлагаемой модификации алгоритма Дейкстры являются следующие положения, ранее обоснованные в работе [33].

1. При достижении очередного узла в графе запоминаются исходящие узлы входящих в этот узел ребер как потенциальные элементы будущих дополнительных путей тестирования к этому узлу.

2. При очередном шаге функционирования методики достигнутый очередной узел графа модели проверяется как потенциальный элемент дополнительного пути тестирования для всех уже достигнутых узлов. Если он является потенциальным элементом дополнительного пути, формируется дополнительный путь к ранее достигнутому узлу через только что достигнутый узел.

3. Если к ранее достигнутому узлу графа модели уже были сформированы дополнительные пути и он участвует в создании нового дополнительного пути к очередному узлу, то к очередному узлу формируется множество дополнительных путей с включением в них всех возможных вариантов дополнительных путей, сформированных ранее. Причем если в дополнительный путь входит сам очередной узел модели, то такой путь во избежание циклов в дополнительные не включается.

4. Все дополнительные пути к узлам модели упорядочиваются в соответствии с минимизацией суммы весов входящих в них ребер и вносятся в таблицу путей тестирования одновременно с кратчайшим путем.

Схема формирования упорядоченного множества путей тестирования на основе модифицированного алгоритма Дейкстры, ранее разработанного автором и представленного в работе [26], приведена на рисунке 3.

Входными параметрами этого этапа методики являются:

а) граф модели тестирования защищенности объекта информатизации — $G(W, V)$, где W — множество узлов графа G модели тестирования защищенности объекта информатизации, на основе которого формируются пути тестирования; V — множество весов ребер в графе G модели тестирования защищенности объекта информатизации;

б) количество узлов в графе G — N;

в) вес ребер, соединяющих произвольные n-й и j-й узлы $V(W_n, W_j)$ графа G.

Для обеспечения поиска не только кратчайшего, но и других дополнительных путей тестирования помимо имеющихся множеств, предусмотренных логикой функционирования алгоритма Дейкстры (P — множество помеченных вершин, L — множество смежных помеченных вершин, множество расстояний до помеченных вершин от начальной вершины), вводятся следующие дополнительные множества:

а) B — множество узлов потенциальных дополнительных путей. В это множество вносятся достигнутые узлы, смежные рассматриваемому. В дальнейшем элементы множества используются при нахождении дополнительных путей;

б) C — множество весов ребер потенциальных дополнительных путей. В это множество вносятся веса ребер, исходящих из узлов, вносимых в множество B и входящих в рассматриваемый узел;

в) Q — множество дополнительных путей в узлы. Содержит дополнительные пути в рассматриваемый узел, сформированные в результате проведения логических операций над входящими в него элементами и элементами множеств B и L.

г) S — множество весов дополнительных путей к узлам. Это множество содержит веса путей из множества Q и используется для ранжирования дополнительных путей при выводе результатов функционирования данного этапа методики.

К блокам, отличающим данный этап методики от известного алгоритма Дейкстры, относятся блоки 16–23, 25 на рисунке 3. В блоках 16–17 реализуется формирование элементов множества узлов B к текущему рассматриваемому узлу за счет использования положения № 1 по модификации алгоритма Дейкстры. Далее, в блоках 18–23, путем пересечения элементов множества B и L, а также Q осуществляется формирование элементов множества Q с учетом положения № 2 по модификации алгоритма Дейкстры. В блоке 25 осуществляется ранжировка дополнительных маршрутов по сумме весов входящих в их состав ребер. Блоки 3–15, 24 соответствуют стандартному алгоритму Дейкстры. По итогам работы нулевому элементу множества Q присваивается значение кратчайшего пути из множества L.

ВТОРОЙ ЭТАП МЕТОДИКИ — ВЫБОР ПУТЕЙ ТЕСТИРОВАНИЯ, ОБЕСПЕЧИВАЮЩИХ РАЦИОНАЛЬНУЮ ПОЛНОТУ ОЦЕНКИ УЯЗВИМОСТЕЙ, ПРИ ОГРАНИЧЕНИЯХ НА РЕСУРСЫ

Содержание данного этапа состоит в выборе из кратчайшего пути и упорядоченного по возрастанию весов множества путей Q (с весами, сформированными в множестве S) на графе модели G такого ранжированного множества ИТВ и формирование из них тестового набора T, который бы обеспечивал максимизацию абсолютной суммарной стоимости обнаруженного ущерба $\pi \to \max$ (относительного значения $\pi_{отн} \to 100$ %) в рамках заданных ограничений на расход ресурса тестирования $R_{гр}$.

В целом этап выбора путей тестирования, обеспечивающих рациональную полноту оценки уязвимостей, при ограничениях на ресурсы состоит из следующей последовательности шагов.

Шаг 0. Определение исходных данных. Множество тестовых ИТВ — пустое ($T = \emptyset$). Счетчик m элементов ИТВ в множестве T равен нулю ($m = 0$). Множество тестовых ИТВ I включает в себя все рассматриваемые ИТВ. Затраты ресурса, необходимого для тестирования защищенности ОИ, равны нулю ($R_{тест} = 0$). Вводим ограничение на затраты ресурса $R_{гр}$ при проведении тестирования.

Рассчитываем сумму ущерба П по всем возможным комбинациям ИТВ $\{i\}$ потенциальных злоумышленников, уязвимостей $\{u\}$ элементов объекта $\{e\}$ и свойств ИБ $\{\sigma\}$:

$$\sum_{\substack{\forall\{i\},\forall\{u\},\\ \forall\{e\},\forall\{\sigma\}}} z(e,u,i,\sigma) = \Pi\,.$$

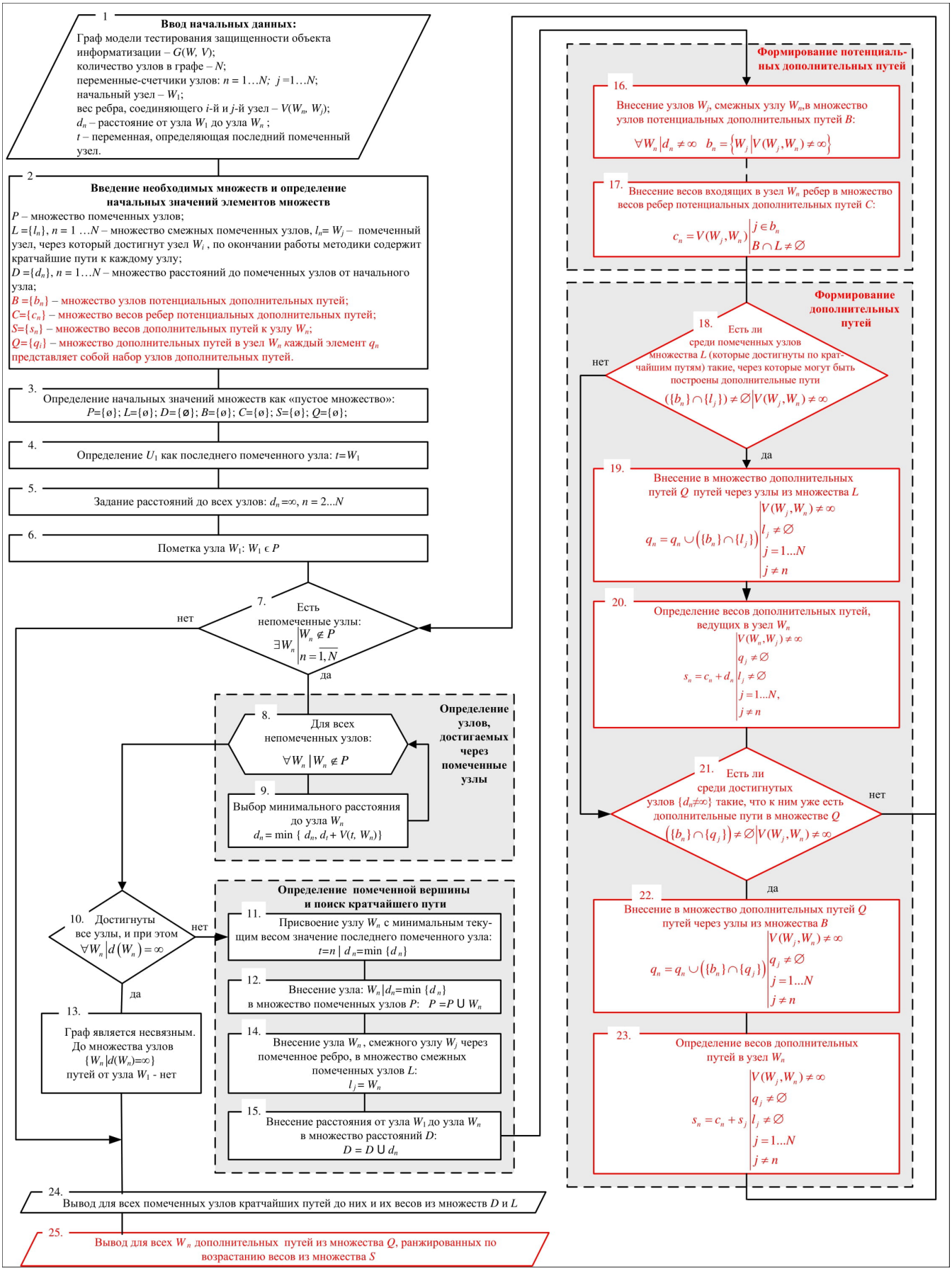

Рис. 3. Схема этапа формирования упорядоченного множества путей тестирования на основе модифицированного алгоритма Дейкстры

Шаг 1. Если множество рассматриваемых ИТВ не пустое ($I \neq \emptyset$), то из него выбирается ИТВ i_j, которое входит в путь q_k ($q_k \in Q$) в графе G с минимальным весом пути s_k ($s_k \in S$):

$$i_j = \{i\} \big| (s_k(q_k) = \min S) \wedge (i_j \in q_k) .$$

При первоначальном прогоне данного шага, множество I будет содержать все возможные ИТВ $\{i\}$ и будет выбран кратчайший путь q_0 в графе G с весом s_0. При дальнейших прогонах – множество I будет убывать, за счет исключения, а из множества Q будут последовательно выбираться дополнительные пути q_k из множества Q, имеющие наименьший вес s_k.

Шаг 2. Определяются затраты ресурса, необходимого на проведение ИТВ i_j. Значение ресурса r_j, расходуемого для проведения j-го ИТВ, для отдельных ребер графа G (рис. 2) пересчитываются из весов ребер $v(R, i_j)$ в соответствии с выражением:

$$r_j = v(R, i_j) \times \sum_{n=1}^{N_I} r_n ,$$

где r_j — затраты ресурса аудитора на проведение j-го тестового ИТВ;
r_n — затраты ресурса аудитора на проведение n-го тестового ИТВ;
N_I — количество тестовых ИТВ;
n — переменная-счетчик.

Шаг 3. Проверяется условие: если при добавлении в тестовый набор j-го ИТВ i_j сумма текущих затрат ресурса на проведение теста $R_{\text{тест}}$ и r_j меньше ограничения на затраты ресурса $R_{\text{гр}}$, то увеличиваем счетчик ИТВ в тестовом наборе на 1 ($m = m + 1$) и добавляем ИТВ i_j в тестовый набор ($t_m = i_j$, где $t_m \in T$) и продолжаем выполнение дальнейших операций. Если $R_{\text{тест}} + r_j > R_{\text{гр}}$, то ИТВ i_j в тестовый набор T не добавляется и из дальнейшего рассмотрения исключается ($I = I \setminus i_j$). В последнем случае возвращаемся к шагу 1.

Шаг 4. При принятии решения о добавлении ИТВ i_j в тестовый набор T в качестве элемента t_m выполняются следующие операции:

1. Производится оценка абсолютного значения ущерба π_m, который может быть выявлен m-м ИТВ в тестовом наборе, а также нарастающего итога по показателю $\pi = \sum_m \pi_m$. Для этого производится суммирование значений «стоимости» ущерба, который наносится ОИ при использовании ИТВ i_j, путем суммирования значений ущерба $z(e_k, \sigma_n)$, в тех путях $\{q \mid i_j \in q\}$, которые содержат в качестве вершины ИТВ i_j:

$$\pi_m = \sum_{(e_k \wedge \sigma_n) \in \{q | i_j \in q\}} z(e_k, \sigma_n) .$$

При этом значения ущерба $z(e_k, \sigma_n)$ для отдельных ребер графа G (рис. 2) пересчитываются из весов ребер $v(e_k, Z)$ в соответствии с выражением:

$$z(e_k, \sigma_n) = \left(\max_{\substack{l=1\ldots N_E \\ n=1\ldots 3}} \{z(e_l, \sigma_n)\} \right) - \left(v(e_k, Z) \times \sum_{l=1}^{N_E} \sum_{n=1}^{3} z(e_l, \sigma_n) \right) + 1 ,$$

где $z(e_k, \sigma_n)$ — «стоимость» ущерба, который наносится ОИ при нарушение σ_n-го свойства ИБ на его элементе e_k;
N_E — количество элементов ОИ, которое соответствует количеству элементов множества E;
$n = 1 \ldots 3$ — счетчик свойств ИБ σ_n;
$\sum_{l=1}^{N_E} \sum_{n=1}^{3} z(e_l, \sigma_n)$ — сумма ущерба по всем элементам ОИ и свойствам ИБ;
$\max\limits_{\substack{l=1\ldots N_E \\ n=1\ldots 3}} \{z(e_l, \sigma_n)\}$ — значение максимального ущерба среди всех комбинаций элементов и свойств ИБ.

2. Производится оценка относительного суммарного ущерба $\pi_{\text{отн } m}$, который может быть выявлен m-м ИТВ в тестовом наборе:

$$\pi_{\text{отн } m} = \frac{\pi_m}{\Pi} ,$$

а также оценка нарастающего итога по показателю:

$$\pi_{\text{отн}} = \sum_m \pi_{\text{отн } m} .$$

Шаг 5. Проверяются условия: если значение суммарного выявленного и потенциально предотвращенного ущерба π достаточно для заказчика тестирования либо относительное значение выявленного и потенциально предотвращенного ущерба $\pi_{\text{отн}} \to 100$ %, то процесс формирования тестового набора останавливается. Если вышеуказанные условия не выполняются, то выполняются дальнейшие операции.

Шаг 6. Производятся операции удаления тех путей тестирования (комбинаций $\{i, u, e, \sigma\}$), которые уже охвачены ИТВ, включенными в тестовый набор T.

1. Из графа G и из множества путей Q удаляются все пути $\{q | i_j \in q\}$, содержащие вершину i_j:

$$G = G \setminus \{q | i_j \in q\} ,$$

$$Q = Q \setminus \{q | i_j \in q\} .$$

2. Из множества весов путей S удаляются все значения весов путей $\{s(q) | i_j \in q\}$, которые содержат вершину i_j:

$$S = S \setminus \{s(q) | i_j \in q\} .$$

Шаг 7. Переход к шагу 1.

Общая схема методики с конкретизацией этапа выбора путей тестирования, обеспечивающих рациональную полноту оценки уязвимостей, при ограничениях на ресурсы представлена на рисунке 4.

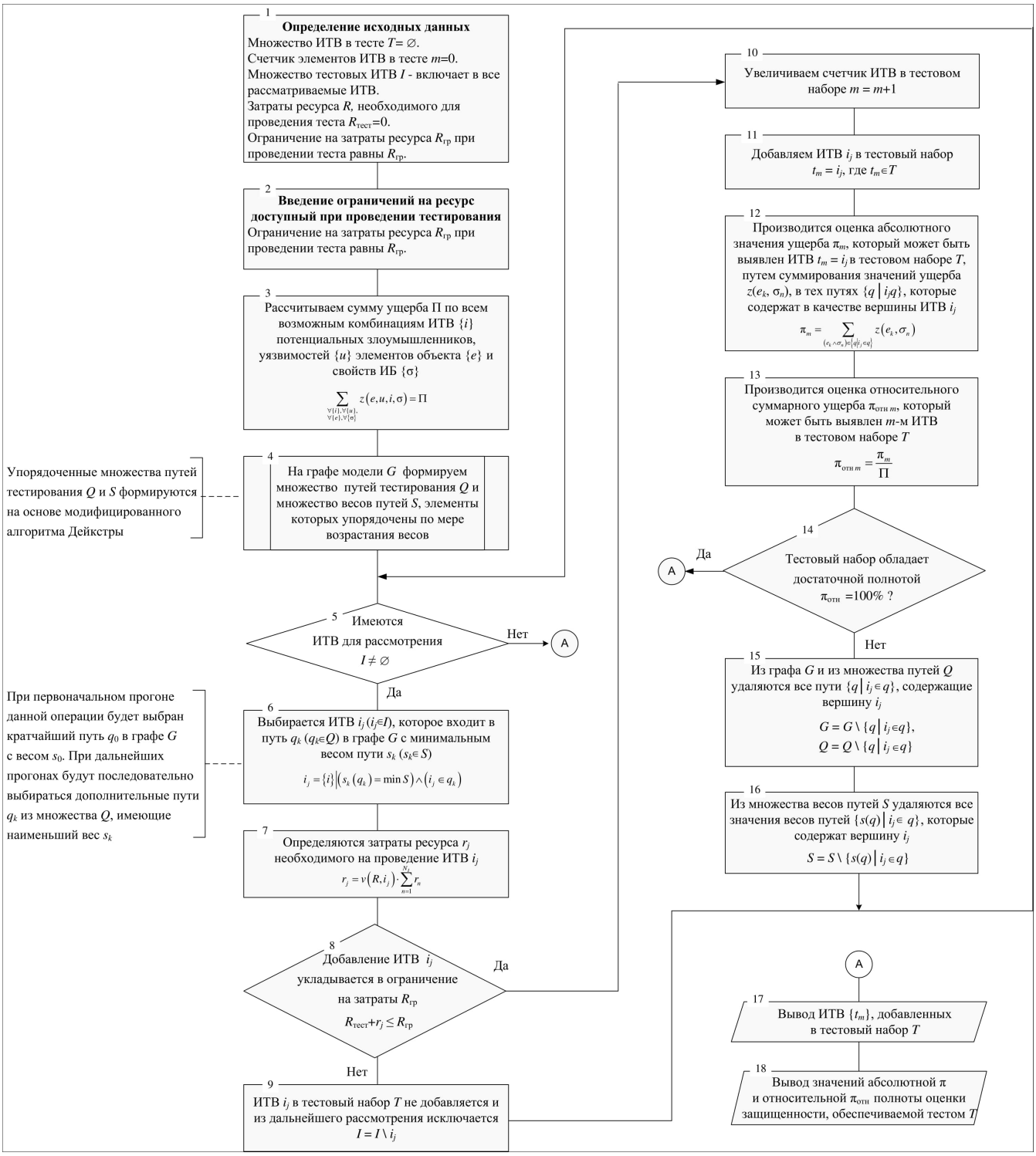

Рис. 4. Схема методики с конкретизацией этапа выбора путей тестирования, обеспечивающих рациональную полноту оценки уязвимостей, при ограничениях на ресурсы

Выводы

Представленная методика на первом этапе позволяет на основе модели тестирования защищенности ОИ ЖТ формировать множество путей тестирования с их ранжированием по степени повышения веса. При этом под весом пути понимается показатель «эффективность/стоимость» отдельной комбинации ресурса r_i, тестового ИТВ i, уязвимости u элемента ОИ ЖТ e и уровня ущерба $z(i, u, e, \sigma)$, наносимого ОИ S по свойству ИБ σ. На втором этапе методики производится выбор из кратчайшего пути и упорядоченного по возрастанию весов множества дополнительных путей такого ранжированного множества ИТВ $\{i\}$ и формирование из них тестового набора T, который бы обеспечивал максимизацию абсолютной суммарной стоимости обнаруженного ущерба $\pi \rightarrow \max$ (относительного значения $\pi_{отн} \rightarrow 100\ \%$) в рамках заданных ограничений на расход ресурса тестирования $R_{гр}$.

Элементами новизны данной методики, которые отличают ее от известных руководств по тестированию на проникновение [28], является то, что, во-первых, методика основана на модели тестирования защищенности ОИ, которая впервые разработана в данном исследовании, во-вторых, в состав методики введены оригинальные операции, которые на первом этапе методики за счет использования модификации известного алгоритма Дейкстры формируют упорядоченное множество путей тестирования, ранжированных по показателю «эффективность/стоимость», а на втором этапе — осуществляют формирование тестового набора из тех ИТВ, которые являются элементами «лучших» путей тестирования, таким образом, чтобы тестовый набор максимизировал абсолютную суммарную стоимость обнаруженного ущерба в рамках заданных ограничений на расход ресурса тестирования.

Данная методика предполагается к внедрению в автоматизированные комплексы тестирования защищенности ОИ ЖТ, архитектура и функциональность которых была изложена в работах [37, 38].

Литература

1. Рябцев, С. С. Метод выявления вредоносных роботов на основе данных процесса коллективного принятия решений в роевых робототехнических системах // Системы управления, связи и безопасности. 2022. № 3. С. 105–137. DOI: 10.24412/2410-9916-2022-3-105-137.

2. Будко, Н. П. Обзор графо-аналитических подходов к мониторингу информационно-телекоммуникационных сетей и их применение для выявления аномальных состояний / Н. П. Будко, Н. В. Васильев // Системы управления, связи и безопасности. 2021. № 6. С. 53–75.
DOI: 10.24412/2410-9916-2021-6-53-75.

3. Построение профиля атакующего на основе анализа сетевого трафика в критических инфраструктурах / Е. В. Федорченко, Е. С. Новикова, Д. А. Гайфулина, И. В. Котенко // Системы управления, связи и безопасности. 2021. № 6. С. 76–89. DOI: 10.24412/2410-9916-2021-6-76-89.

4. Обеспечение информационной безопасности киберфизических объектов на основе прогнозирования и обнаружения аномалий их состояния / В. И. Васильев, А. М. Вульфин, В. Е. Гвоздев, [и др.] // Системы управления, связи и безопасности. 2021. № 6. С. 90–119.
DOI: 10.24412/2410-9916-2021-6-90-119.

5. Израилов, К. Е. Модель классификации уязвимостей интерфейсов транспортной инфраструктуры «умного города» / К. Е. Израилов, Д. С. Левшун, А. А. Чечулин // Системы управления, связи и безопасности. 2021. № 5. С. 199–223. DOI: 10.24412/2410-9916-2021-5-199-223.

6. Горбачев, А. А. Модель функционирования и алгоритм проактивной защиты сервиса электронной почты от сетевой разведки / А. А. Горбачев, С. П. Соколовский, С. В. Усатиков // Системы управления, связи и безопасности. 2021. № 3. С. 60–109.
DOI: 10.24412/2410-9916-2021-3-60-109.

7. Методика оценки актуальных угроз и уязвимостей на основе технологий когнитивного моделирования и Text Mining / В. И. Васильев, А. М. Вульфин, А. Д. Кириллова, Н. В. Кучкарова // Системы управления, связи и безопасности. 2021. № 3. С. 110–134.
DOI: 10.24412/2410-9916-2021-3-110-134.

8. Методика разработки программы аудита информационной безопасности с учетом весовых коэффициентов значимости свидетельств аудита на основе метода анализа иерархий / В. А. Воеводин, П. В. Маркин, М. С. Маркина, Д. С. Буренок // Системы управления, связи и безопасности. 2021. № 2. С. 96–129.
DOI: 10.24412/2410-9916-2021-2-96-129.

9. Заколдаев, Д. А. Формальная модель обеспечения информационной безопасности при управлении ресурсами на производствах / Д. А. Заколдаев, А. Ю. Гришенцев // Системы управления, связи и безопасности. 2021. № 1. С. 33–61. DOI: 10.24411/2410-9916-2021-10102.

10. Санькова, Г. В. Информационные технологии в перевозочном процессе: Учебное пособие / Г. В. Санькова, Т. А. Одуденко. — Хабаровск: ДВГУПС, 2012. — 111 с.

11. Исаков, О. А. Вопросы совершенствования АСУ железнодорожного транспорта. — Саарбрюккен: Lambert Academic Publishing, 2011. — 224 с.

12. Методологические аспекты обеспечения информационной безопасности перевозочного процесса / С. Е. Ададуров, А. П. Глухов, А. А. Корниенко, Е. И. Белова // Управление товарными потоками и перевозочным процессом на железнодорожном транспорте на основе клиентоориентированности и логистических технологий: Коллективная монография членов и научных партнеров Объединенного ученого совета ОАО «РЖД» / под ред. Б. М. Лапидуса и А. Т. Осьминина. — Санкт-Петербург: ЛЕМА, 2019. — С. 251–263. — (Бюллетень Объединенного ученого совета ОАО «РЖД» № 4–6, 2019).

13. Котенко, И. В. Анализ защищенности инфраструктуры железнодорожного транспорта на основе аналитического моделирования / И. В. Котенко, А. А. Чечулин, Д. С. Левшун // Защита информации. Инсайд. 2017. № 6 (78). С. 48–57.

14. Определение уровня безопасности значимых объектов критической информационной инфраструктуры железнодорожного транспорта / А. П. Глухов, В. В. Василенко, А. А. Сидак, [и др.] // Двойные технологии. 2020. № 1 (90). С. 84–88.

15. Международная кибербезопасность на железнодорожном транспорте: методологические подходы и нормативная методическая база / С. Е. Ададуров, С. В. Диасамидзе, А. А. Корниенко, А. А. Сидак // Вестник Научно-исследовательского института железнодорожного транспорта. 2015. № 6. С. 9–15.

16. Котенко, И. В. Об архитектуре многоуровневой интеллектуальной системы обеспечения информационной безопасности автоматизированных систем на железнодорожном транспорте / И. В. Котенко, И. Б. Саенко // Методы и технические средства обеспечения безопасности информации (МиТСОБИ): Сборник материалов 23-й научно-технической конференции (Санкт-Петербург, Россия, 30 июня–03 июля 2014 г). — Санкт-Петербург: Изд-во Политехнического ун-та, 2014. — С. 97–98.

17. Котенко, И. В. Предложения по созданию многоуровневой интеллектуальной системы обеспечения информационной безопасности автоматизированных систем на железнодорожном транспорте / И. В. Котенко, И. Б. Саенко // Вестник Ростовского государственного университета путей сообщения (Вестник РГУПС). 2013. № 3 (51). С. 69–79.

18. Построение многоуровневой интеллектуальной системы обеспечения информационной безопасности для автоматизированных систем железнодорожного транспорта / И. В. Котенко, И. Б. Саенко, А. В. Чернов, М. А. Бутакова // Труды СПИИРАН. 2013. Вып. 7 (30). С. 7–25.

19. Управление безопасностью кибер-физических систем на основе оперативного ситуационного информирования об инцидентах / М. А. Бутакова, А. В. Чернов, П. С. Шевчук, С. М. Ковалев // Труды Ростовского государственного университета путей сообщения (Труды РГУПС). 2016. № 5. С. 14–16.

20. Методологические аспекты упреждающего управления информационной безопасностью железнодорожного транспорта / А. П. Глухов, Д. Н. Бирюков, В. В. Василенко, [и др.] // Двойные технологии. 2019. № 3 (88). С. 86–92.

21. О безопасности критической информационной инфраструктуры / С. Е. Ададуров, А. П. Глухов, А. А. Корниенко, Е. И. Белова // Автоматика, связь, информатика. 2020. № 4. С. 2–4. DOI: 10.34649/AT.2020.4.4.001.

22. Макаренко, С. И. Аудит информационной безопасности: основные этапы, концептуальные основы, классификация мероприятий // Системы управления, связи и безопасности. 2018. № 1. С. 1–29.
DOI: 10.24411/2410-9916-2018-10101.

23. Макаренко, С. И. Аудит безопасности критической инфраструктуры специальными информационными воздействиями: Монография. — Санкт-Петербург: Наукоемкие технологии, 2018. — 122 с.

24. Макаренко, С. И. Тестирование на проникновение на основе стандарта NIST SP 800-115 // Вопросы кибербезопасности. 2022. № 3 (49). С. 44–57. DOI: 10.21681/2311-3456-2022-3-44-57.

25. Макаренко, С. И. Модель аудита защищенности объекта критической информационной инфраструктуры тестовыми информационно-техническими воздействиями / С. И. Макаренко, Г. Е. Смирнов // Труды учебных заведений связи. 2021. Т. 7, № 1. С. 94–104. DOI: 10.31854/1813-324X-2021-7-1-94-104.

26. Макаренко, С. И. Методика обоснования тестовых информационно-технических воздействий, обеспечивающих рациональную полноту аудита защищенности объекта критической информационной инфраструктуры / С. И. Макаренко, Г. Е. Смирнов // Вопросы кибербезопасности. 2021. № 6 (46). С. 12–25.
DOI: 10.21681/2311-3456-2021-6-12-25.

27. Макаренко, С. И. Критерии и показатели оценки качества тестирования на проникновение // Вопросы кибербезопасности. 2021. № 3 (43). С. 43–57.
DOI: 10.21681/2311-3456-2021-3-43-57.

28. Макаренко, С. И. Анализ стандартов и методик тестирования на проникновение / С. И. Макаренко, Г. Е. Смирнов // Системы управления, связи и безопасности. 2020. № 4. С. 44–72. DOI: 10.24411/2410-9916-2020-10402.

29. Татт, У. Т. Теория графов = Graph Theory / У. Т. Татт; пер. с англ. Г. П. Гаврилова. — Москва: Мир. Редакция литературы по математическим наукам, 1988. — 424 с.

30. Свами, М. Графы, сети и алгоритмы = Graphs, Networks, and Algorithms / М. Свами, К. Тхуласираман; пер. с англ. М. В. Горбатовой, [и др.]; под ред. В. А. Горбатова. — Москва: Мир. Редакция литературы по новой технике, 1984. — 455 с.

31. Кормен, Т. Алгоритмы: построение и анализ = Introduction to Algorithms / Т. Кормен, Ч. Лейзерсон, Р. Ривест; пер. с англ. К. Белова, [и др.]. — Москва: МЦНМО, 2000. — 960 с. — (Классические учебники: Computer science).

32. Макаренко, С. И. Метод обеспечения устойчивости телекоммуникационной сети за счет использования ее топологической избыточности // Системы управления, связи и безопасности. 2018. № 3. С. 14–30.
DOI: 10.24411/2410-9916-2018-10302.

33. Цветков, К. Ю. Формирование резервных путей на основе алгоритма Дейкстры в целях повышения устойчивости информационно-телекоммуникационных сетей / К. Ю. Цветков, С. И. Макаренко, Р. Л. Михайлов // Информационно-управляющие системы. 2014. № 2 (69). С. 71–78.

34. Макаренко, С. И. Модифицированный алгоритм Беллмана-Форда с формированием кратчайших и резервных путей и его применение для повышения устойчивости телекоммуникационных систем / С. И. Макаренко, М. Н. Квасов // Инфокоммуникационные технологии. 2016. Т. 14, № 3. С. 264–274. DOI: 10.18469/ikt.2016.14.3.06.

35. Макаренко, С. И. Усовершенствованный протокол маршрутизации OSPF, обеспечивающий повышенную устойчивость сетей связи // Труды учебных заведений связи. 2018. Т. 4, № 2. С. 82–90.

36. Макаренко, С. И. Усовершенствование функций маршрутизации и сигнализации протокола PNNI с целью повышения устойчивости сети связи // Труды учебных заведений связи. 2020. Т. 6, № 2. С. 45–59.
DOI: 10.31854/1813-324X-2020-6-2-45-59.

37. Смирнов, Г. Е. Использование тестовых информационно-технических воздействий для аудита защищенности информационных систем железнодорожного транспорта / Г. Е. Смирнов, С. И. Макаренко // Интеллектуальные технологии на транспорте. 2020. № 3 (23). С. 20–29.

38. Смирнов, Г. Е. Использование тестовых информационно-технических воздействий для превентивного аудита защищенности информационно-телекоммуникационных сетей / Г. Е. Смирнов, С. И. Макаренко // Экономика и качество систем связи. 2020. № 3 (17). С. 43–58.

DOI: 10.24412/2413-2527-2022-331-5-18

Justification Method of Test Information-Technical Impacts for Security Analysis of Informatization Objects of Railway Transport

G. E. Smirnov
Saint Petersburg Electrotechnical University
Saint Petersburg, Russia
science.cybersec@yandex.ru

Abstract. **The article discusses the main information objects and automated systems of railway transport. It is shown that these systems are objects of a critical information infrastructure. In accordance with the legislation of the Russian Federation and the analysis of their real security is an important task. In the article is proposed to analyze the security of the objects and systems through using test information-technology impacts, which are predicted to be used by hackers. The justification method of information-technology impacts based on the Dijkstra's algorithm, which makes it possible to form a set of paths ranked by the total path metric is consisting of two stages: the stage of forming an ordered set of testing paths and the stage of choosing testing paths that ensure the rational completeness of vulnerability assessment with restrictions on resources.**

Keywords: **information security, audit, testing, Dijkstra's algorithm, information and technical impact, critical information infrastructure, informatization facility, railway transport.**

REFERENCES

1. Ryabtsev S. S. A Method for Detecting Byzantine Robots Based on Data from the Collective Decision-Making Process in Swarm Robotic Systems [Metod vyyavleniya vredonosnykh robotov na osnove dannykh protsessa kollektivnogo prinyatiya resheniy v roevykh robototekhnicheskikh sistemakh], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2022, No. 3, Pp. 105–137.
DOI: 10.24412/2410-9916-2022-3-105-137.

2. Budko N. P., Vasiliev N. V. Review of Graph-Analytical Approaches to Monitoring of Information and Telecommunication Networks and Their Application to Identify Abnormal States [Obzor grafo-analiticheskikh podkhodov k monitoringu informatsionno-telekommunikatsionnykh setey i ikh primenenie dlya vyyavleniya anomalnykh sostoyaniy], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2021, No. 6, Pp. 53–75.
DOI: 10.24412/2410-9916-2021-6-53-75.

3. Fedorchenko E. V., Novikova E. S., Gaifulina D. A., Kotenko I. V. Attacker Profiling Based on the Network Traffic Analysis [Postroenie profilya atakuyushchego na osnove analiza setevogo trafika v kriticheskikh infrastrukturakh], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2021, No. 6, Pp. 76–89.
DOI: 10.24412/2410-9916-2021-6-76-89.

4. Vasilyev V. I., Vulfin A. M., Gvozdev V. E., et al. Ensuring Information Security of Cyber-Physical Objects Based on Predicting and Detecting Anomalies in Their State [Obespechenie informatsionnoy bezopasnosti kiberfizicheskikh obektov na osnove prognozirovaniya i obnaruzheniya anomaliy ikh sostoyaniya], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2021, No. 6, Pp. 90–119. DOI: 10.24412/2410-9916-2021-6-90-119.

5. Izrailov K. E., Levshun D. S., Chechulin A. A. Vulnerability Classification Model for Smart City Transport Infrastructure Interfaces [Model klassifikatsii uyazvimostey interfeysov transportnoy infrastruktury «umnogo goroda»], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2021, No. 5, Pp. 199–223.
DOI: 10.24412/2410-9916-2021-5-199-223.

6. Gorbachev A. A., Sokolovsky S. P., Usatikov S. V. Functioning Model and Algorithm of Email Service Proactive Protection from Network Intelligence [Model funktsionirovaniya i algoritm proaktivnoy zashchity servisa elektronnoy pochty ot setevoy razvedki], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2021, No. 3, Pp. 60–109. DOI: 10.24412/2410-9916-2021-3-60-109.

7. Vasilyev V. I., Vulfin A. M., Kirillova A. D., Kuchkarova N. V. Methodology for Assessing Current Threats and Vulnerabilities Based on Cognitive Modeling Technologies and Text Mining [Metodika otsenki aktualnykh ugroz i uyazvimostey na osnove tekhnologiy kognitivnogo modelirovaniya i Text Mining], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2021, No. 3, Pp. 110–134.
DOI: 10.24412/2410-9916-2021-3-110-134.

8. Voevodin V. A., Markin P. V., Markina M. S., Burenok D. S. Technique for Developing an Information Security Audit Program Taking into Account the Weight Coefficients of Certificates Audit Based on the Hierarchy Analysis Method [Metodika razrabotki programmy audita informatsionnoy bezopasnosti s uchetom vesovykh koeffitsientov znachimosti svidetelstv audita na osnove metoda analiza ierarkhiy], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2021, No. 2, Pp. 96–129.
DOI: 10.24412/2410-9916-2021-2-96-129.

9. Zakoldaev D. A., Grishentsev A. Y. Formal Model of Information Security in the Management of Resources in Production [Formalnaya model obespecheniya informatsionnoy bezopasnosti pri upravlenii resursami na proizvodstvakh], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2021, No. 1, Pp. 33–61.
DOI: 10.24411/2410-9916-2021-10102.

10. Sankova G. V., Odudenko T. A. Information technologies in the transportation process: Study guide [Informatsionnye tekhnologii v perevozochnom protsesse: Uchebnoe

posobie]. Khabarovsk, Far Eastern State Transport University, 2012, 111 p.

11. Isakov O. A. Questions of improving the automated control system of railway transport [Voprosy sovershenstvovaniya ASU zheleznodorozhnogo transporta]. Saarbrücken, LAP Lambert Academic Publishing, 2011, 224 p.

12. Adadurov S. E., Glukhov A. P., Kornienko A. A., Belova E. I. Methodological Aspects of Ensuring Information Security of the Transportation Process [Metodologicheskie aspekty obespecheniya informatsionnoy bezopasnosti perevozochnogo protsessa]. *In: Lapidus B. M., Osminin A. T. (eds) Management of goods flows and the transportation process on railway transport based on customer orientation and logistics technologies: A collective monograph of members and scientific partners of the Joint Scientific Council of Russian Railways JSC [Upravlenie tovarnymi potokami i perevozochnym protsessom na zheleznodorozhnom transporte na osnove klientoorientirovannosti i logisticheskikh tekhnologiy: kollektivnaya monografiya chlenov i nauchnykh partnerov Obedinennogo uchenogo soveta OAO «RZhD»]*. Saint Petersburg, LEMA Publishing House, 2019, Pp. 251–263.

13. Kotenko I. V., Chechulin A. A., Levshun D. S. Security Analysis of Railway Transport Infrastructure on the Base of Analytical Modeling [Analiz zashchishchennosti infrastruktury zheleznodorozhnogo transporta na osnove analiticheskogo modelirovaniya], *Zaŝita Informacii. Inside [Zashchita informatsii. Insayd]*, 2017, No. 6 (78), Pp. 48–57.

14. Gluhov A. P., Vasilenko V. V., Sidak A. A., et al. Determination of the Security Level of Significative Objects of Critical Information Infrastructure of Railway Transport [Opredelenie urovnya bezopasnosti znachimykh obektov kriticheskoy informatsionnoy infrastruktury zheleznodorozhnogo transporta], *Dual Technologies [Dvoynye tekhnologii]*, 2020, No. 1 (90), Pp. 84–88.

15. Adadurov S. E., Diasamidze S. V., Kornienko A. A., Sidak A. A. International Cybersecurity on Railway Transport: Methodological Approaches and Normal Procedural Framework [Mezhdunarodnaya kiberbezopasnost na zheleznodorozhnom transporte: metodologicheskie podkhody i normativnaya metodicheskaya baza], *Russian Railway Science Journal [Vestnik Nauchno-issledovatelskogo instituta zheleznodorozhnogo transporta]*, 2015, No. 6, Pp. 9–15.

16. Kotenko I. V., Saenko I. B. On the Architecture of a Multi-Level Intelligent Information Security System of Automated Systems in Railway Transport [Ob arkhitekture mnogourovnevoy intellektualnoy sistemy obespecheniya informatsionnoy bezopasnosti avtomatizirovannykh sistem na zheleznodorozhnom transporte], *Methods and Technical Means of Information Security(MiTSOBI): Collection of Materials of the 23rd Scientific and Technical Conference [Metody i tekhnicheskie sredstva obespecheniya bezopasnosti informatsii (MiTSOBI): Sbornik materialov 23-y nauchno-tekhnicheskoy konferentsii], Saint Petersburg, Russia, June 30–July 03, 2014*. Saint Peterburg, St. Petersburg State Polytechnic University, 2014, Pp. 97–98.

17. Kotenko I. V., Saenko I. B. Proposals on Creation of a Multi-Level Intelligent Information Security System of Automated Systems on Railway Transport [Predlozheniya po sozdaniyu mnogourovnevoy intellektualnoy sistemy obespecheniya informatsionnoy bezopasnosti avtomatizirovannykh sistem na zheleznodorozhnom transporte], *Vestnik Rostovskogo Gosudarstvennogo Universiteta Putey Soobshcheniya (Vestnik RGUPS),* 2013, No. 3 (51), Pp. 69–79.

18. Kotenko I. V., Saenko I. B., Chernov A. V., Butakova M. A. The Construction of a Multi-Level Intelligent Information Security System for Automated Systems of Railway Transport [Postroenie mnogourovnevoy intellektualnoy sistemy obespecheniya informatsionnoy bezopasnosti dlya avtomatizirovannykh sistem zheleznodorozhnogo transporta], *SPIIRAS Proceedings [Trudy SPIIRAN]*, 2013, Is. 7 (30), Pp. 7–25.

19. Butakova M. A., Chernov A. V., Shevchuk P. S., Kovalev S. M. Cyber-Physical Systems Security Management Based on Operational Situational Incidents Information [Upravlenie bezopasnostyu kiber-fizicheskikh sistem na osnove operativnogo situatsionnogo informirovaniya ob intsidentakh], *Trudy Rostovskogo gosudarstvennogo universiteta putej soobshcheniya (Trudy RGUPS)*, 2016, No. 5, Pp. 14–16.

20. Glukhov A. P., Biryukov D. N., Vasilenko V. V., et al. Methodological Aspects of Proactive Management of Railroad Transport Information Security [Metodologicheskie aspekty uprezhdayushchego upravleniya informatsionnoy bezopasnostyu zheleznodorozhnogo transporta], *Dual Technologies [Dvoynye tekhnologii]*, 2019, No. 3 (88), Pp. 86–92.

21. Adadurov S. E., Glukhov A. P., Kornienko A. A., Belova E. I. Principles of Railway Transport Critical Information Infrastructure Security Supporting [O bezopasnosti kriticheskoy informatsionnoy infrastruktury], *Automation, Communications, Informatics [Avtomatika, svyaz, informatika]*, 2020, No. 4, Pp. 2–4. DOI: 10.34649/AT.2020.4.4.001.

22. Makarenko S. I. Audit of Information Security — The Main Stages, Conceptual Framework, Classification of Types [Audit informatsionnoy bezopasnosti: osnovnye etapy, kontseptualnye osnovy, klassifikatsiya meropriyatiy], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2018, No. 1, Pp. 1–29.
DOI: 10.24411/2410-9916-2018-10101.

23. Makarenko S. I. Security audit of critical infrastructure with special information impacts: Monograph [Audit bezopasnosti kriticheskoy infrastruktury spetsialnymi informatsionnymi vozdeystviyami: Monografiya]. Saint Petersburg, Naukoemkie tekhnologii Publishing House, 2018, 122 p.

24. Makarenko S. I. Penetration Testing in Accordance with NIST SP 800-115 Standard [Testirovanie na proniknovenie na osnove standarta NIST SP 800-115], *Cybersecurity Issues [Voprosy kiberbezopasnosti]*, 2022, No. 3 (49), Pp. 44–57.
DOI: 10.21681/2311-3456-2022-3-44-57.

25. Makarenko S. I., Smirnov G. E. Model of Security Audit of a Critical Information Infrastructure Object with Use the Test Cyber Attacks [Model audita zashchishchennosti obekta kriticheskoy informatsionnoy infrastruktury testovymi informatsionno-tekhnicheskimi vozdeystviyami], *Proceedings of Telecommunication Universities [Trudy uchebnykh zavedeniy svyazi]*, 2021, Vol. 7, No. 1, Pp. 94–104.
DOI: 10.31854/1813-324X-2021-7-1-94-104.

26. Makarenko S. I., Smirnov G. E. Selection Method of Test Cyber Attacks That Ensure the Rational Completeness of the Penetration Testing of a Critical Information Infrastructure Object [Metodika obosnovaniya testovykh informatsionno-tekhnicheskikh vozdeystviy, obespechivayushchikh ratsionalnuyu polnotu audita zashchishchennosti obekta kriticheskoy informatsionnoy infrastruktury], *Cybersecurity Issues [Voprosy kiberbezopasnosti]*, 2021, No. 6 (46), Pp. 12–25.

DOI: 10.21681/2311-3456-2021-6-12-25.

27. Makarenko S. I. Criteria and Parameters for Estimating Quality of Penetration Testing [Kriterii i pokazateli otsenki kachestva testirovaniya na proniknovenie], *Cybersecurity Issues [Voprosy kiberbezopasnosti]*, 2021, No. 3 (43), Pp. 43–57. DOI: 10.21681/2311-3456-2021-3-43-57.

28. Makarenko S. I., Smirnov G. E. Analysis of Penetration Testing Standards and Methodologies [Analiz standartov i metodik testirovaniya na proniknovenie], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2020, No. 4, Pp. 44–72.
DOI: 10.24411/2410-9916-2020-10402.

29. Tutte W. T. Graph Theory [Teoriya grafov]. Moscow, Mir Publishers, 1988, 424 p.

30. Swamy M. N. S., Thulasiraman K. Graphs, Networks, and Algorithms [Grafy, seti i algoritmy]. Moscow, Mir Publishers, 1984, 455 p.

31. Cormen T. H., Leiserson C. E., Rivest R. L. Introduction to Algorithms [Algoritmy: postroenie i analiz]. Moscow, Moscow Center for Continuous Mathematical Education, 2000, 960 p.

32. Makarenko S. I. Stability Method of Telecommunication Network with Using Topological Redundancy [Metod obespecheniya ustoychivosti telekommunikatsionnoy seti za schet ispolzovaniya ee topologicheskoy izbytochnosti], *Systems of Control, Communication and Security [Sistemy upravleniya, svyazi i bezopasnosti]*, 2018, No. 3, Pp. 14–30.
DOI: 10.24411/2410-9916-2018-10302.

33. Tsvetcov K. U., Makarenko S. I., Mikhailov R. L. Forming Reserve Paths Based on Dijkstra Algorithm in Order to Enhance Stability of Telecommunication Networks [Formirovanie rezervnykh putey na osnove algoritma Deykstry v tselyakh povysheniya ustoychivosti informatsionno-telekommunikatsionnykh setey], *Information and Control Systems [Informatsionno-upravlyayushchie sistemy]*, 2014, No. 2 (69), Pp. 71–78.

34. Makarenko S. I., Kvasov M. N. Modified Bellman-Ford Algorithm with Forming the Shortest and Fallback Paths and Its Application for Telecommunication Network Stability Improvement [Modifitsirovannyy algoritm Bellmana-Forda s formirovaniem kratchayshikh i rezervnykh putey i ego primenenie dlya povysheniya ustoychivosti telekommunikatsionnykh sistem], *Infocommunication Technologies [Infokommunikatsionnye tekhnologii]*, 2016, Vol.14, No. 3, Pp. 264–274.
DOI: 10.18469/ikt.2016.14.3.06.

35. Makarenko S. I. The Improved OSPF Protocol for High Network Stability [Usovershenstvovannyy protokol marshrutizatsii OSPF, obespechivayushchiy povyshennuyu ustoychivost setey svyazi], *Proceedings of Telecommunication Universities [Trudy uchebnykh zavedeniy svyazi]*, 2018, Vol. 4, No. 2, Pp. 82–90.

36. Makarenko S. I. Improved Routing and Signaling Functions of PNNI Protocol for High Network Stability [Usovershenstvovanie funktsiy marshrutizatsii i signalizatsii protokola PNNI s tselyu povysheniya ustoychivosti seti svyazi], *Proceedings of Telecommunication Universities [Trudy uchebnykh zavedeniy svyazi]*, 2020, Vol. 6, No. 2, Pp. 45–59. DOI: 10.31854/1813-324X-2020-6-2-45-59.

37. Smirnov G. E., Makarenko S. I. The Use of Test Information and Technical Impacts for Security Audit of Information Systems of Railway Transport [Ispolzovanie testovykh informatsionno-tekhnicheskikh vozdeystviy dlya audita zashchishchennosti informatsionnykh sistem zheleznodorozhnogo transporta], *Intellectual Technologies on Transport [Intellektualnye tekhnologii na transporte]*, 2020, No. 3 (23), Pp. 20–29.

38. Smirnov G. E., Makarenko S. I. The use of test information and technical impacts for preventive audit security audit of information and telecommunication networks [Ispolzovanie testovykh informatsionno-tekhnicheskikh vozdeystviy dlya preventivnogo audita zashchishchennosti informatsionno-telekommunikatsionnykh setey], *Economics and quality of communication systems [Ekonomika i kachestvo sistem svyazi]*, 2020, No. 3 (17), Pp. 43–58.

DOI: 10.24412/2413-2527-2022-331-19-24

Об ожидаемом размере подмножества Парето случайного множества точек

к.ф.-м.н. А. Н. Баушев
Петербургский государственный университет
путей сообщения Императора Александра I
Санкт-Петербург, Россия
baushev@pgups.ru

к.ф.-м.н. О. Л. Семёнова
Санкт-Петербургский государственный университет
Санкт-Петербург, Россия
o_semenova@mail.ru

Аннотация. **Статья посвящена оцениванию математического ожидания мощности множества Парето в конечном множестве точек, образованном симметрично зависимыми случайными элементами пространства $\mathbb{R}^d$. Получены оценки, аналогичные хорошо известным оценкам для случая независимых случайных векторов.**

Ключевые слова: **случайное множество точек, множество Парето, симметрично зависимые случайные величины, случайные перестановки, рекорды в последовательности случайных величин.**

Введение

Задача построения и исследования множества Парето играет ключевую роль в современных подходах к решению задач многокритериальной оптимизации [1–8]. Ей посвящено большое количество работ, которые можно разбить на два класса, в зависимости от того, конечным или бесконечным является рассматриваемое исходное множество, с незначительными отличиями в терминологии. Мы будем иметь дело со случаем конечного исходного множества и говорить о *множестве Парето,* а не о «границе Парето» или о «фронте Парето», как это принято в работах, в которых исходное множество описывается как подмножество пространства $\mathbb{R}^d$, точки которого удовлетворяют определенным аналитическим соотношениям [5].

Отметим, что несмотря на то, что можно легко построить примеры, в которых множество Парето совпадает с исходным множеством, в типичных ситуациях множество Парето имеет существенно меньшую мощность, чем исходное множество, как это было показано в пионерской работе [9]. «Типичная ситуация» в этой работе определялась как ситуация, в которой множество весов рассматриваемых объектов может быть интерпретировано как результат процедуры независимого выбора из вероятностного распределения в пространстве $\mathbb{R}^d$ с независимыми и одинаково распределенными координатами, имеющими непрерывную функцию распределения.

Заметим, что случай, когда координаты независимы, но имеют различные непрерывные строго монотонные функции распределения $F_i,\ i = 1, \dots, d$ при помощи преобразования Смирнова $(x_1, \dots, x_d) \to \big(F_1(x_1), \dots, F_d(x_d)\big)$ сводится к случаю независимых координат, имеющих равномерное распределение на отрезке [0,1], поскольку такое преобразование, помимо прочего, сохраняет отношение частичного порядка на $\mathbb{R}^d$.

Однако попытки отказаться от «независимости» приводят к необходимости так или иначе описывать модели зависимостей между элементами исходного множества. Отметим, что имеются два направления для обобщения результатов работы [9]. В первом из них происходит отказ от условия независимости координат, но сохраняется условие независимого выбора при генерировании исходного множества, а во втором, наоборот, сохраняется условие независимости координат, но при этом исходное случайное множество интерпретируется как реализация случайного вектора с $\mathbb{R}^d$-значными компонентами или, другими словами, как случайная $n \times d$ матрица с независимыми столбцами, где n — число точек в исходном множестве.

В настоящей работе мы примыкаем ко второму из этих направлений, предполагая, что строки матрицы представляют собой симметрично зависимые $\mathbb{R}^d$-значные случайные векторы, а столбцы независимы.

Симметрично зависимые случайные величины и векторы естественным образом возникают в различных прикладных задачах [9]. Приводимая далее лемма 3 также в какой-то мере описывает достаточно широкий спектр ситуаций, в которых могут появляться симметрично зависимые случайные величины.

Основные результаты нашей работы оказываются вполне аналогичными результатам работы [9]. Однако мы использовали другой подход, основанный на теории рекордов [7, 10, 11], который позволил обобщить результаты [9] на случай симметрично зависимых случайных векторов.

Предварительные результаты

Пусть $(X, \leq)$ — частично упорядоченное множество. Множеством *минимальных элементов* множества X или *множеством Парето* для множества X называется множество $Min(X) = \{x \in X | \nexists\, y \in X: y < x\}$.

Отображение $\varphi: Y \mapsto Min(Y)$, определенное на совокупности подмножеств множества X, называется *фильтрацией* совокупности подмножеств отношением частичного порядка $\leq$. Неподвижные точки этого отображения, то есть подмножества $Y \subset X$, состоящие из попарно несравнимых элементов, называются *фильтрованными* множествами. Таким образом, множество Парето для множества X — это максимальное по включению фильтрованное подмножество множества X.

В дальнейшем мы будем иметь дело с естественным отношением частичного порядка на пространстве $\mathbb{R}^d$. По определению $x = (x_1, \dots, x_d) \leq y = (y_1, \dots, y_d)$ равносильно выполнению системы неравенств $x_1 \leq y_1, \dots, x_d \leq y_d$. Мы также будем иметь дело с отношениями частичного порядка $\leq_k$, для которых неравенство

$$x = (x_1, \dots, x_d) \leq_k y = (y_1, \dots, y_d)$$

равносильно выполнению системы неравенств $x_k \le y_k, \dots, x_d \le y_d$. Через φ_k мы будем обозначать соответствующие фильтрации при $k = 2, \dots, d$.

Мы будем рассматривать множества $X \subset \mathbb{R}^d$, обладающие следующим свойством *координатной единственности*: для любых $x, y \in X$ и любого $j \in \{1, \dots, d\}$ выполняется одно из неравенств $x_j < y_j$ или $y_j < x_j$. Такие множества мы будем называть U-множествами (англ. *Uniqueness* — «уникальность, единственность»).

Лемма 1. Пусть элементы U-множества $X \subset \mathbb{R}^d$ занумерованы по возрастанию первой координаты: $X = \{x^1, \dots x^n\}$. Положим $X_k = \{x^1, \dots, x^k\}$, $k = 2, \dots n$. Для того, чтобы $x^k \in \varphi(X)$ необходимо и достаточно, чтобы $x^k \in \varphi_2(X_k)$.

Доказательство.

1) Необходимость. Если $x^k \notin \varphi_2(X_k)$, то найдется $r < k$ такое, что $x_2^r \le x_2^k, \dots, x_d^r \le x_d^k$, причем по крайней мере одно из этих неравенств является строгим. Поскольку $x_1^r \le x_1^k$, отсюда следует, что $x^k \notin \varphi(X)$.

2) Достаточность. Доказательство от противного. Предположим, что $x^k \in \varphi_2(X_k)$, но $x^k \notin \varphi(X)$. Тогда найдется $r \in \{1, \dots, n\} \backslash \{k\}$ такое, что $x_1^r \le x_1^k, \dots, x_d^r \le x_d^k$, а по условию координатной единственности все эти неравенства являются строгими. Так как элементы занумерованы по возрастанию первой координаты, то должно выполняться неравенство $r < k$, но тогда получаем противоречие с условием $x^k \in \varphi_2(X_k)$. ■

Следствие. Пусть элементы U-множества $X \subset \mathbb{R}^d$ занумерованы по возрастанию j-й координаты:

$$1 \le j \le d-1 : X = \{x^1, \dots, x^n\}, X_k = \{x^1, \dots, x^k\}, k = 2, \dots n.$$

Для того, чтобы $x^k \in \varphi_j(X)$ необходимо и достаточно, чтобы $x^k \in \varphi_{j+1}(X_k)$.

Доказательство следствия вполне аналогично доказательству леммы 1.

Лемма 1 служит основой для рекурсивной конструкции, которая используется в дальнейшем при построении множества Парето. Граничные условия для этой конструкции описываются леммой 2.

Лемма 2. Пусть элементы U-множества $X \subset \mathbb{R}^2$ занумерованы по возрастанию первой координаты: $X = \{x^1, \dots, x^n\}$. Для того, чтобы множество X было фильтрованным, необходимо и достаточно выполнения неравенств $x_2^1 > x_2^2 > \dots > x_2^n$.

Доказательство леммы 2 непосредственно вытекает из определения фильтрованного множества.

Пусть элементы U-множества $X = \{x^1, \dots, x^n\} \subset \mathbb{R}^2$ занумерованы по возрастанию первой координаты. Элемент x^k называется *рекордным* если $x_2^k < \min\{x_2^1, \dots, x_2^{k-1}\}$. По определению элемент x^1 считается рекордным.

Следствие. Мощность подмножества Парето в U-множестве $X \subset \mathbb{R}^2$ совпадает с числом рекордных элементов в последовательности $x^1, \dots, x^n$, полученной в результате сортировки множества X по первой координате.

В дальнейшем мы будем иметь дело с $\mathbb{R}^d$-значными случайными векторами $X_1, \dots, X_n$. Для обозначения вероятностей событий и математических ожиданий случайных величин мы будем использовать соответственно символы $\mathbf{P}\{\cdot\}$ и $\mathbf{E}(\cdot)$, а для совместного распределения векторов $X_1, \dots, X_n$ использовать символ $P_{X_1,\dots,X_n}$.

Обозначим через Π множество всех перестановок множества $\{1, 2, \dots, n\}$, то есть взаимно однозначных отображений множества $\{1, 2, \dots, n\}$ на себя.

Случайные векторы $X_1, \dots, X_n$ называются *симметрично зависимыми*, если их совместное распределение инвариантно относительно любой перестановки индексного множества, то есть если $P_{X_{\pi(1)},\dots,X_{\pi(n)}} = P_{X_1,\dots,X_n}$ для любой $\pi = (\pi(1), \dots, \pi(n)) \in \Pi$.

Отметим некоторые простейшие факты, связанные с понятием симметричной зависимости.

Предложение 1. Пусть $X_1, \dots, X_n$ — $\mathbb{R}^d$-значные случайные векторы.

1. Для того, чтобы $X_1, \dots, X_n$ были симметрично зависимыми необходимо и достаточно, чтобы их совместное распределение было инвариантным относительно любой транспозиции (перестановки, меняющей местами только два индекса).

2. Если $X_1, \dots, X_n$ симметрично зависимы, то для любого непустого подмножества $I \subset \{1, \dots, n\}$ случайные величины $(X_i)_{i \in I}$ также симметрично зависимы.

3. Если $X_1, \dots, X_n$ независимы и одинаково распределены, то $X_1, \dots, X_n$ симметрично зависимы.

4. Если $X_1, \dots, X_n$ симметрично зависимы и X_1 не зависит от $X_2, \dots, X_n$, то $X_1, \dots, X_n$ независимы и одинаково распределены.

5. Если $X_1, \dots, X_n$ симметрично зависимы и $\mathbb{R}^d$-значный случайный вектор Y не зависит от $X_1, \dots, X_n$, то случайные векторы $Z_1 = Y + X_1, \dots, Z_n = Y + X_n$ симметрично зависимы.

6. Если $X_1, \dots, X_n$ симметрично зависимы, то для любой измеримой функции $f: \mathbb{R}^d \to \mathbb{R}$ случайные величины $f(X_1), \dots, f(X_n)$ симметрично зависимы. В частности, координаты $X_{1j}, \dots, X_{nj}$ симметрично зависимы для каждого $j \in \{1, \dots, d\}$.

Случайные величины со значениями в множестве Π называются *случайными перестановками* элементов множества $\{1, 2, \dots, n\}$. В дальнейшем мы будем использовать термин «случайная перестановка» в более узком смысле, подразумевая, что случайная перестановка σ имеет равномерное распределение на множестве Π, то есть что $\mathbf{P}\{\sigma = \pi\} = 1/n!$ для любой $\pi \in \Pi$.

Предложение 2. Пусть $\pi \in \Pi$, σ, σ' — случайные перестановки элементов множества $\{1, 2, \dots, n\}$. Тогда:

1. $\pi\sigma$ и $\sigma\pi$ — случайные перестановки.

2. Если σ и σ' независимы, то $\sigma\sigma'$ и $\sigma'\sigma$ — случайные перестановки.

Доказательство. Докажем п. 2. Зафиксируем $\pi' \in \Pi$. По формуле полной вероятности имеем

$$\mathbf{P}\{\sigma\sigma' = \pi'\} = \sum_{\pi \in \Pi} \mathbf{P}\{\sigma\sigma' = \pi' | \sigma = \pi\} \mathbf{P}\{\sigma = \pi\} =$$

$$= \sum_{\pi \in \Pi} \mathbf{P}\{\sigma' = \pi'\pi^{-1}\} \frac{\mathbf{1}}{\boldsymbol{n!}} = \frac{1}{n!}$$

так как $\pi'\pi^{-1}$ так же пробегает множество Π, когда его пробегает π. ■

Лемма 3. Пусть $\mathrm{X}_1, \dots, \mathrm{X}_n$ — случайные $\mathbb{R}^d$-значные векторы с произвольным совместным распределением и пусть σ — случайная перестановка индексов, не зависящая от $\mathrm{X}_1, \dots, \mathrm{X}_n$. Тогда случайные векторы $\mathrm{X}_{\sigma(1)}, \dots, \mathrm{X}_{\sigma(n)}$ симметрично зависимы.

Доказательство. Для любых борелевских множеств $B_1, \dots, B_n \subset \mathbb{R}^d$ и любой перестановки $\pi' \in \Pi$ по формуле полной вероятности мы имеем равенства

$$\mathbf{P}\{\mathrm{X}_{\pi'\sigma(1)} \in B_1, \dots, \mathrm{X}_{\pi'\sigma(n)} \in B_n\} =$$

$$= \sum_{\pi \in \Pi} \mathbf{P}\{\mathrm{X}_{\pi'\sigma(1)} \in B_1, \dots, \mathrm{X}_{\pi'\sigma(n)} \in B_n | \sigma = \pi\}\mathbf{P}\{\sigma = \pi\} =$$

$$= \frac{1}{n!} \sum_{\pi \in \Pi} \mathbf{P}\{\mathrm{X}_{\pi'\pi(1)} \in B_1, \dots, \mathrm{X}_{\pi'\pi(n)} \in B_n\}.$$

Когда π пробегает множество Π, $\pi'\pi$ также пробегает множество Π, следовательно, правая часть последнего равенства не зависит от π'. ■

С последовательностью $\mathbb{R}^d$-значных случайных векторов $\mathrm{X}_1, \dots, \mathrm{X}_n$ мы будем связывать матрицу их координат $(X_{ij} | \, i = 1, \dots, n; j = 1, \dots, d)$, в которой индекс строки — номер вектора в последовательности, а индекс столбца — номер координаты.

Будем говорить, что случайные $\mathbb{R}^d$-значные векторы обладают свойством *стохастической координатной единственности* и называть их *SU*-векторами (Stochastic Uniqueness), если $\mathbf{P}\{X_{\boldsymbol{kj}} = X_{lj}\} = 0$ для любых различных $k, l \in \{1, \dots, n\}$ и для любого $j \in \{1, \dots, d\}$.

Лемма 4. Пусть $\mathrm{X}_1, \dots, \mathrm{X}_n$ — $\mathbb{R}^d$-значные симметрично зависимые случайные *SU*-векторы, σ_j — перестановка их индексов в соответствии с возрастанием *j*-й координаты ($j \in \{1, \dots, d\}$). Тогда σ_j — случайная перестановка, то есть σ_j имеет равномерное распределение на множестве Π.

Доказательство. Зафиксируем $\pi \in \Pi$. Случайное событие $\{\sigma_j = \pi\}$ происходит тогда и только тогда, когда происходит событие $\{X_{\pi(1),j} < \dots < X_{\pi(n),j}\}$. Из условия симметричной зависимости следует, что вероятность последнего события не зависит от π, а из условия стохастической координатной единственности — что

$$\sum_{\pi' \in \Pi} \mathbf{P}\{X_{\pi'(1),j} < \dots < X_{\pi'(n),j}\} = 1.$$

Таким образом $n!\ \mathbf{P}\{\sigma_j = \pi\} = 1$, что и требовалось. ■

Замечание. Свойство стохастической координатной единственности играет существенную роль при выводе основных результатов. Для $\mathbb{R}^d$-значных симметрично зависимых случайных векторов $\mathrm{X}_1, \dots, \mathrm{X}_n$ координатные величины $\mathrm{X}_{1j}, \dots, \mathrm{X}_{nj}$ симметрично зависимы, поэтому вероятность $\mathbf{P}\{X_{\boldsymbol{kj}} = X_{lj}\}$ не зависит от конкретного выбора пары $k, l \in \{1, \dots, n\}$. Обозначим через F_j функцию распределения величины $X_{\mathbf{1}j} - X_{2j}$. Тогда требование условия стохастической координатной единственности равносильно следующему требованию: функции F_j не имеют атомов в нуле при любом $j \in \{1, \dots, d\}$.

При выводе основных результатов наряду с условием стохастической координатной единственности мы будем использовать также условие независимости координат векторов $\mathrm{X}_1, \dots, \mathrm{X}_n$, то есть условие взаимной независимости столбцов матрицы $(X_{ij} | \, i = 1, \dots, n; j = 1, \dots, d)$.

$\mathbb{R}^d$-значные случайные векторы $\mathrm{X}_1, \dots, \mathrm{X}_n$, удовлетворяющие обоим этим условиям, мы будем называть *SUI*-векторами (англ. *Independence* — «независимость»), а соответствующие им координатные матрицы — *SUI*-матрицами.

Основные результаты

Символом $\#M$ мы в дальнейшем обозначаем число элементов множества *M*. Введем также обозначение для *n*-го гармонического числа:

$$H(n) = 1 + \frac{1}{2} + \dots + \frac{1}{n}, \qquad n = 1, 2, \dots .$$

Пусть $X = \{\mathrm{X}_1, \dots, \mathrm{X}_n\}$, где $\mathrm{X}_1, \dots, \mathrm{X}_n$ — $\mathbb{R}^d$-значные случайные *SUI*-векторы.

Теорема 1. Пусть $d = 2$. Тогда $\mathbf{E}[\#\varphi(X)\,] = H(n)$.

Доказательство. Пусть $X^1, \dots, X^2$ — элементы множества *X*, перенумерованные в соответствии с возрастанием первой координаты. Тогда, согласно леммам 3 и 4, последовательность вторых координат этих векторов $X_2^1 \dots, X_2^n$ является последовательностью симметрично зависимых случайных величин, для которой выполняется свойство стохастической единственности. По следствию из леммы 2 $\mathbf{E}[\#\varphi(X)\,] = \mathbf{E}R_n$, где R_n — число рекордных значений в последовательности $X_2^1 \dots, X_2^n$. Положим

$$\tau_1 = 1, \tau_k = \begin{cases} 1, \text{если } X_2^k < min\{X_2^1, \dots, X_2^{k-1}\} \\ 0, \text{если } X_2^k \geq min\{X_2^1, \dots, X_2^{k-1}\} \end{cases}, k \geq 2.$$

Тогда $\mathbf{E}R_n = 1 + \mathbf{P}\{\tau_2 = 1\} + \dots + \mathbf{P}\{\tau_n = 1\}$. Так как величины $X_2^1, \dots, X_2^k$ симметрично зависимы, то события $\left\{X_2^i = min\{X_2^1, \dots, X_2^k\}\right\}, i = 1, \dots, k$ равновероятны, а из условия стохастической единственности следует, что эти события являются попарно дизъюнктными с вероятностью 1. Поэтому $\mathbf{P}\{\tau_k = 1\} = \frac{\mathbf{1}}{k}$ при $k = 2, \dots, n$. ■

Следствие. Пусть $d = 2$, ξ — случайно выбранная точка из множества *X*, удовлетворяющего условиям теоремы 1. Тогда $\mathbf{P}\{\xi \in \varphi(X)\,\} = H(n)/n$.

Доказательство. Пусть $\kappa = \kappa(\xi)$ — номер точки ξ, полученный ею в результате сортировки множества *X*. Случайный выбор точки означает, что событие $\{\kappa = \mathrm{k}\}$ имеет вероятность $1 / n$ для каждого $k = 1, \dots, n$. Если это событие произошло, то событие $\xi \in \varphi(X)$ происходит в том и только том случае, когда $X_2^k = min\{X_2^1, \dots, X_2^k\}$. По формуле полной вероятности получаем

$$\mathbf{P}\{\xi \in \varphi(X)\,\} = \sum\nolimits_{k=1}^{n} \mathbf{P}\{\xi \in \varphi(X) | \kappa = \mathrm{k}\,\}\mathbf{P}\{\kappa = \mathrm{k}\} =$$

$$= \sum\nolimits_{k=1}^{n} \frac{1}{k} \times \frac{1}{n} = \frac{H(n)}{n}. \ ■$$

Заметим, что для любого непустого подмножества $X' \subset X$ мощности $n' = \#X'$ и случайно выбранной точки ξ из множества X'

$$\mathbf{P}\{\xi \in \varphi(X')\,\} = \frac{H(n')}{n'} = \frac{\mathbf{E}[\#\varphi(X')\,]}{n'}.$$

Следовательно,

$$\mathbf{P}\{\xi \in \varphi(X')\} = \frac{\mathbf{E}[\#\varphi(X')]}{n'}. \quad (1)$$

Равенство (1) справедливо, как нетрудно видеть, для любого $d \geq 2$.

Пусть $d > 2$, $X = \left(X_{ij} \middle| i = 1, \dots, n; j = 1, \dots, d\right)$ — *SUI*-матрица, строки которой симметрично зависимы и занумерованы по возрастанию первой координаты. Рассмотрим подматрицу X' полученной матрицы, образованную столбцами с номерами $2, \dots, d$. Эта подматрица также будет *SUI*-матрицей, строки которой симметрично зависимы согласно леммам 3 и 4. Пусть X_{k*} — k-я строка матрицы X, ${X'}_{k*}$ — подстрока в матрице X'. По лемме 1 событие $\{X_{k*} \in \varphi(X)\}$ происходит тогда и только тогда, когда происходит событие

$$\{{X'}_{k*} \in \varphi(\mathcal{M}_k)\}, \text{ где } \mathcal{M}_k = \{{X'}_{1*}, \dots, {X'}_{k*}\}.$$

Обозначим через τ_k индикатор этого события. Имеем

$$\mathbf{E}[\#\varphi(X)] = \sum_{k=1}^{n} \mathbf{P}\{\tau_k = 1\} = \sum_{k=1}^{n} \frac{\mathbf{E}[\#\varphi_2(\mathcal{M}_k)]}{k}. \quad (2)$$

Итерируя (2) (d - 2) раза и используя теорему 1, мы получим следующее равенство:

$$\mathbf{E}[\#\varphi(X)] = \sum_{k=1}^{N} \frac{1}{k} \left[\sum_{k_1=1}^{k} \frac{1}{k_1} \left[\dots \left[\frac{1}{k_{d-4}} \sum_{k_{d-3}=1}^{k_{d-4}} \frac{H(k_{d-3})}{k_{d-3}} \right] \right] \dots \right]. \quad (3)$$

Формула (3) слишком громоздка для использования на практике, поэтому оставшаяся часть статьи посвящена получению простых оценок для $\mathbf{E}[\#\varphi(X)]$.

Лемма 5.

1. $\mathbf{E}[\#\varphi(X)] \leq H^{d-1}(n)$.
2. $\mathbf{E}[\#\varphi(X)] \geq \frac{1}{2^{\frac{(d-2)(d-1)}{2}}} H^{d-1}(n)$.

Доказательство.

1. Из (2) имеем

$$\mathbf{E}[\#\varphi(X)] \leq \mathbf{E}[\#\varphi_2(X)] H(n) \leq \dots \leq H^{d-1}(n).$$

2. Заметим, что

$$S(n) := \sum_{k=1}^{n} \frac{H(k)}{k} > \frac{1}{2} H^2(n). \quad (4)$$

Действительно, так как

$$\sum_{1 \leq j \leq k \leq n} \frac{1}{kj} = \sum_{1 \leq k \leq j \leq n} \frac{1}{kj},$$

имеем

$$S(n) = \frac{1}{2} \left(\sum_{1 \leq j \leq k \leq n} \frac{1}{kj} + \sum_{1 \leq k \leq j \leq n} \frac{1}{kj} \right) = \frac{1}{2} \left(\sum_{1 \leq j,k \leq n} \frac{1}{kj} + \sum_{1 \leq k = j \leq n} \frac{1}{kj} \right) > \frac{1}{2} H^2(n).$$

Покажем теперь, что для любых $r, n \geq 1$

$$\sum_{k=1}^{n} \frac{H^r(k)}{k} \geq \frac{1}{2^r} H^{r+1}(n). \quad (5)$$

Так как функция $f(x) = x^r$ выпукла, то по неравенству Йенсена, принимая во внимание (4), имеем

$$\sum_{k=1}^{n} \frac{H^r(k)}{k} = \sum_{k=1}^{n} \frac{1}{kH(n)} \left(H(k) H^{1/r}(n) \right)^r \geq$$

$$\geq H^{(1/r - 1)r}(n) \times (S(n))^r \geq$$

$$\geq H^{1-r}(n) \times \frac{H^{2r}(n)}{2^r} = \frac{1}{2^r} H^{r+1}(n).$$

Применяя последовательно неравенство (5) в равенстве (3) (d - 2) раза, получим:

$$\mathbf{E}[\#\varphi(X)] \geq \frac{H^{d-1}(n)}{2^{1+\dots+(d-2)}} = \frac{1}{2^{\frac{(d-2)(d-1)}{2}}} H^{d-1}(n). \blacksquare$$

Принимая во внимание, что $H(n) = \Theta(\log n)$ мы получаем следующий результат.

Теорема 2. Пусть $d \geq 2$. Тогда

$$\mathbf{E}[\#\varphi(X)] = \Theta([\log n]^{d-1}).$$

Следствие. Пусть $d \geq 2$, ξ — случайно выбранная точка в множестве X. Тогда

$$\mathbf{P}\{\xi \in \varphi(X)\} = \Theta\left(\frac{[\log n]^{d-1}}{n} \right).$$

ЛИТЕРАТУРА

1. Campigotto, P. Active Learning of Pareto Fronts / P. Campigotto, A. Passerini, R. Battiti // IEEE Transactions on Neural Networks and Learning Systems. 2014. Vol. 25, Is. 3. Pp. 506–519. DOI: 10.1109/TNNLS.2013.2275918.

2. Couckuyt, I. Fast Calculation of Multiobjective Probability of Improvement and Expected Improvement Criteria for Pareto Optimization / I. Couckuyt, D. Deschrijver, T. Dhaene // Journal of Global Optimization. 2014. Vol. 60, Is. 3. Pp. 575–594. DOI: 10.1007/s10898-013-0118-2.

3. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II / K. Deb, A. Pratap, S. Agarwal, T. Meyarivan // IEEE Transactions on Evolutionary Computation. 2002. Vol. 6, Is. 2. Pp. 182–197. DOI: 10.1109/4235.996017.

4. Zuluaga, M. ε-PAL: An Active Learning Approach to the Multi-Objective Optimization Problem / M. Zuluaga, A. Krause, M. Püschel // Journal of Machine Learning Research. 2016. Vol. 17. Art. No. 15-047. 32 p.

5. Messac, A. The Normalized Normal Constraint Method for Generating the Pareto Frontier / A. Messac, A. Ismail-Yahaya, C. A. Mattson // Structural and Multidisciplinary Optimization. 2003. Vol. 25, Is. 2. Pp. 86–98. DOI: 10.1007/s00158-002-0276-1.

6. Andersson, J. A Survey of Multiobjective Optimization in Engineering Design. Technical report LiTH-IKP-R-1097. Department of Mechanical Engineering Linköping University, Linköping, Sweden. 2000. 34 p.

7. Kallenberg, O. Probabilistic Symmetries and Invariance Principles. — New York: Springer New York, 2005. — 524 p. — (Probability and Its Applications).
DOI: 10.1007/0-387-28861-9.

8. Goldie, C. M. Records in a Partially Ordered Set / C. M. Goldie, S. Resnick // The Annals of Probability. 1989. Vol. 17, No. 2. Pp. 678–699. DOI: 10.1214/aop/1176991421.

9. On the Average Number of Maxima in a Set of Vectors and Applications / J. L. Bentley, H. T. Kung, M. Schkolnick, C. D. Thompson // Journal of the Association for Computing Machinery. 1978. Vol. 25, No. 4. Pp. 536–543.
DOI: 10.1145/322092.322095.

10. Невзоров, В. Б. Рекорды. Математическая теория. — Москва: Фазис, 2000. — 244 с. — (Стохастика; Вып. 4).

11. Arnold, B. C. Records / B. C. Arnold, N. Balakrishnan, H. N. Nagaraja. — New York: John Wiley & Sons, 1998. — (Wiley Series in Probability and Statistics).
DOI: 10.1002/9781118150412.

DOI: 10.24412/2413-2527-2022-331-19-24

On the Expected Size of the Pareto Subset of a Random Set of Points

PhD A. N. Baushev
Emperor Alexander I St. Petersburg
State Transport University
Saint Petersburg, Russia
baushev@pgups.ru

PhD O. L. Semenova
Saint Petersburg State University
Saint Petersburg, Russia
o_semenova@mail.ru

Abstract. **The article is devoted to an estimation of the expected size of the Pareto set in a finite set of points drawn by symmetrically dependent random variables in $\mathbb{R}^d$. The received estimates are fully analogized to the well-known estimates in the case of independent random variables.**

Keywords: **a random set of points, the Pareto set, symmetrically dependent random variables, random permutations, records in a sequence of random variables.**

REFERENCES

1. Campigotto P., Passerini A., Battiti R. Active Learning of Pareto Fronts, *IEEE Transactions on Neural Networks and Learning Systems*, 2014, Vol. 25, Is. 3, Pp. 506–519. DOI: 10.1109/TNNLS.2013.2275918.

2. Couckuyt I., Deschrijver D., Dhaene T. Fast Calculation of Multiobjective Probability of Improvement and Expected Improvement Criteria for Pareto Optimization, *Journal of Global Optimization*, 2014, Vol. 60, Is. 3, Pp. 575–594. DOI: 10.1007/s10898-013-0118-2.

3. Deb K., Pratap A., Agarwal S., Meyarivan T. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, 2002, Vol. 6, Is. 2, Pp. 182–197. DOI: 10.1109/4235.996017.

4. Zuluaga M., Krause A., Püschel M. ε-PAL: An Active Learning Approach to the Multi-Objective Optimization Problem, *Journal of Machine Learning Research*, 2016, Vol. 17, Art. No. 15-047, 32 p.

5. Messac A., Ismail-Yahaya A., Mattson C. A. The Normalized Normal Constraint Method for Generating the Pareto Frontier, *Structural and Multidisciplinary Optimization*, 2003, Vol. 25, Is. 2, Pp. 86–98. DOI: 10.1007/s00158-002-0276-1.

6. Andersson J. A Survey of Multiobjective Optimization in Engineering Design. Technical report LiTH-IKP-R-1097. Department of Mechanical Engineering Linköping University, Linköping, Sweden, 2000, 34 p.

7. Kallenberg O. Probabilistic Symmetries and Invariance Principles. New York, Springer New York, 2005, 524 p. DOI: 10.1007/0-387-28861-9.

8. Goldie C. M., Resnick S. Records in a Partially Ordered Set, *The Annals of Probability*, 1989, Vol. 17, No. 2, Pp. 678–699. DOI: 10.1214/aop/1176991421.

9. Bentley J. L., Kung H. T., Schkolnick M., Thompson C. D. On the Average Number of Maxima in a Set of Vectors and Applications, *Journal of the Association for Computing Machinery*, 1978, Vol. 25, No. 4, Pp. 536–543. DOI: 10.1145/322092.322095.

10. Nevzorov V. B. Records. Mathematical theory [Rekordy. Matematicheskaya teoriya]. Moscow, Fazis Publishing House, 2000, 244 p.

11. Arnold B. C., Balakrishnan N., Nagaraja H. N. Records. New York, John Wiley & Sons, 1998. DOI: 10.1002/9781118150412.

DOI: 10.24412/2413-2527-2022-331-25-36

Russian version of the article © V. N. Kustov, A. I. Grokhotov, E. V. Golovkov is published
in *Intelligent Technologies in Transport*, 2021, No. 4 (28), Pp. 46–56.
DOI: 10.24412/2413-2527-2021-428-46-56.

A Simulation Software Model of the ⊕HUGO Stegosystem

Grand PhD V. N. Kustov, A. I. Grokhotov, E. V. Golovkov
Emperor Alexander I St. Petersburg State Transport University
Saint Petersburg, Russia
kvnvika@mail.ru, grohotov.aleksei@mail.ru, jyk22@mail.ru

Abstract. **In this article, the authors consider the problems of modern steganography. Starting with the presentation of a historical example of steganography, the authors classify contemporary steganography methods. The authors also offer a structural diagram of the steganographic system, which is based on further research. Further, the authors describe a simulation software model called a «⊕Highly Undetectable steGOsystem» or «⊕HUGO stegosystem» for short, implementing a steganographic method of transmitting a secret message embedded in a fixed digitized image. The article also discusses the principle of operation of the simulation software model and its steganographic justification. As an implementation algorithm, the authors used a cryptographic gamming algorithm using the function of bijective addition modulo two, conventionally denoted — ⊕. The authors determine the difficulty of detecting container change in this embedding method by calculating the Pearson correlation coefficient. The authors show that this model successfully improved information security when transmitting classified information in various electronic document management systems. The developed software model is much more efficient than the algorithm LSB, which is determined by higher performance and provides higher resistance to detection.**

Keywords: **simulation software model, highly undetectable stegosystem, stegosystem ⊕HUGO, cryptographic agorhythm of gamming, bijective addition modulo two, Pearson correlation coefficient.**

Introduction

Steganography is a science that studies methods to increase information security by hiding the very fact of the transfer of classified information. The main goal is to transmit an encrypted message in open, publicly available information, in secret, the very existence of which will be known only to the sending and receiving parties.

The first recorded steganographic methods consisted of manipulations with the information carrier. For example, clay tablets of ancient Sumerians were discovered by archaeologists. A clay tablet was the carrier of information, cuneiform was used at that time, and the steganographic method of hiding information consisted of cunning and ingenuity. On such tablets, the hidden text was stuffed with the first layer of the letter. After the sender applied a new layer of clay, a non-secret message was knocked out on it with a wedge by him. In this method, the container is a clay tablet, the hidden text is a secret message, and the key is knowledge, agreement on this method of transmission.

The second well-known steganography method is the story of the tyrant from Greece, Herodotus, who, while in captivity, used his slave to transmit a secret message through him. The method of concealing information was that the slave's head was shaved bald, after which a secret message was applied to the scalp by the sender. Over time, the hair on the slave's head grew, which protected the message from being read by third parties. This method was also used in the Roman Empire, as shown in Figure 1.

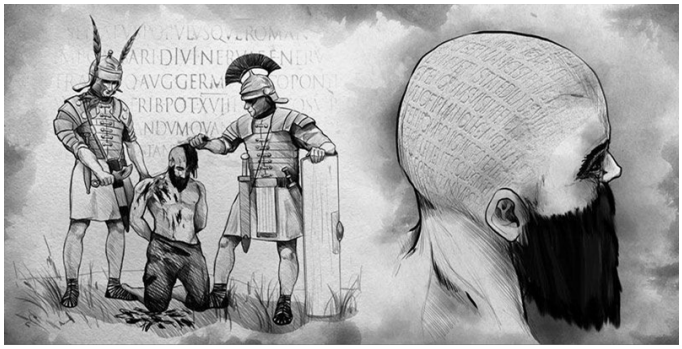

Fig. 1. The Story of Herodotus

England, Aeneas Tactician, described encrypting information that the sender used to save money by ordinary people. The sender used this method because sending letters over long distances was expensive, while sending a newspaper costs many times less. In this regard, a method was invented, which consisted in piercing small holes above the letters in old newspapers. After that, the sender sent the newspaper, and the recipient, writing out these letters, received an encrypted message.

Digital steganography as a separate science appeared not so long ago, so it has no established terminology. The authors can cite one of the most common definitions that can be formulated as follows: Digital steganography is the science of secretly and reliably hiding some bit sequences in other sequences of a similar nature. In this formulation, there are primary criteria for the applied steganographic methods. It uses the concept of invisibility. It can be defined as stability to the analysis of information by a person or a program that detects changes in the structure of information and reliability - which means preserving the integrity of information when exposed to various kinds of noise.

A steganographic system consists of the means and methods necessary to form a hidden data transmission channel. In the

process of its creation, it is required to take into account the introductory provisions of digital steganography:

- the optimal ratio of the complexity of the implementation of the stegosystem to the security of the system;
- performance of optimal throughput;
- maintaining the integrity and completeness of classified information during transmission;
- the stegosystem is entirely open to the intruder, excluding the private key;
- if the violator discloses information about data transmission by the steganographic method, it should be impossible to extract secret information without knowing the key.

In digital steganography, the main tasks are developing new, more advanced, highly undetectable methods and stegosystems, improving and modifying existing ones, and creating based on more efficient steganographic systems for storing and transmitting the information.

Block diagram of the steganographic information protection system

In general, a stegosystem can be compared to a communication system. Figure 2 shows a generalized block diagram of a steganographic system.

The basic concepts in stegosystems are:

1. A *hidden message* is an information that is encrypted in the stegosystem.

2. The container (or *covering object*) is open information in which the sender will embed the hidden message. The presence of a secret message in the container should not cause noticeable changes in the container.

3. A *key* — as in cryptography, is secret information that is used when encrypting/decrypting a message. The key can be public, and then it will be openly distributed by a trusted third party over the network to embed the message in the container or private. The recipient will use it to receive the message from the container.

4. *Steganographic algorithm* — this concept refers to two types of transformation: the first is a direct algorithm, which from a message, container, key will have a container with a message encrypted in it, the second is a reverse algorithm, which forms a pair: a container with a message, key, will have the original message at the output.

5. *Precoder* — performs the translation of secret information into the form necessary for encryption into the container.

6. *Stegocoder* — responsible for embedding a secret message in a container.

7. A *stegochannel* is a communication channel through which a container with an encrypted message is transmitted inside. The container can be damaged by directed attacks by intruders or be distorted under the influence of interference.

8. A *stegodetector* is software that analyzes the structure of a container for changes. Such changes may be intentional when an embedded encrypted message is detected in the container and errors and distortions during transmission.

9. A *stegoencoder* — restores a message from a container using a key.

Principles of steganographic transformations are:

- the container must contain a structure that can be changed with the condition that the functionality of the object will not be affected;
- when analyzing the structure of the container, the changes should not be recognized by attackers.

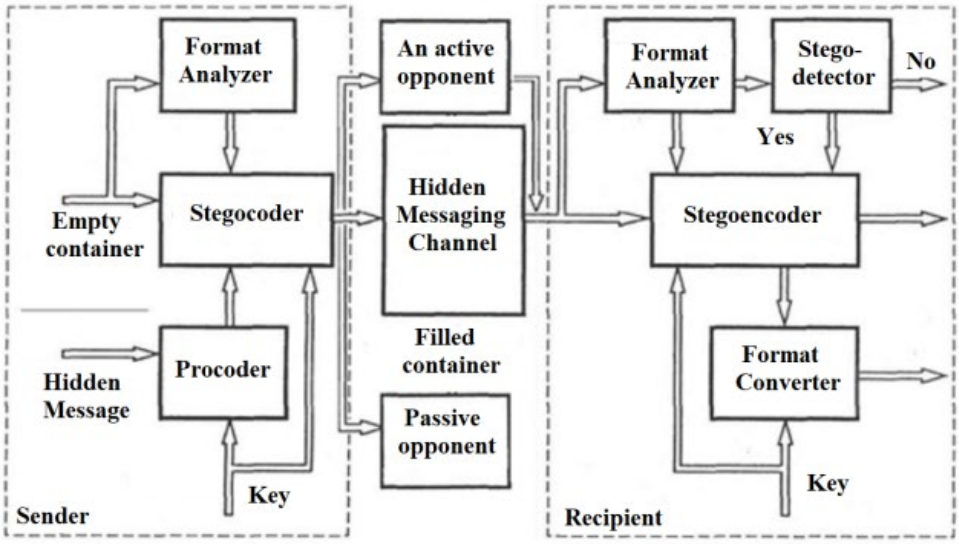

Fig. 2. Block diagram of a steganographic system

Classification of steganography methods

The key principles based on which different methods of steganography are formed:

- incomplete accuracy — some files do not need full transmission accuracy, and adjustments can be made to them by senders;
- invisible to humans — some file structures contain redundancy; when changing the structure of such a file, insignificant changes occur; human senses cannot distinguish that, and there is no special equipment to detect such changes.

A general idea of methods is allocating insignificant parts in the container structure and replacing such parts with information from the message.

The general block diagram of the classification of steganographic methods is shown in Figure 3.

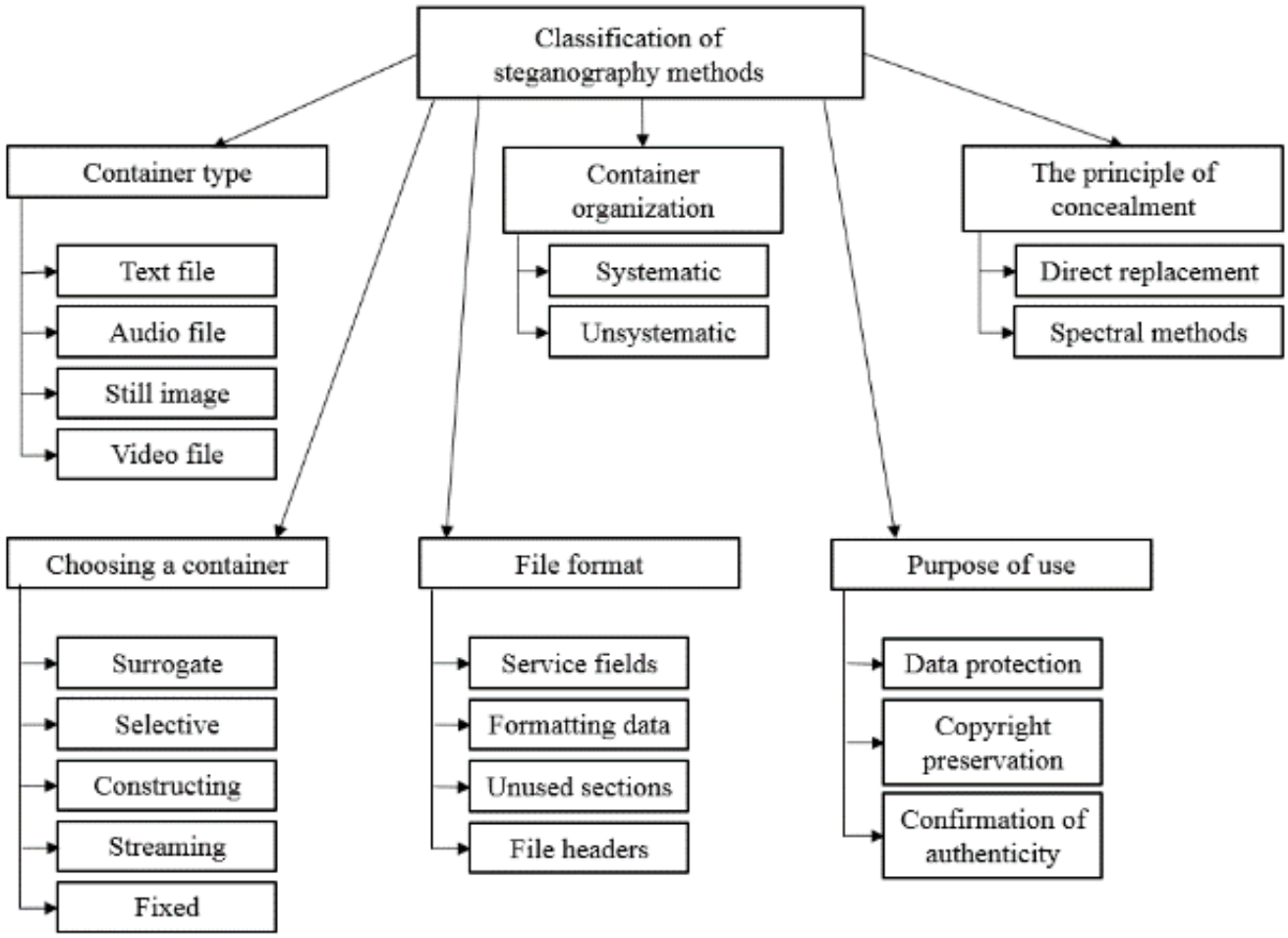

Fig. 3. Block diagram of the classification of steganographic methods

Classification of methods according to the practice of container selection:

- surrogate — an empty container is not taken, preference is given to the first one that comes along, the container, in this case, is most often not optimal;
- selective — in this method, a large number of empty containers are created, after which the optimal one remains, which most repeats the static noise characteristics of an empty container;
- constructed — the stegosystem itself forms empty containers; in this case, the noise of the container masks the hidden message;
- streaming (continuous) containers — such methods cannot know the characteristics of the container in advance, and the embedding of the secret message will be in real-time;
- fixed (limited length) containers — methods with predefined features of an empty container.

Classification of methods by container organization method:

- systematic — in such practices, it is possible to determine where the sender will embed the secret information and where the noise data will be;
- unsystematic — in such containers, it is necessary to process the file to receive a secret message fully.

Classification of methods based on the use of unique properties of file presentation formats:

- service fields, such as headers, which are not taken into account in programs, and mostly filled with zeros;
- special formatting of data;
- use of unused sections on media;
- removal of file headers-identifiers, etc.

Classification of methods according to the principle of hiding methods used is divided into:

- methods of direct replacement — represent the replacement of unimportant bits of an empty container with bits of a hidden message, based on the excess of the information environment in the spatial or temporal domain;
- spectral methods — use spectral representations of elements of the embedding environment to hide the message.

Classification according to the purpose of using steganographic methods:

- protection of non-public data;
- copyright preservation;
- confirmation of authenticity.

The methods are divided by container types:

- text files;
- audio files;
- images;
- videos.

Research underground of HUGO technology

Let's look in chronological order at the well-known scientific publications that form the basis of the authors' research.

Research in the field of highly undetectable stegosystems intensified at the beginning of the twentieth century when an article was published [1]. A modification of the F5 algorithm was proposed, providing high resistance to visual attacks with a low degree of detection. Thus, the attacker's task to detect a hidden message embedded in the covering object has become more complicated.

In the following paper [2], the authors use so-called wet paper codes and introduce the concept of perturbed quantization to describe a new approach to passive safety of steganography. The authors present a heuristic algorithm that provides higher steganographic security for covering objects in JPEG format.

In the article [3], the authors determine the largest embedded payload that the attacker cannot detect. The authors claim that the average undetectable ability to embed hidden messages for black-and-white covering objects in JPEG format is at least 0.05 bits/per non-zero DCT coefficient.

In further studies [4], the authors established a connection between synthesizing a stegosystem, minimizing distortion during implementation, and statistical physics. A distinctive feature of this work from previous works is that the authors introduced an arbitrary character of the distortion function. It allowed the authors to describe the changes in the implementation as spatially dependent. The research method proposed by the authors reduced the task of synthesizing a stegosystem to the study of finding the minimum values of the distortion function potentials that determine the statistical undetectability of a hidden message.

The authors described another new approach to using additive steganographic embedding in the spatial domain of the covering object [5]. The authors propose to determine the level of change in pixel values in the high-frequency regions of the covering object by its weight and aggregation using the inverse Helder norm to determine individual pixel changes. It makes it possible to increase the stability of the proposed scheme for steganalysis significantly.

In the paper [6], the authors used a different strategy in which the covering object is represented as a sequence of independently distributed quantized Gaussians. The probabilities of making changes to the pixels of the covering object are calculated to minimize the overall discrepancy for a given embedding operation and a given payload.

In the article [7], the authors propose a universal approach to the description of distortions, called universal wavelet relative distortion (UNIWARD), and apply it to embed a hidden message in the spatial and frequency domains of the encompassing object.

In most stegosystems for still digital images using raster formats, the changes' amplitude is usually limited to the minimum value when implementing a hidden message. However, in the article [8], the authors explore ways to increase the embedding volume in highly textured areas of the covering object by significantly growing the embedding amplitude, which leads to an increase in payload.

The opinion that adding additional information to a hidden message increases the security of the stegosystem has long been indisputable. Further confirmation of this is the article [9]. The authors investigate the use of additional information in a set of several JPEG images for the same scene, provided there is no access to the pre-recording.

They further developed the results obtained by the authors in the previous publication in the article [10]. As in the latter case, the secret message is hidden in the covering object by adding a noise signal to it, a heteroscedastic noise naturally introduced by the recipient. The main requirement of this method is that the covering image is available in raw form (this operation is called «sensor capture»). A significant payload can be embedded for monochrome n objects or low-quality JPEG while providing a high level of security.

Simulation software model for embedding a hidden message in a covering object by gamming

In this section, the authors present a prototype of a simulation software model for implementing the process of transmitting hidden messages in digital still images on the way from the sender to the recipient using their discrete transformations and concealment algorithms.

In the model under consideration, the principle of operation is based on the well-known least significant bits (LSB) method. It is its complement, with the correction of its inherent shortcomings.

The authors propose to carry out a preliminary conversion of the file into a form that resembles noise in many ways. It makes it possible to increase the system's security since in the event of a leak of a secret message, it will be possible to restore it only with the private key with which the image is initially converted.

Purpose and structure of the simulation software model

The purpose of the simulation model development is to increase the level of security of information transmitted in the hidden electronic document management system using steganographic methods. The structure of this model is illustrated in Figure 4.

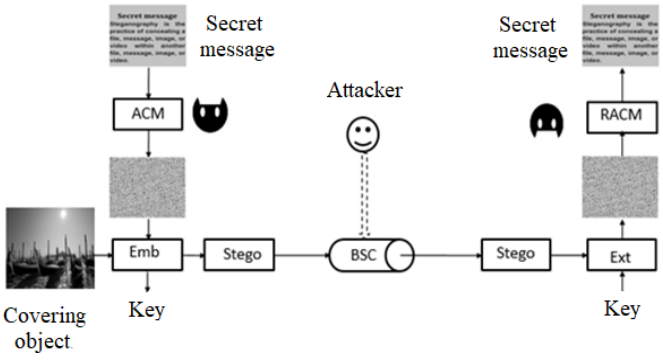

Fig. 4. The structure of the simulation software model

The model was developed by the previously mentioned scheme of the steganographic system, which can be seen in Figure 2.

The functions of the procoder in this model are recommended to be performed using the ACM (Arnold Cat Map) method. This method converts a secret graphic digitized image in PNG format into a form that most resembles noise. An example of how these method works are shown in Figure 5.

Further, according to the scheme, the secret message converted into noise gets to the input to the stegocoder module,

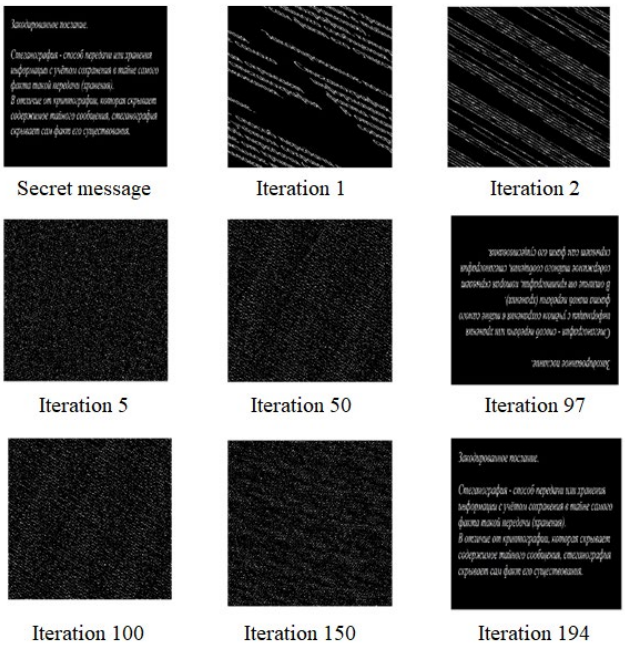

Fig. 5. The process of converting a hidden message using the ACM method

which performs the function of embedding the message into the container, which is also a graphic image. This module is represented as Emb, which means embed.

This module has two outputs. The first thing you can see in the diagram is that the output element is the key needed to restore the image during the decoding operation by the recipient. Without this key, it is impossible to restore the original image. An example of the image (covering object) in which the sender embedded the message is shown in Figure 6. An example of the secret key is shown in Figure 7, and the graphic image that the sender encrypted at the same time is shown in Figure 8.

Fig. 6. The covering object

Fig. 7. The Key (with labels where the sender embedded the secret image)

How the level of secrecy of documents is determined: basic rules It is possible to correctly solve the problem of installing the secrecy stamp and preserving state secrets only with a systematic approach to it. Therefore, you should rely on a number of rules. It is necessary not to allow self-will in this area, otherwise there will be problems in working with documentation or with the leakage of important information. Let's list the rules in question.

1. The rule of a systematic approach to determining the level of secrecy of documents. The main essence of this principle is to take into account the general problem of secrecy. It is necessary to take into account the existing duality: on the one hand, there is a goal to ensure the reliable preservation of state secrets, on the other - it is impossible to unreasonably and massively classify data. Therefore, it is unacceptable both to overestimate the secrecy rating and to underestimate it. Any extremes should be avoided.
2. The rule of objectivity when assigning the secrecy stamp. There is a list of information to be classified, which you need to rely on in your work. A subjective approach is unacceptable
3. The rule of optimizing the volume of secret data in papers. In separate documentation, secret information should be kept to a minimum and strictly in the volume necessary to solve the issue under consideration.

Fig. 8. The secret message

The second output is a container with a secret message encrypted inside. In the diagram, this container is marked as a stego module. This container will be transmitted via a binary symmetric channel (BSC).

BSC is a simple binary channel through which it is possible to transmit only 0 and 1, with the condition that on the other side, the receiver does not always receive the value that the sender sent. This channel illustrates the simplest example of a communication channel with a condition for noise during data transmission.

At the exit from the communication channel, the Stego container enters the Ext (Extract) module, which means to extract. This module extracts a secret message from the container so far in noise, using the secret key.

Next, the secret message from the container gets into the RACM module (Reverse Arnold Cat Map), which is the reverse of the ACM module, and restores the original hidden image from the noise view.

THE STRUCTURE OF THE SOFTWARE PACKAGE AND THE FUNCTIONS IMPLEMENTED IN IT

The software package (SP) is developed in the free NetBeans integrated application development environment in the Java programming language.

When creating the SP, the various steganographic stages of the application were divided into classes. The structure of the developed application is shown in Figure 9.

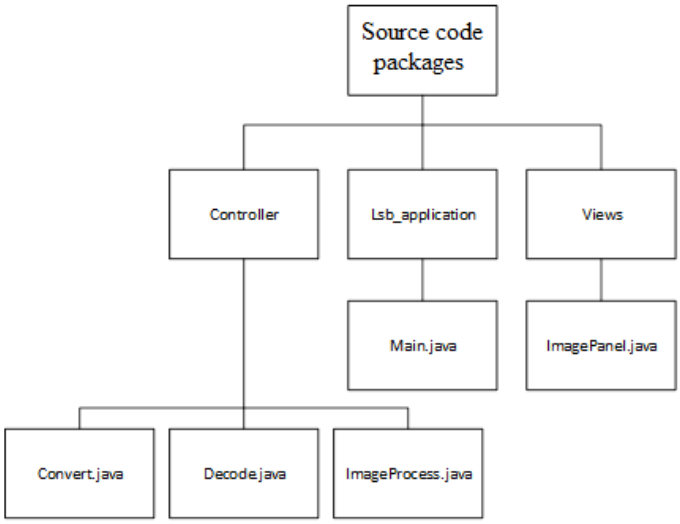

Fig. 9. The structure of software package

The SP consists of five classes, and let's briefly list their functionality.

The Main class is the main class in which the graphical user interface (GUI) SP is created, and the events of the buttons pressed by the user are processed. The GUI view is shown in Figure 10.

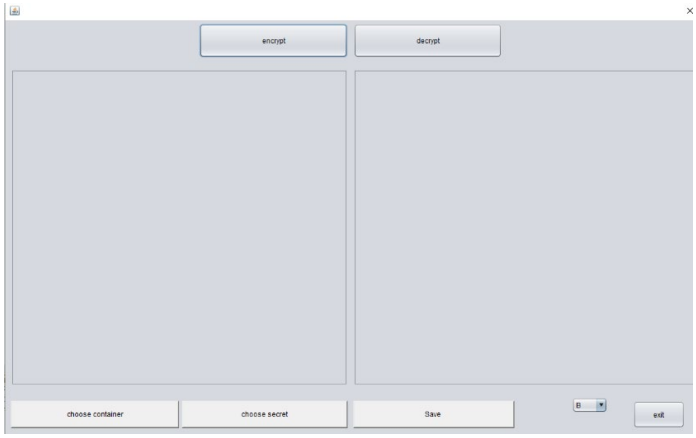

Fig. 10. The GUI view of the SP

The following functions are implemented in this class:

• decryptActionPerformed — this function handles the decrypt button click event. Two images are read and transmitted to the decryption module: a container with a secret message embedded in it and a key image with the coordinates of the embedded message bits. After that, the image with the key replaced an image that is obtained at the output of the stegodecoder module;

• encryptActionPerformed — this function processes the event of pressing the encrypt button. During processing, it reads the image of the container and the secret message, after which it transmits these images to the stegocoder module. Then the function displays the image of the container with the embedded secret message and the image of the key with the coordinates of the embedded bits of information;

• exitActionPerformed — this function handles the exit event of the program by pressing the exit button.

• load_containerActionPerformed — implements the process of selecting a graphic image by the user for the role of a container and displays this image on the screen;

• load_secretActionPerformed — implements the process of choosing a picture of secret information by the user, which will be encrypted into a container and displayed on the program screen.

• saveActionPerformed — this function saves two images that are currently displayed on the program panels.

The image panel class is responsible for initializing panels for images. It implements functions:

• setImage — add an image to the panel;

• getImage — read the image from the panel;

• removeImage — clears the panel from the image;

• extractBytes2 — this function converts the image from the panel into an array of bytes. And for further, it is used when embedded in the container.

The Convert class was developed by it to implement two auxiliary functions:

• intToBytes — in this function, the numeric format Integer is converted into an array of bytes;

• buildStego — this function implements the formation of an array of bytes for embedding in a container by receiving a byte array of a message as input and adding a service header to it – the length of this array.

The Decode class performs the function of the stegodecoder module, the following functions are initialized in this class:

• extractHiddenBytes_B;

• extractHiddenBytes_G;

• extractHiddenBytes_B.

All three of these functions perform decoding, and their difference is that of the bytes of what color you need to get a secret message.

The last ImageProcess class performs the function of a stegocoder module; it has the same structure as Decode – a separate procedure is implemented for each of the three cases, depending on the byte of what color the embedding of secret information will take place. And also, in this class, the suitability of the pixel for use for encryption is checked. List of functions in this class:

• B_Hide;

• G_Hide;

• R_Hide.

THE PROCESS OF EMBEDDING A HIDDEN MESSAGE IN A CONTAINER

As mentioned above, SP is based on the steganographic method of least significant bits.

For embedding, the last bits of the bytes responsible for the colors in the image are used. The program provides three encryption options: in blue, red, and green bytes. The selection is made by switching the drop-down list in the program interface, as shown in Figure 11.

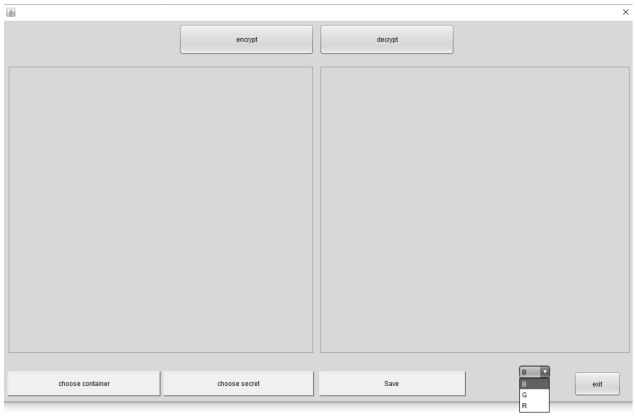

Fig. 11. Choosing the byte color for embedding

Next, the secret image is converted into an array of bytes. After which, a header containing the length of the embedded information is added to the beginning of this array.

When embedding an image, the container is analyzed to select suitable pixels at the borders of the image color change. It is done to avoid areas of uniform color because changes in them are simply detected by steganalysis.

The image analysis process takes place by selecting a 3x3 pixel area of the image, after which the central pixel is analyzed for possible embedding. Let's consider the analysis of the area using the example of the area shown in Figure 12.

A_1	A_2	A_3
A_4	B_i	A_5
A_6	A_7	A_8

Fig. 12. The pixel area under study

Let's introduce the notation: A_i is the color value of the i pixel, B_i is the pixel under study. Then the check will take place according to the formula:

$$B_i = \begin{cases} 1, \left| \frac{B_i + \sum_{i=1}^{8} A_i}{9} - B_i \right| > 8, \\ 0, \left| \frac{B_i + \sum_{i=1}^{8} A_i}{9} - B_i \right| < 8\,. \end{cases} \quad (1)$$

Authors should also note that more specific neighborhood types can be used, such as three consecutive bytes or a 3x3 cross.

The embedding operation takes place according to the ⊕HUGO algorithm. For a visual description of it, let's assume that the secret message M, a subset of bytes of the covering object C selected for the embedding operation and satisfying the condition in formula (1), and a subset of the corresponding m bytes of stego S are the final byte strings. We can describe the execution of the ⊕HUGO algorithm by the following sequence of actions.

For the embedding operation (Emb), we perform:

1. The next half byte m_i of the hidden message is added using the exclusive operation XOR with the right half of the next byte c_i satisfying the condition in formula (1), from the subset C. The result of the operation is written to the right half of the corresponding next byte of the covering object (stego) s_i. Formally, this operation can be written as the ratio: $s_i = m_i \oplus c_i$.

2. Item 1 is executed for all half-bytes of the secret message.

In the example shown in Figure 13, for the embedding operation of the first half-byte m_1 of a hidden message, the embedding process can be represented as the following relations: $s_1 = m_1 \oplus c_1$. The representation of this operation in hexadecimal code looks like this: 1 = B⊕A. In binary, it looks like this: 0001 = 1011⊕1010.

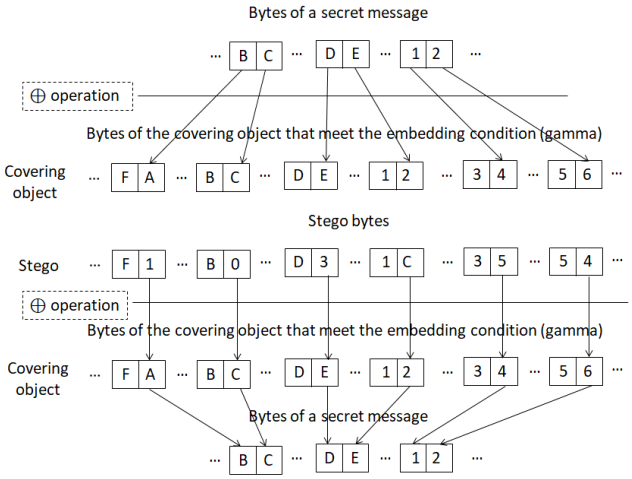

Fig. 13. The scheme for performing operations by the ⊕HUGO algorithm

In this paper, in the process of embedding a secret message, the above method is used, with one difference that only the last bit of a byte, a pixel satisfying condition (1), is used for encryption.

In parallel with the process of embedding information into a pixel, a black or red pixel is placed on a pure white image. At the same time, a black or red pixel is placed in the container's images by the pixel's coordinate into which the information is embedded. This image will serve as a key for the stegodecoder.

At the output of the stegocoder module, there will be a container with a built-in message and a secret key.

THE PROCESS OF EXTRACTING A MESSAGE FROM A CONTAINER

Consider the operation of extracting (Ext) a hidden message.

As a recipient, we have two images, and it is necessary to perform a decoding operation.

The first step is to add these two images to the program, see Figure 14 and Figure 15.

The decoding process takes place according to the following algorithm:

1. Determining the length of the encrypted message. It is done by decrypting the header. For this, both images and the length of the header are transmitted to the decryption function.

2. The decryption process takes place by determining the pixel marked on the key, after which the color value according

to which the encryption took place is read in the container using these coordinates. Then, the value of the last bit is added using the operation ⊕ with one if a black pixel was marked on the key and with zero if red.

3. The next step is the decryption of a classified message. For this, the same images and header length obtained in the second step are transmitted to the decryption function.

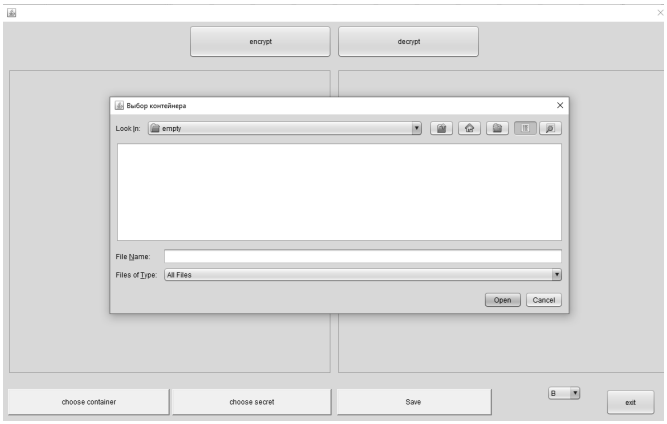

Fig. 14. Procedure for selecting a file to upload to the program

Fig. 15. The result of loading the container and the secret key into the program

The decoding process takes place according to the following algorithm:

1. Determining the length of the encrypted message. It is done by decrypting the header. For this, both images and the length of the header are transmitted to the decryption function.

2. The decryption process takes place by determining the pixel marked on the key, after which the color value according to which the encryption took place is read in the container using these coordinates. Then, the value of the last bit is added using the operation ⊕ with one if a black pixel was marked on the key and with zero if red.

3. The next step is the decryption of a classified message. For this, the same images and header length obtained in the second step are transmitted to the decryption function.

The authors should note that the program removes the pixels already passed from the secret key, and it is done to avoid repeated operations during decryption. Authors should also note that the same functions ⊕ are used for the embedding and extraction processes in implementing the HUGO algorithm. It is due to a remarkable property of this operation called bijectivity (reversibility).

ANALYSIS OF SIMULATION RESULTS

Two attempts to determine the information-theoretical stability of stegosystems are known from the literature. Kashin's definition [11] is based on the following requirement: the entropy of an empty covering object (a container with noise) relative to it should be small. We emphasize that we are talking about relative entropy. Thus, Kashin considers the opponent's task to distinguish an empty covering object from a stego as a task of statistical testing hypotheses. Another approach is described in the work of J. Zöllner, et al. [12]. It is based on the following requirement: knowledge of the covering object and its corresponding stego does not reduce the entropy of the hidden message. Note that here the opponent's task essentially boils down to extracting some information about a secret message (obviously, detecting a steganographic channel is a particular case of this task).

The work of Anderson and Petitcolas [13] and its early version [14] are also known. Some mathematical statements are formulated. For example, an estimate of the capacity of steganographic channels from above through the entropy difference. However, these works are of an overview nature and do not provide mathematically rigorous definitions of the concepts under consideration.

The authors should note that the above stability estimates of the stegosystem are based on the entropy approach. It requires precise determination of the laws of distribution of random variables C' and S. This requirement is the main obstacle that makes it difficult, and in some cases impossible, to apply this approach.

In this case, simpler ratios can be used to statistically evaluate the effectiveness of the developed stegosystem [15] implementing the HUGO algorithm.

The authors should note that most of the earlier estimates of the stability of the stegosystem are based on the entropy approach and require precise determination of the laws of distribution of random variables C (covering object) and S (stego). This requirement is the main obstacle that makes it difficult, and in some cases impossible, to apply this approach.

In this case, simpler ratios can be used to statistically evaluate the effectiveness of the developed stegosystem implementing the ⊕HUGO algorithm.

The main widely used metric for displaying the difference between empty and filled covering objects is the peak signal-to-noise ratio (PSNR), calculated by the formula:

$$PSNR = 10\log_{10}\frac{255^2}{MSE}.$$

It is the ratio between the maximum possible signal value and the power of noise that distorts the signal value [16].

The root-mean-square error (MSE) determines the difference between the pixel intensities of this and the covering object. MSE (denoted by the symbol Б) is calculated from the following ratio:

$$6 = \frac{1}{M \times N}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(f(i,j) - f'(i,j)\right)^2,$$

where $f(i, j)$ is the brightness of the pixel of the covering object, and $f'(i, j)$ is the brightness of the corresponding stego pixel, N is the length of the digital image in pixels, M is the width of the digital image in pixels.

A high value of Ϭ indicates poor quality of the original image and vice versa.

Capacity is a percentage of the size of the initial covering object V_c and the secret message V_m, calculated by the formula:

$$Capacity = \frac{V_m}{V_c}.$$

The r_{cs} correlation displays the degree of identity of the paired linear relationship between C_i and S_i covering object.

Usually, r_{cs} is calculated from the ratio [17]:

$$r_{CS} = \frac{cov_{cs}}{(n-1)6_c6_s},$$

where cov_{cs} is called covariance and is calculated from the ratio:

$$cov_{cs} = \sum_{j=1}^{K}(c_j - \dot{c})(s_j - \dot{s}).$$

If we expand the product of 6_c6_s, we get a formula for calculating them:

$$6_c6_s = \sqrt{\sum_{j=1}^{K}(c_j - \dot{c})^2 \sum_{j=1}^{K}(s_j - \dot{s})^2}. \qquad (2)$$

Assuming $K = N \times M$, $n = 2$, we obtain the following relation for the correlation coefficient:

$$r_{CS} = \frac{\sum_{j=1}^{K}(c_j - \dot{c})(s_j - \dot{s})}{6_c6_s},$$

and given the ratio (2) for 6_c6_s, we obtain the final expression for calculating the correlation coefficient r_{cs} in the following form:

$$r_{CS} = \frac{\sum_{j=1}^{K}(c_j - \dot{c})(s_j - \dot{s})}{\sqrt{\sum_{j=1}^{K}(c_j - \dot{c})^2 \sum_{j=1}^{K}(s_j - \dot{s})^2}},$$

where c_j is the value of byte j of the covering object C_i; s_j is the value of byte j of S_i; $\dot{c}$ and $\dot{s}$ are the average values of bytes C_i and S_i, respectively; 6_c — MSE for C_i; 6_s — MSE for S_i; n is the number of observations compared (in this case $n = 2$); K is the number of bytes in C_i and S_i.

The simulation results shown in Figure 16 and in Table 1 allow us to assert the practical indistinguishability of the covering object even with the capacity values of 0.593308641975 since the value of the Pearson correlation coefficient does not exceed the value 0.99993720336. It makes solving the problem of steganalysis complicated even for a very experienced steganalytic.

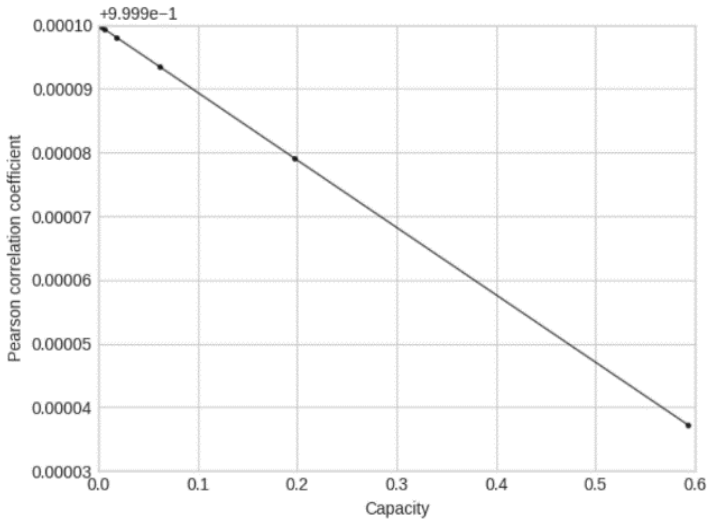

Fig. 16. Dependence of the Pearson correlation coefficient on the capacity

Table 1

Results of calculating the dependence of the Pearson correlation coefficient on the capacity

No.	Capacity	Pearson correlation coefficient
1	0.002093827160	0.99999977887
2	0.005797530864	0.99999939581
3	0.018182716049	0.99999807112
4	0.061323456790	0.99999350505
5	0.197604938271	0.99997905860
6	0.593308641975	0.99993720336

Conclusion

In this article, the authors presented a simulation software model called «Highly Undetectable SteGOsystem» or ⊕HUGO for short, implementing a steganographic method of transmitting a secret embedded in a still digitized image. The authors developed the principle of operation of the program and its steganographic justification based on a cryptographic gamming algorithm. This algorithm uses functions of bijective addition modulo two, conventionally denoted — ⊕.

The authors demonstrated the difficulty of detecting the fact of container change in this embedding method by calculating the Pearson correlation coefficient. Users can implement this model to improve information security when transferring classified information in various electronic document management systems. The developed simulation software model is much more efficient than the least significant bit algorithm (LSB), which is determined by higher performance and by providing higher resistance to detection.

References

1. Westfield A. F5 — A Steganographic Algorithm: High Capacity Despite Better Steganalysis. In: *Moskowitz I. S. (ed.) Proceedings of the Fourth International Workshop of Information Hiding (IH 2001), Pittsburgh, PA, USA, April 25–27, 2001. Lecture Notes in Computer Science*, 2001, Vol. 2137, Pp. 289–302. DOI: 10.1007/3-540-45496-9_21.

2. Fridrich J., Goljan M., Soukal D. Perturbed Quantization Steganography, *Multimedia Systems*, 2005, Vol. 11, Is. 2, Pp. 98–107. DOI: 10.1007/s00530-005-0194-3.

3. Fridrich J., Pevný T., Kodovský J. Statistically Undetectable JPEG Steganography: Dead Ends Challenges, and Opportunities, *Proceedings of the Ninth Workshop on Multimedia and Security (MM&Sec '07), Dallas, TX, USA, September 20–21, 2007*. New York, Association for Computing Machinery, 2007, Pp. 3–14. DOI: 10.1145/1288869.1288872.

4. Filler T., Fridrich J. Gibbs Construction in Steganography, *IEEE Transactions on Information Forensics and Security*, 2010, Vol. 5, Is. 4, Pp. 705–720.
DOI: 10.1109/TIFS.2010.2077629.

5. Holub V., Fridrich J. Designing Steganographic Distortion Using Directional Filters, *Proceedings of the Fourth IEEE International Workshop on Information Forensics and Security (WIFS 2012), Costa Adeje, Spain, December 02–05, 2012*. Institute of Electrical and Electronics Engineers, 2012, Pp. 234–239. DOI: 10.1109/WIFS.2012.6412655.

6. Fridrich J., Kodovský J. Multivariate Gaussian Model for Designing Additive Distortion for Steganography, *Proceedings of the 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), Vancouver, Canada, May 26–31, 2013*. Institute of Electrical and Electronics Engineers, 2013, Pp. 2949–2953. DOI: 10.1109/ICASSP.2013.6638198.

7. Holub V., Fridrich J., Denemark T. Universal Distortion Function for Steganography in an Arbitrary Domain, *EURASIP Journal on Information Security*, 2014, Art. No. 1, 13 p.
DOI: 10.1186/1687-417X-2014-1.

8. Sedighi V., Fridrich J., Cogranne R. Content-Adaptive Pentary Steganography Using the Multivariate Generalized Gaussian Cover Model. In: *Alattar A. M., et al. (eds) Proceedings of the SPIE/IS&T Electronic Imaging 2015 Media Watermarking, Security, and Forensics, San Francisco, CA, USA, February 09–11, 2015. Proceedings of SPIE*, 2015, Vol. 9409, Art. No. 94090H, 13 p. DOI: 10.1117/12.2080272.

9. Denemark T., Fridrich J. Steganography with Multiple JPEG Images of the Same Scene, *IEEE Transactions on Information Forensics and Security*, 2017, Vol. 12, Is. 10, Pp. 2308–2319. DOI: 10.1109/TIFS.2017.2705625.

10. Denemark T., Bas P., Fridrich J. Natural Steganography in JPEG Compressed Images, *Electronic Imaging*, 2018, Is. 7, Art No. 316, 10 p.
DOI: 10.2352/ISSN.2470-1173.2018.07.MWSF-316.

11. Cachin C. An Information-Theoretic Model for Steganography. In: *Aucsmith D. (ed.) Proceedings of the Second International Workshop of Information Hiding (IH 1998), Portland, OR, USA, April 14–17, 1998. Lecture Notes in Computer Science*, 1998, Vol. 1525, Pp. 306–318.
DOI: 10.1007/3-540-49380-8_21.

12. Zöllner J., Federrath H., Klimant H., et al. Modeling the Security of Steganographic Systems. In: *Aucsmith D. (ed.) Proceedings of the Second International Workshop of Information Hiding (IH 1998), Portland, OR, USA, April 14–17, 1998. Lecture Notes in Computer Science*, 1998, Vol. 1525, Pp. 344–354.
DOI: 10.1007/3-540-49380-8_24.

13. Anderson R., Petitcolas F. A. P. On the Limits of Steganography, *IEEE Journal of Selected Areas in Communications*, 1998, Vol. 16, Is. 4, Pp. 474–482.
DOI: 10.1109/49.668971.

14. Anderson R. Stretching the Limits of Steganography. In: *Anderson R. (ed.) Proceedings of the First International Workshop of Information Hiding (IH 1996), Cambridge, United Kingdom, May 30–June 01, 1996. Lecture Notes in Computer Science*, 1996, Vol. 1174, Pp. 39–48.
DOI: 10.1007/3-540-61996-8_30.

15. Kustov V. N., Protsko D. K. Ispolzovanie diskretnogo veyvlet-preobrazovaniya dlya vnedreniya informatsii v izobrazheniya [Using Discrete Wavelet Transform to Embed Information in Images], *Nauchnye tendentsii: Voprosy tochnykh i tekhnicheskikh nauk: Sbornik nauchnykh trudov po materialam XVII Mezhdunarodnoy nauchnoy konferentsii [Scientific Trends: Issues of Exact and Technical Sciences: Collection of scientific papers based on the materials of the XVII International Scientific Conference], Saint Petersburg, Russia, June 12, 2018*. Saint Petersburg, International United Academy of Sciences, 2018, Pp. 15–20. DOI: 10.18411/spc-12-06-2018-05. (In Russian)

16. Peak signal-to-noise ratio, *Wikipedia*. Available at: http://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio (accessed 07 Dec 2021).

17. Correlation, *Wikipedia*. Available at: http://en.wikipedia.org/wiki/Correlation (accessed 07 Dec 2021).

DOI: 10.24412/2413-2527-2022-331-25-36
Русскоязычная версия статьи © В. Н. Кустов, А. И. Грохотов, Е. В. Головков опубликована в журнале «Интеллектуальные технологии на транспорте». 2021. № 4 (28). С. 46–56.
DOI: 10.24412/2413-2527-2021-428-46-56.

Имитационная программная модель ⊕HUGO стегосистемы

д.т.н. В. Н. Кустов, А. И. Грохотов, Е. В. Головков
Петербургский государственный университет путей сообщения Императора Александра I
Санкт-Петербург, Россия
kvnvika@mail.ru, grohotov.aleksei@mail.ru, jyk22@mail.ru

***Аннотация*. В статье рассмотрены проблемы современной стеганографии. Начиная с представления исторического примера, классифицированы современные методы стеганографии. Предложена структурная схема стеганографической системы, которая основана на дальнейших исследованиях. Описана имитационная программная модель, называемая ⊕Highly Undetectable steGOsystem, или, сокращенно, «стегосистема ⊕HUGO», реализующая стеганографический метод передачи секретного сообщения, встроенного в неподвижное цифровое изображение. Также рассматривается принцип работы имитационной программной модели и ее стеганографическое обоснование. В качестве алгоритма реализации применен криптографический алгоритм гаммирования, использующий функцию биективного сложения по модулю два, условно обозначаемую ⊕. Авторы определяют сложность обнаружения изменения контейнера в этом методе встраивания путем вычисления коэффициента корреляции Пирсона. Показано, что данная модель успешно повысила информационную безопасность при передаче секретной информации в различных системах электронного документооборота. Разработанная программная модель намного эффективнее алгоритма LSB, что определяется более высокой производительностью и обеспечивает более высокую устойчивость к обнаружению.**

***Ключевые слова:* имитационная программная модель, высоконеобнаруживаемая стегосистема, стегосистема ⊕HUGO, криптографический алгоритм гаммирования, биективное сложение по модулю два, коэффициент корреляции Пирсона.**

Литература

1. Westfield, A. F5 — A Steganographic Algorithm: High Capacity Despite Better Steganalysis // Proceedings of the Fourth International Workshop of Information Hiding (IH 2001), (Pittsburgh, PA, USA, 25–27 April 2001) / I. S. Moskowitz (ed.) Lecture Notes in Computer Science. 2001. Vol. 2137. Pp. 289–302. DOI: 10.1007/3-540-45496-9_21.

2. Fridrich, J. Perturbed Quantization Steganography / J. Fridrich, M. Goljan, D. Soukal // Multimedia Systems. 2005. Vol. 11, Is. 2. Pp. 98–107. DOI: 10.1007/s00530-005-0194-3.

3. Fridrich, J. Statistically Undetectable JPEG Steganography: Dead Ends Challenges, and Opportunities / J. Fridrich, T. Pevný, J. Kodovský // Proceedings of the Ninth Workshop on Multimedia and Security (MM&Sec '07), (Dallas, TX, USA, 20–21 September 2007). — New York: Association for Computing Machinery, 2007. — Pp. 3–14.
DOI: 10.1145/1288869.1288872.

4. Filler, T. Gibbs Construction in Steganography / T. Filler, J. Fridrich // IEEE Transactions on Information Forensics and Security. 2010. Vol. 5, Is. 4. Pp. 705–720.
DOI: 10.1109/TIFS.2010.2077629.

5. Holub, V. Designing Steganographic Distortion Using Directional Filters / V. Holub, J. Fridrich // Proceedings of the Fourth IEEE International Workshop on Information Forensics and Security (WIFS 2012), (Costa Adeje, Spain, 02–05 December 2012). — Institute of Electrical and Electronics Engineers, 2012. — Pp. 234–239. DOI: 10.1109/WIFS.2012.6412655.

6. Fridrich, J. Multivariate Gaussian Model for Designing Additive Distortion for Steganography / J. Fridrich, J. Kodovský // Proceedings of the 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), (Vancouver, Canada, 26–31 May 2013). — Institute of Electrical and Electronics Engineers, 2013. — Pp. 2949–2953.
DOI: 10.1109/ICASSP.2013.6638198.

7. Holub, V. Universal Distortion Function for Steganography in an Arbitrary Domain / V. Holub, J. Fridrich, T. Denemark // EURASIP Journal on Information Security. 2014. Art. No. 1. 13 p. DOI: 10.1186/1687-417X-2014-1.

8. Sedighi, V. Content-Adaptive Pentary Steganography Using the Multivariate Generalized Gaussian Cover Model / V. Sedighi, J. Fridrich, R. Cogranne // Proceedings of the SPIE/IS&T Electronic Imaging 2015 Media Watermarking, Security, and Forensics (San Francisco, CA, USA, 09–11 February 2015) / A. M. Alattar, [et al.] (eds) Proceedings of SPIE. 2015. Vol. 9409. Art. No. 94090H. 13 p.
DOI: 10.1117/12.2080272.

9. Denemark, T. Steganography with Multiple JPEG Images of the Same Scene / T. Denemark, J. Fridrich // IEEE Transactions on Information Forensics and Security. 2017. Vol. 12, Is. 10. Pp. 2308–2319.
DOI: 10.1109/TIFS.2017.2705625.

10. Denemark, T. Natural Steganography in JPEG Compressed Images / T. Denemark, P. Bas, J. Fridrich // Electronic Imaging. 2018. Is. 7. Art No. 316. 10 p.
DOI: 10.2352/ISSN.2470-1173.2018.07.MWSF-316.

11. Cachin, C. An Information-Theoretic Model for Steganography // Proceedings of the Second International Workshop of Information Hiding (IH 1998), (Portland, OR, USA, 14–17 April 1998) / D. Aucsmith (ed.) Lecture Notes in Computer Science. 1998. Vol. 1525. Pp. 306–318.
DOI: 10.1007/3-540-49380-8_21.

12. Modeling the Security of Steganographic Systems / J. Zöllner, H. Federrath, H. Klimant, [et al.] // Proceedings of the Second International Workshop of Information Hiding (IH 1998), (Portland, OR, USA, 14–17 April 1998) / D. Aucsmith (ed.) Lecture Notes in Computer Science. 1998. Vol. 1525. Pp. 344–354. DOI: 10.1007/3-540-49380-8_24.

13. Anderson, R. On the Limits of Steganography / R. Anderson, F. A. P. Petitcolas // IEEE Journal of Selected Areas in Communications. 1998. Vol. 16, Is. 4. Pp. 474–482. DOI: 10.1109/49.668971.

14. Anderson, R. Stretching the Limits of Steganography // Proceedings of the First International Workshop of Information Hiding (IH 1996), (Cambridge, United Kingdom, 30 May–01 June 1996) / R. Anderson (ed.) Lecture Notes in Computer Science. 1996. Vol. 1174. Pp. 39–48. DOI: 10.1007/3-540-61996-8_30.

15. Кустов, В. Н. Использование дискретного вейвлет-преобразования для внедрения информации в изображения / В. Н. Кустов, Д. К. Процко // Научные тенденции: Вопросы точных и технических наук: Сборник научных трудов по материалам XVII Международной научной конференции (Санкт-Петербург, Россия, 12 июня 2018 г.). — Санкт-Петербург: Изд-во ЦНК МОАН, 2018. — С. 15–20. DOI: 10.18411/spc-12-06-2018-05.

16. Peak signal-to-noise ratio // Wikipedia. URL: http://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio (дата обращения 07.12.2021).

17. Correlation // Wikipedia. URL: http://en.wikipedia.org/wiki/Correlation (дата обращения 07.12.2021).

DOI: 10.24412/2413-2527-2022-331-37-45

Cluster Load Balancing Algorithms Based on Shortest Queue Models

PhD V. A. Goncharenko, Grand PhD V. A. Lokhvitsky
Mozhaisky Military Space Academy
Saint Petersburg, Russia
vlango@yandex.ru, lokhv_va@mail.ru

Abstract. **Various algorithms for redistributing tasks in cluster computing systems are described. The results of calculating the probabilistic-time characteristics of the system with connection to the shortest queue and transitions between queues are presented. A number of models with different performance and node failures, with delays in the transition between nodes are described. The results of analytical and simulation modeling of the considered systems are compared.**

Keywords: **cluster, load balancing, shortest queue models, queue theory, dispatching, transition between queues, join-the-shortest-queue.**

INTRODUCTION

Cluster technologies are currently widely used to solve the problems of ensuring the stability of the functioning and survivability of computing systems (CS) [1]. At the same time, there are cluster systems for various purposes – to increase fault tolerance by duplicating calculations (HA-clusters, High-availability), to ensure a uniform load of cluster nodes by redistributing it (LB-clusters, Load Balancing) or to ensure high performance by parallelizing calculations between cluster nodes (HPC clusters, High performance computing). It is also possible to organize the work of a computing cluster in a mixed mode – with switching functions.

Let's consider in more detail the problem of optimal load redistribution in cluster computing systems. It is relevant in solving problems of both optimizing bandwidth and increasing fault tolerance of distributed computing systems [1, 2]. Examples of such systems can be database query processing systems, Web factories, firewalls, mail and Web traffic content analysis systems, where sufficiently high response times are required.

One of the load balancing mechanisms is dispatching incoming service requests. This mechanism redistributes the workload between several servers of the cluster system, which in general may have different performance. If they fail, the load is redistributed to other nodes of the cluster. At the same time, in a distributed system, there may be delays in transferring the load from one processing node to another.

The objective of the article is to consider algorithms and analytical and simulation models of load balancing with heterogeneous cluster architecture and various methods of dispatching organization.

ALGORITHMS FOR DISPATCHING TASKS IN CLUSTERS

Consider the models of a cluster computing system (Fig. 1), where the distribution of tasks between nodes is carried out by a hardware or software dispatcher (switching processor, specialized load balancing server, special software). Each node has the necessary means to organize a queue of tasks. The dispatcher has either a centralized or distributed implementation, when the dispatcher functions are performed in each of the nodes under consideration. The homogeneity of the CS nodes is not mandatory, i. e. nodes of different performance are allowed.

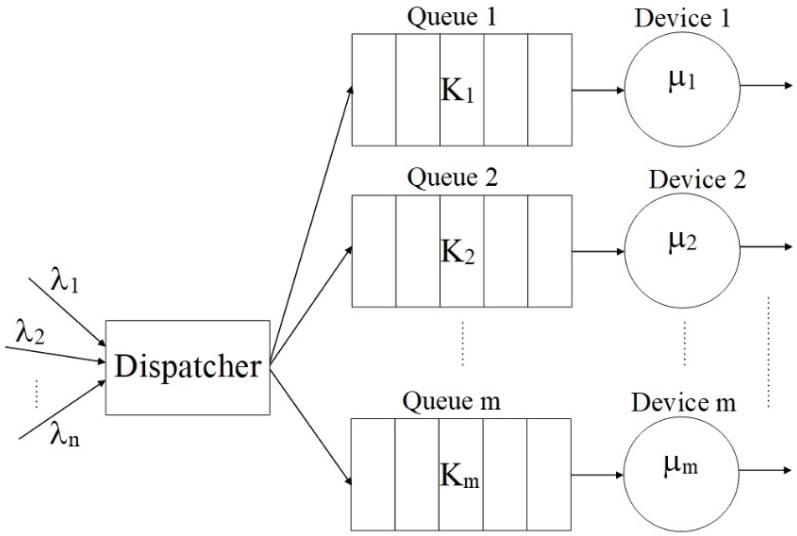

Fig. 1. Model with dispatching

There are deterministic, stochastic and adaptive dispatch algorithms.

1. *Deterministic algorithms.*

The dispatcher directs the received task to a specific server:

a) fixed dispatching (each task flow is sent to its «own» predefined server);

b) cyclic dispatching (each newly received task is sent to the next server by number, for example, in Round Robin, WRR, DRR cyclic algorithms).

2. *Stochastic algorithms*.

The dispatcher directs tasks to one of the cluster nodes with equal probability (as a generalization– with a given probability, depending on performance and other factors). The algorithm does not take into account the current degree of node load.

3. *Adaptive algorithms*.

The dispatcher directs the next incoming task based on the ratio of queue lengths to individual servers (as a generalization — based on the ratio of productivity or serviceability of servers [3].

Obviously, adaptive algorithms do a better job with load balancing [4], but require additional information. A feature of the algorithms is the possibility of making a decision on load redistribution based on operational dynamically changing information, for example, information about queue lengths to servers [5, 6].

A large number of publications are devoted to the study of the problem of the shortest queue [7–10]. For the first time such a model was considered in [11]. At the same time, there are no exact analytical calculations in the literature for models with more than two servers – approximation methods are used [12]. Thus, approximations of the average response time for the case of K queues are presented in [13], assuming that different queue lengths can differ by no more than one. The boundaries for the average residence time of requirements in a two-channel system were obtained in [7] using linear programming methods.

In [14], an approximation was developed to generalize the shortest queue model, namely, the model with the shortest expected delay in routing clients to servers with different operating speeds.

Below we will consider various strategies for organizing the work of adaptive dispatch algorithms [3]:

1. The dispatcher receives or does not receive additional information about the performance of nodes.

2. The dispatcher directs the task to the node with the shortest queue length or (if additional information is available) to the node with the lowest delay (the ratio of queue length to node performance).

3. If the queue lengths (delays) are equal, the dispatcher directs the task:

a) to the node specified for each task flow;

b) to the next node after the last node that received the task;

c) to any node with equal probability;

d) to the node with the highest performance;

e) to any node with a probability proportional to performance.

4. If the node capacities are equal, the dispatcher directs the task:

a) to the node specified for each task flow;

b) to the next node after the last node that received the task;

c) to any node with equal probability.

5. In addition to dispatching input tasks, it is possible to organize the transition of tasks between queues. After servicing the next task, when the difference between the shortest queue and the longest queues is more than ΔL (sensitivity threshold):

a) redistributes the last task of the nearest of the longest queues preceding the shortest queue to the shortest queue;

b) redistributes to the shortest queue the last task of one of the longest queues, selected equally likely;

c) no longer redistributes tasks from the longest queues.

Figure 2 shows the classification of algorithms for dispatching input tasks depending on the selected model.

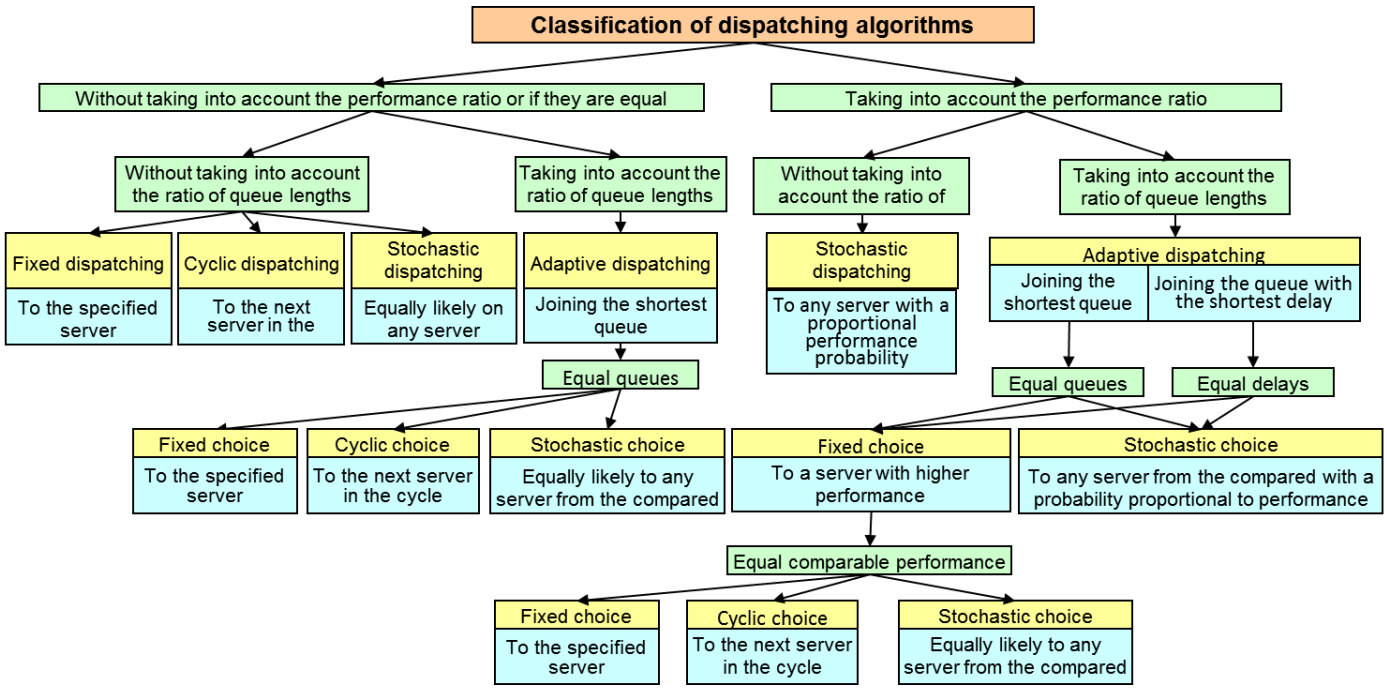

Fig. 2. Classification of dispatching algorithms

A MODEL WITH JOINING THE SHORTEST QUEUE AND TRANSITIONS BETWEEN QUEUES

Despite the considerable interest in models with the shortest queue, the analytical results are still very modest, even with the simplest assumptions about the input flow and service flows. At the same time, models and algorithms have been developed that, in addition to joining the shortest queue, allow requests to move between queues during the waiting process. For the first time such a two-channel model was considered in [15]. In [5], expressions are obtained for the main characteristics of the model with connection to the shortest queue and transition between queues based on a two-channel system with one input flow, different channel capacities and an infinite queue. The algorithm of functioning and a device for modeling a two-channel system with connection to the shortest queue and transition between queues are described in [6].

Let's call this system a system with *join the shortest queue* and *transition between queues* — JSQ/TBQ. Consider the case of a two-channel system ($m = 2$) with a limited capacity of queue buffers K_i, $i = 1, 2$. Figure 3 shows the block diagram of this system [16].

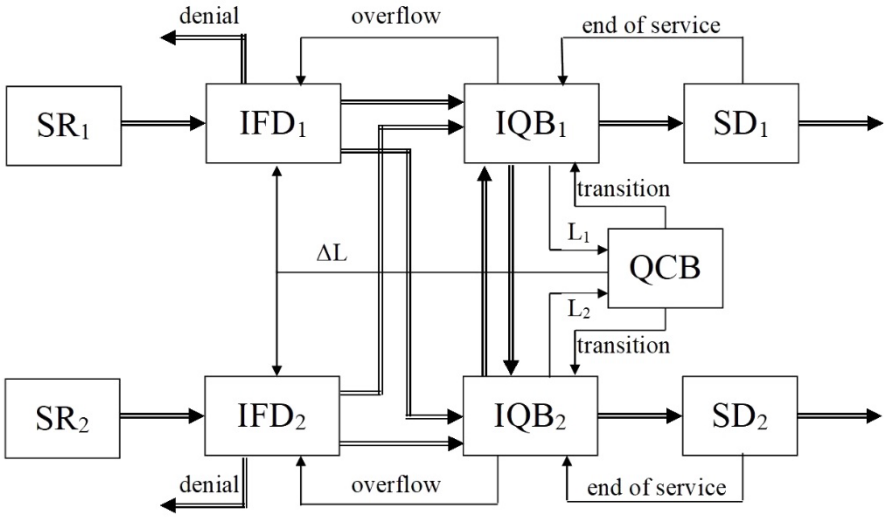

Fig. 3. Block diagram of the JSQ/TBQ system

The figure uses abbreviations:

SR_i — i-th source of requests;

IFD_i — i-th input flow dispatcher;

IQB_i — i-th input queue block;

QCB — queue comparison block;

SD_i — i-th service device.

The total input flow of requests from the sources of requests (SR_i) will be distributed in such a way as to load both nodes most optimally, since any task that enters the system will join the shortest queue. To do this, the input flow dispatchers (IFD_i) use information about the difference in the lengths of the queues of the input queue blocks (IQB_i) $\Delta L = L1 - L2$ from the queue comparison block (QCB). In order to reduce the difference in queue lengths that occurs during the waiting for service due to the random nature of the request service process, a mechanism for transferring requests between queues is used. We will assume that the transfer of requests from queue to queue is carried out at $|\Delta L| \geq 2$.

We describe the algorithm of the system functioning [6, 7].

Step 1. Requests received from SR_i to IFD_i, depending on the state of the system:

a) are sent to the queue of the first node if $\Delta L<0$;

b) are sent to the queue of the second node if $\Delta L>0$;

c) are removed from the system if the queues are full.

Step 2. In cases a) and b) of step 1, the task enters the corresponding service channel and becomes in the service queue in the input queue block (IQB). In case c), the request simply does not enter the system and is deleted.

Step 3. The ratio of queue lengths is reported to the dispatcher by the QCB, which receives information about the lengths of queues $L1$ and $L2$ from both IQB. In case of inequality of queues depending on the signal ΔL, the dispatcher sends the request to the shortest queue. If the queues are equal, then the request is sent to the channel to which it was received.

Step 4. In case of queue overflow, IQB signals to the dispatcher, who closes access to this channel for requests and transfers it to the neighboring channel. When both queues overflow, in addition to the overflow signal, the IFD_i receives a queue equality signal from the QCB. The receipt of request s in the system is stopped until the seats in the queues are vacated.

Step 5. From the IQB, the request is sent to the service device (SD_i) for maintenance, the end of which it signals to the IQB in order to accept the next service request and replenish the queue if there was a limit number of requests in it.

Step 6. If there is a difference in the queue lengths of more than one request, the BSO generates a signal for the transition of the last request from a longer queue to the end of a shorter one. In the QCB, after the transfer of the request is completed, the ΔL is changed.

Thus, the alignment of queue lengths occurs not only due to the redistribution of the incoming input flow, but also due to the transfer of requests between queues. A special case of the system is with one incoming flow and one fiberboard.

The request distribution strategy can be of two types. The first type is when the ratio of the service rates of SD_1 and SD_2 is known, the second is when there is no a priori information about their ratio. In the first case, if the queues are equal, the request is sent to the queue to the SD with greater rate, in the second — with equal probability.

CALCULATION OF THE CHARACTERISTICS
OF THE SHORTEST QUEUE TWO-SERVER MODEL

Consider a two-channel system JSQ/TBQ with two input flows, queue end drives and different node performance (Fig. 4).

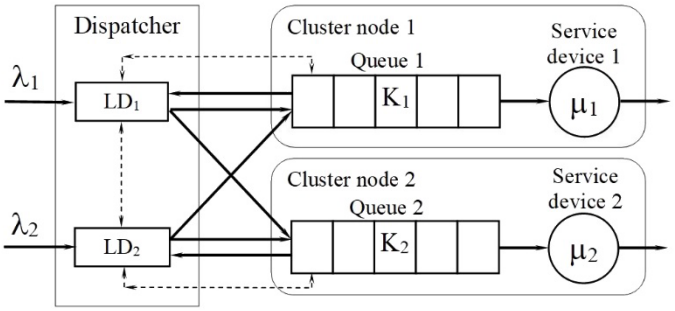

Fig. 4. Model of two-channel Queuing system JSQ/TBQ with finite storage devices

The general dispatcher is distributed, consists of local dispatchers LD_1 and LD_2, exchanging information about the status of queues. Requests come from two different input flows and are sent to the node with the smallest queue. If the queue lengths are equal, the incoming request is sent to a node with a higher service rate, if the same or unknown ratio of service rates is equal — to a node with the same number. During the waiting process, the last task from the longest queue goes to the shortest queue with a queue difference equal to the sensitivity threshold. In the simplest case, the sensitivity threshold is two. The transition time to the next queue, both when a request is received and during the waiting process, is generally not equal to zero. If both queues overflow, the incoming request is rejected.

The transition graph of the system is shown in Figure 5.

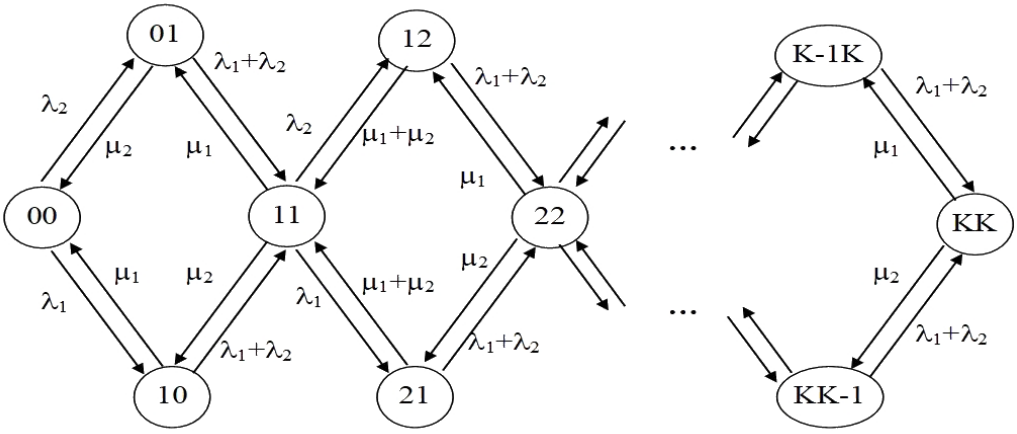

Fig. 5. The transition graph of the system

The states characterize the number of requests in each node. Each arrow is set in accordance with the rate of transitions. At the same time, the number of tasks in each node does not differ from each other by more than 1, which corresponds to the dispatching algorithm. We denote by $P_{i,i}$, $P_{i,i+1}$, $P_{i+1,i}$ the stationary probabilities of the state of the system. Based on the transition graph, in accordance with the conservation laws of queue theory [17], we will compile a system of equations and transform it to the following form:

$$\left.\begin{aligned} p_{01} &= p_{00}\frac{\rho(\lambda_1+\lambda_2)+\lambda_2}{(2\rho+1)\mu_2} \\ p_{10} &= p_{00}\frac{\rho(\lambda_1+\lambda_2)+\lambda_1}{(2\rho+1)\mu_1} \\ &\vdots \\ p_{ii} &= \rho^{2i-1}(p_{10}+p_{01}), i = 1 \div K \\ p_{i,i+1} &= \rho^{2i-1}(p_{10}+p_{01})\frac{\rho^2\mu_1+\lambda_2}{\lambda_1+\lambda_2+\mu_1+\mu_2} \\ p_{i+1,i} &= \rho^{2i-1}(p_{10}+p_{01})\frac{\rho^2\mu_2+\lambda_1}{\lambda_1+\lambda_2+\mu_1+\mu_2} \\ &\vdots \\ p_{KK} &= \rho^{2K-1}(p_{10}+p_{01}). \end{aligned}\right\} \quad (1)$$

Herein $\rho=\lambda/\mu$ — system load factor; $\lambda=\lambda_1+\lambda_2$ — the total arrival rate; $\mu=\mu_1+\mu_2$ — total service rate.

We take $\lambda_1 = r\lambda$, $\mu_1 = s\mu$, where r and s are the coefficients of the asymmetry of the input flow and the service flow. Based on this and the conditions $p_{ij}+p_{ji}=p_{i+j}$, $p_{ii}=p_{2i}$, we bring the system (1) to the following form:

$$\left.\begin{aligned} p_1 &= p_0\frac{\rho^2+\rho(r+s-2rs)}{(2\rho+1)(s-s^2)} \\ p_2 &= \rho\times p_1 \\ &\vdots \\ p_i &= \rho^{i-1}p_1 \\ &\vdots \\ p_{2K} &= \rho^{2K-1}p_1. \end{aligned}\right\} \quad (2)$$

From (2) and the normalization condition (the sum of all probabilities of states is equal to one), we find the probability of a free state of the system:

$$p_0 = \frac{1}{\left[1+\frac{1-\rho^{2K}}{1-\rho}\times\frac{\rho(\rho+r+s-2rs)}{(1+2\rho)(s-s^2)}\right]}. \quad (3)$$

If the node capacities are equal, the formula for p_0 completely coincides with the similar formula for a two-channel system *M/M/2/K*:

$$p_0 = \frac{1-\rho}{1+\rho-2\rho^{2K+1}}. \quad (4)$$

With an infinite queue accumulator, formula (3) takes the form:

$$p_0 = 1/\left[1+\frac{\rho(\rho+r+s-2rs)}{(1-\rho)(1+2\rho)(s-s^2)}\right].$$

And with equal productivity $\mu_1 = \mu_2$:

$$p_0 = (1-\rho)/(1+\rho).$$

Average response time *T* in the system under consideration:

$$T = p_0\frac{(\rho+r+s-2rs)}{\mu(1+2\rho)(s-s^2)}\times\left[\frac{1-\rho^{2K}}{(1-\rho)^2}-\frac{2K\rho^{2K}}{1-\rho}\right]. \quad (5)$$

From (5), you can get the average number of requests in the system: $N = \lambda T$.

It is also of interest what proportion of the total number of requests is served in the first and which in the second node, how it varies depending on the coefficients *s* and *r*.

The probabilities of servicing requests in the corresponding node P_{serv1} and P_{serv2} depend on three events – on the probability of joining the request from the common input flow to the corresponding queue, on the probability of transferring the request from the neighboring queue and the probability of transferring the request to the neighboring queue:

$$P_{serv1} = P_{join1} + P_{trans}^{2\to1} - P_{trans}^{1\to2}; \quad (6)$$

$$P_{serv2} = P_{join2} + P_{trans}^{1\to2} - P_{trans}^{2\to1}, \quad (7)$$

where P_{join1} — probability of joining the input request to the first queue;

P_{join2} — probability of joining the input request to the second queue;

$P_{trans}^{2\to1}$ — the probability of the transition of the request from the second stage to the first;

$P_{trans}^{1\to2}$ — the probability of the transition of the request from the first stage to the second.

$$P_{join1} = p_{01} + p_{12}+\ldots+p_{k-1k} + r\left(p_{00} + p_{11}+\ldots+p_{k-1,k-1}\right);$$

$$P_{join2} = p_{10} + p_{21}+\ldots+p_{kk-1} + (1-r)\left(p_{00} + p_{11}+\ldots+p_{k-1,k-1}\right).$$

It is clear from formulas (8)–(9) that

$$P_{serv1} + P_{serv2} = P_{join1} + P_{join2}.$$

If both queues overflow, the request will be denied service:

$$P_{serv1} + P_{serv2} + p_{den} = 1.$$

The probability of denial of service is defined as the probability that all places in the queue are occupied, i. e.

$$p_{den} = p_{KK} = p_0\times\rho^{2K-1}\times\frac{\rho(\rho+r+s-2rs)}{(1+2\rho)(s-s^2)}.$$

For the probabilities of request transitions between queues after joining the shortest queue, we have:

$$P_{trans}^{2\to1} = \left(p_{12} + p_{23} + \cdots + p_{K-1,K}\right)\times\frac{\mu_1}{\mu} = $$

$$= G\times(\rho s + 1 - r)\times s;$$

$$P_{trans}^{1\to2} = \left(p_{21} + p_{32} + \cdots + p_{K,K-1}\right)\times\frac{\mu_2}{\mu} = $$

$$= G\times(\rho(1-s)+r)\times(1-s),$$

where

$$G = p_0\rho^3\frac{(\rho+r+s-2rs)\times(1-\rho^{2K-2})}{(s-s^2)(1+2\rho)(1+\rho)(1-\rho^2)}.$$

We will determine what the parameters of the system under study should be in order to meet the requirements of the optimality of the service process. It follows from (3) that p_0 will be the maximum at the minimum of the function

$$z = \frac{\rho(\rho+r+s-2rs)}{(1+2\rho)(s-s^2)}.$$

Let *r* and ρ be constant. Then we have

$$\frac{dz}{ds} = \frac{\rho(1-2r)s^2 - \rho(1-2s)(r+\rho)}{(1+2\rho)(s-s^2)^2}.$$

Equate the numerator to 0:

$$\rho(1-2r)s^2 - \rho(1-2s)(r+\rho) = 0;$$

$$(1-2r)s^2 + 2(r+\rho)s - (r+\rho) = 0.$$

The first root of the quadratic equation:

$$s = \frac{\sqrt{\rho^2 - r^2 + \rho + r} - \rho - r}{1-2r}. \quad (8)$$

The second root is negative.

Thus, the maximum value of p_0 will be for the corresponding load with a certain ratio of node service rates determined by expression (8). Let's make a table (Table 1) the maxima p_0 corresponding to the optimal values of the coefficient *s* at different load factors ρ.

Table 1

The maximum values of p_0 at different *r* and ρ and optimal values of *s*, *K* = 25

		ρ = 0.1	**ρ = 0.5**	**ρ = 0.9**
***r* = 0.1**	s_{opt}	0.309	0.396	0.427
	p_{0max}	0.838	0.343	0.054
***r* = 0.3**	s_{opt}	0.414	0.449	0.464
	p_{0max}	0.822	0.336	0.053
***r* = 0.5**	s_{opt}	0.500	0.500	0.500
	p_{0max}	0.818	0.333	0.053
***r* = 0.7**	s_{opt}	0.586	0.551	0.536
	p_{0max}	0.822	0.336	0.053
***r* = 0.9**	s_{opt}	0.691	0.604	0.573
	p_{0max}	0.838	0.343	0.054

For r = 0.5, expression (8) does not make sense, since at this point *s* cannot be optimal, and p_0 cannot be greater than that of the M/M/2/K system. The boundary value will be $s = 0.5$, at which the p_0 of the system under study coincides with the p_0 of the system M/M/2/K.

We define the boundaries within which the values of the coefficients r and s should lie, so that the system under study is not inferior to the system M/M/2/K in probabilistic characteristics. To do this, we compare formulas (3) and (4) and set the condition:

$$1-\rho+(1-\rho^{2K})\times\frac{\rho(\rho+r+s-2rs)}{(1+2\rho)(s-s^2)}\le 1+\rho-2\rho^{2K+1}.$$

As a result of the transformations, we get

$$(4\rho+2)s^2-(1+2r+4\rho)s+(\rho+r)\le 0.$$

From the square inequality we obtain two roots that define the boundaries (sectors) of the optimal values of the coefficients *s* and *r*:

$$s_1=\frac{2r+4\rho+1+(2r-1)}{4(2\rho+1)}=\frac{\rho+r}{2\rho+1}; \qquad (9)$$

$$s_2=\frac{2r+4\rho+1-(2r-1)}{4(2\rho+1)}=0.5. \qquad (10)$$

Let's make a table of the boundary values of *s* and *r* for different ρ (Table 2). For example, using the values (8)–(10), we will plot a graph for ρ = 0.1 (Fig. 6).

Table 2

Boundary values of the coefficient *s*

	ρ = 0.1	**ρ = 0.3**	**ρ = 0.5**	**ρ = 0.7**	**ρ = 0.9**
***r* = 0.0**	0.080	0.188	0.250	0.292	0.321
***r* = 0.5**	0.500	0.500	0.500	0.500	0.500
***r* = 1.0**	0.917	0.813	0.750	0.708	0.679

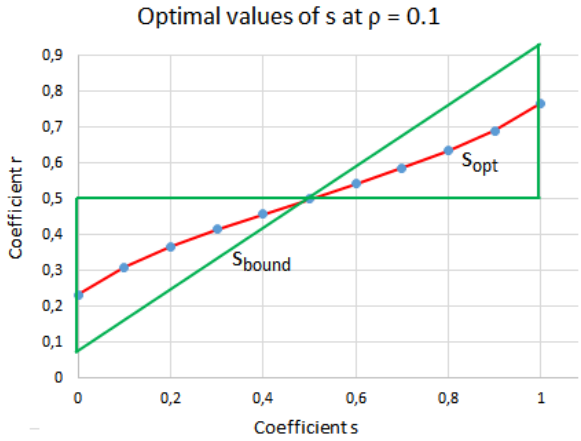

Fig. 6. Optimal values of the coefficient *s*

The graph shows that at $r = 0.5$ there is only one optimal point $s = 0.5$, at which the characteristics of the system under study and the system *M/M/2/K* coincide. As the load increases, the angle of the sector, and hence the range of optimal values of *r* and *s*, decreases.

A MODEL WITH A DELAY IN TRANSMISSION BETWEEN QUEUES

Let's now briefly consider the case when the delay in transferring tasks from the local dispatcher or from the queue to another queue is not zero. This is possible in global cluster systems, in which the transmission time is comparable to the service time in the nodes, and not taking into account this delay time will introduce a significant error in the calculations. The transfer of an request to a neighboring queue occurs when a serviced request drops out of a node with less than 1 number of request s in the queue and the difference in queues reaches 2 (Figure 7). Also, the transition occurs when the request arrives at its «own» node, in which there is 1 more in the queue than in the neighboring queue. In principle, the trigger threshold may be higher to prevent frequent transitions between queues.

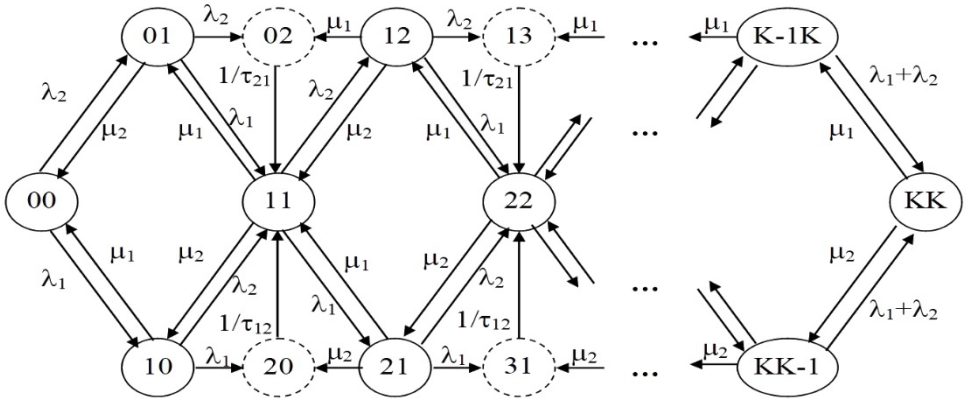

Fig. 7.The transition graph of the system

The temporary states *i* - 1, *i* + 1 and *i* + 1, *i* - 1 are marked with a dotted line, because the system, after a random delay time in transmitting a request from channel to channel (τ_{12} or τ_{21}), enters the equilibrium state *i, i*. Then the rate of the transition from states *i, i* - 1 and *i* - 1, *i* to state *i, i* through intermediate states *i* - 1, *i* + 1 and *i* + 1, *i* - 1 can be expressed from the equations:

$$\frac{1}{\lambda_1^*}=\frac{1}{\lambda_1}+\tau_{12};\ \frac{1}{\lambda_2^*}=\frac{1}{\lambda_2}+\tau_{21};$$

$$\frac{1}{\mu_1^*}=\frac{1}{\mu\lambda_1}+\tau_{21};\ \frac{1}{\mu_2^*}=\frac{1}{\mu_2}+\tau_{12}.$$

We get:

$$\lambda_1^* = \frac{\lambda_1}{1+\lambda_1\tau_{12}}\ ;\ \lambda_2^* = \frac{\lambda_2}{1+\lambda_2\tau_{21}}\ ;$$

$$\mu_1^* = \frac{\mu_1}{1+\mu_1\tau_{21}}\ ;\ \mu_2^* = \frac{\mu_2}{1+\mu_2\tau_{12}}.$$

Composing a transition graph and a system of equations based on it, we find expressions for the stationary probabilities of the states of the system. So, for $\lambda_1 = \lambda_2 = \lambda$, $\mu_1 = \mu_2 = \mu$ and delay $\tau_{12} = \tau_{21} = \tau_d$ the probability of downtime p_0 has the form:

$$p_0 = \frac{1}{\left[1 + \frac{1-R^K}{1-R} \times (2\rho + \rho^2 \frac{2+\lambda\tau_d}{1+\lambda\tau_d}\right]}, \qquad (11)$$

where

$$R = \rho^2 \frac{2+\lambda\tau_d}{1+\lambda\tau_d} \times \frac{1+\mu\tau_d}{2+\mu\tau_d}.$$

For $\tau_d = 0$, formula (11) reduces to expression (4).

Simulation models of adaptive dispatching simulation results

Unfortunately, not all models can be studied analytically. Therefore, during the research, the following variants of two- and three-channel simulation models of cluster systems with a finite queue storage in the GPSS World language were also developed and investigated:

1) models with the shortest queue and transition between queues;

2) models with node failures;

3) models with delayed transmission of requests between nodes;

4) models with the shortest delay in the system (the ratio of queue length to node performance).

During the simulation, various load variants were tested: subcritical (ρ = 0.5), critical (ρ = 0.95) and supercritical (ρ = 1.5; 2.0).

A comparative analysis of 4 two-channel simulation models was carried out: M1 (the system with the lowest delay), M2 (the system with the shortest queue), M3 (the M/M/2/K system), M4 (two single M/M/1/K systems). The M1–M2 models also have a mechanism for setting a non-zero delay in the transfer of requests between nodes. Three analytical models M2–M4 are also considered.

In the model with the lowest delay M1, the application is attached to the node with the lowest ratio of queue length to service intensity. In general, the model shows better results compared to the model with the shortest queue, but the implementation of the dispatcher will be more difficult due to the calculation of the node with the least delay, which may affect the decision time on the distribution of the next request.

Simulation also confirmed that it is impossible to achieve an advantage over a two-channel system with the same service intensities. Models with dispatching asymptotically approach the characteristics of the M/M/2/K system. But the gain can be achieved with a heterogeneous system architecture. In addition, models with adaptive dispatching allow us to study systems with global clustering [1].

Conclusions

With an increase in the number of nodes and, accordingly, queues, the queue selection strategy and the analytical description of the model become much more complicated.

The inclusion in the block diagram of a model with the shortest queue of connections for the transition of requirements between queues significantly improves its characteristics, which is explained by the greater adaptability of the model to load balancing.

Response time in the system JSQ/TBQ can achieve an advantage in comparison with the M/M/2/K system with different channel capacities and a certain optimal ratio.

Taking into account the performance of nodes during load redistribution significantly improves the time characteristics of job maintenance, but complicates the implementation of the dispatcher.

Acknowledgments

The study was carried out with the financial support of the Russian Foundation for Basic Research, project No. 18-29-22064\18.

References

1. Zaleshchansky B. D., Chernikhov D. Ya. Klasternaya tekhnologiya i zhivuchest globalnykh avtomatizirovannykh system [Cluster technology and the survivability of global automated systems]. Moscow, Finance and Statistics Publishers, 2005. 384 p. (In Russian)

2. Dodonov A. G., Kuznetsova M. G., Gorbachik E. S. Vvedenie v teoriyu zhivuchesti vychislitelnykh system [Introduction to the theory of survivability of computing systems]. Kyiv, Naukova Dumka Publishers, 1990, 184 p. (In Russian)

3. Goncharenko V. A. Modeli adaptivnogo pereraspredeleniya nagruzki v klasternykh vychislitelnykh sistemakh [Models of Adaptive Load Redistribution in Cluster Computing Systems, *Izvestiya vysshikh uchebnykh zavedeniy. Priborostroenie [Journal of Instrument Engineering]*, 2008, Vol. 51, No. 3, Pp. 32–37. (In Russian)

4. Doniants V. N., Udalova T. V. Pereraspredelenie vychislitelnoy nagruzki v lokalnykh setyakh EVM [Redistribution of Computing Load in Local Computer Networks. In: *Lazarev V. G., Chernyaev V. G. (eds.) Upravlenie protsessami i resursami v raspredelennykh sistemakh: Sbornik nauchnykh trudov [Process and resource management in distributed systems*: *Collection of scientific papers]*. Moscow, Nauka Publishers, 1989, Pp. 57–64. (In Russian)

5. Gortsev A. M. Dvukhkanalnaya sistema massovogo obsluzhivaniya s perekhodom trebovaniy iz odnoy ocheredi v druguyu [A Two-Channel Queuing System with the Transition of Requirements from One Queue to Another], *Avtomatika i telemekhanika [Automation and Remote Control]*, 1981, No. 6, Pp. 189–192. (In Russian)

6. Goncharenko V. A., Filimonikhin G. V. Ustroystvo dlya modelirovaniya dvukhkanalnoy sistemy massovogo obsluzhivaniya [Device for Modeling a Two-Channel Queuing System]. Certificate of Authorship SU No. 1509928, published at September 23, 1989, 6 p. (In Russian).

7. Halfin S. The Shortest Queue Problem, *Journal of Applied Probability*, 1985, Vol. 22, Is. 4, Pp. 865–878.
DOI: 10.2307/3213954.

8. Dester P. S., Fricker C., Tibi D. Stationary Analysis of the Shortest Queue Problem, *Queueing Systems: Theory and Applications*, 2017, Vol. 87, No. 3–4, Pp. 211–243. DOI: 10.1007/s11134-017-9556-8.

9. Adan I. J. B. F., Wessels J., Zijm W. H. M. Analysis of the Symmetric Shortest Queue Problem, *Communications in Statistics. Stochastic Models*, 1990, Vol. 6, Is. 4. Pp. 691–713. DOI: 10.1080/15326349908807169.

10. Cohen J. W. Analysis of the Asymmetrical Shortest Two-Server Queueing Model, *Journal of Applied Mathematics and Stochastic Analysis*, 1998, Vol. 11, Is. 2, Pp. 115–162. DOI: 10.1155/S1048953398000112.

11. Haight F. A. Two Queues in Parallel, *Biometrika*, 1958, Vol. 45, Is. 3–4, Pp. 401–410. DOI: 10.1093/biomet/45.3-4.401.

12. Nelson R. D., Tanatawi A. N. Approximating Task Response Times in ForkJoin Queues. In: *Gelenbe E. (ed.) High Performance Computer Systems: Proceedings of the International Symposium on High Performance Computer Systems (Paris, France, December 14–16, 1987)*. Amsterdam, North Holland Publishing Company, 1988, Pp. 157–167.

13. Nelson R. D., Philips T. K. An Approximation to the Response Time for Shortest Queue Routing, *ACM SIGMETRICS Performance Evaluation Review*, 1989, Vol. 17, Is. 1, Pp. 181–189. DOI: 10.1145/75372.75392.

14. Lui J. S. C., Muntz R. R. Algorithmic Approach to Bounding the Mean Response Time of a Minimum Expected Delay Routing System, *ACM SIGMETRICS Performance Evaluation Review*, 1992, Vol. 20, Is. 1, Pp. 140–151. DOI: 10.1145/149439.133099.

15. Saaty T. L. Elementy teorii massovogo obsluzhivaniya i ee prilozheniya [Elements of Queueing Theory: With Applications]. Moscow, Soviet Radio Publishing House, 1971, 520 p.

16. Goncharenko V. A. Analiz adaptivnykh algoritmov dispetcherizatsii zadaniy v klasterakh informatsionno-vychislitelnykh setey [Analysis of Adaptive Algorithms for Dispatching Tasks in Clusters of Information and Computing Networks]. In: *Kudryashov I. A. (ed.) Sbornik algoritmov i programm tipovykh zadach. Vypusk 24 [Collection of Algorithms and Programs for Typical Tasks. Issue 24]*. Moscow, Ministry of Defense of the Russian Federation, 2006, Pp. 222–233. (In Russian)

17. Ryzhikov Yu. I. Algoritmicheskiy podkhod k zadacham massovogo obsluzhivaniya [Algorithmic approach to queuing tasks]: Monograph. Saint Petersburg, Mozhaisky Military Space Academy, 2013, 496 p. (In Russian)

DOI: 10.24412/2413-2527-2022-331-37-45

Алгоритмы балансировки нагрузки кластеров на основе моделей с кратчайшей очередью

к.т.н. В. А. Гончаренко, д.т.н. В. А. Лохвицкий
Военно-космическая академия имени А. Ф. Можайского
Санкт-Петербург, Россия
vlango@mail.ru, lokhv_va@mail.ru

Аннотация. **Описаны различные алгоритмы перераспределения заданий в кластерных вычислительных системах. Приведены результаты расчета вероятностно-временных характеристик системы с присоединением к кратчайшей очереди и переходами между очередями. Описан ряд моделей с разными производительностями и отказами узлов, с задержками при переходе между узлами. Выполнено сопоставление результатов аналитического и имитационного моделирования рассматриваемых систем.**

Ключевые слова: **кластер, балансировка нагрузки, модели с кратчайшей очередью, теория очередей, диспетчеризация, переход между очередями, присоединение к кратчайшей очереди.**

Литература

1. Залещанский, Б. Д. Кластерная технология и живучесть глобальных автоматизированных систем / Б. Д. Залещанский, Д. Я. Чернихов. — Москва: Финансы и статистика, 2005. — 384 с.

2. Додонов, А. Г. Введение в теорию живучести вычислительных систем. / А. Г. Додонов, М. Г. Кузнецова, Е. С. Горбачик; АН УССР, Ин-т проблем регистрации информации. — Киев: Наукова думка, 1990. — 181 с.

3. Гончаренко, В. А. Модели адаптивного перераспределения нагрузки в кластерных вычислительных системах // Известия высших учебных заведений. Приборостроение. 2008. Т. 51, № 3. С. 32–37.

4. Донианц, В. Н. Перераспределение вычислительной нагрузки в локальных сетях ЭВМ / В. Н. Донианц, Т. В. Удалова // Управление процессами и ресурсами в распределенных системах / АН СССР, Ин-т проблем передачи информации; отв. ред. В. Г. Лазарев, В. Г. Черняев. — Москва: Наука, 1989. — С. 57–64.

5. Горцев, А. М. Двухканальная система массового обслуживания с переходом требований из одной очереди в другую // Автоматика и телемеханика. 1981. № 6. С. 189–192.

6. Авторское свидетельство № 1509928 СССР, G 06 F 15/20. Устройство для моделирования двухканальной системы массового обслуживания: № 4364699/24-24: заявл. 13.01.1988: опубл. 23.09.1989 / Гончаренко В. А., Филимонихин Г. В.; заявитель Военный инженерный краснознаменный институт имени А. Ф. Можайского. — 6 с.

7. Halfin, S. The Shortest Queue Problem // Journal of Applied Probability. 1985. Vol. 22, Is. 4. Pp. 865–878. DOI: 10.2307/3213954.

8. Dester, P. S. Stationary Analysis of the Shortest Queue Problem / P. S. Dester, C. Fricker, D. Tibi // Queueing Systems: Theory and Applications. 2017. Vol. 87, No. 3–4. Pp. 211–243. DOI: 10.1007/s11134-017-9556-8.

9. Adan, I. J. B. F. Analysis of the Symmetric Shortest Queue Problem / I. J. B. F. Adan, J. Wessels, W. H. M. Zijm // Communications in Statistics. Stochastic Models. 1990. Vol. 6, Is. 4. Pp. 691–713. DOI: 10.1080/15326349908807169.

10. Cohen, J. W. Analysis of the Asymmetrical Shortest Two-Server Queueing Model // Journal of Applied Mathematics and Stochastic Analysis. 1998. Vol. 11, Is. 2. Pp. 115–162. DOI: 10.1155/S1048953398000112.

11. Haight, F. A. Two Queues in Parallel // Biometrika. 1958. Vol. 45, Is. 3–4. Pp. 401–410. DOI: 10.1093/biomet/45.3-4.401.

12. Nelson, R. D. Approximating Task Response Times in ForkJoin Queues / R. D. Nelson, A. N. Tantawi // High Performance Computer Systems: Proceedings of the International Symposium on High Performance Computer Systems (Paris, France, 14–16 December 1987) / E. Gelenbe (ed.). — Amsterdam: North Holland Publishing Company, 1988. — Pp. 157–167.

13. Nelson, R. D. An Approximation to the Response Time for Shortest Queue Routing / R. D. Nelson, T. K. Philips // ACM SIGMETRICS Performance Evaluation Review. 1989. Vol. 17, Is. 1. Pp. 181–189. DOI: 10.1145/75372.75392.

14. Lui, J. S. C. Algorithmic Approach to Bounding the Mean Response Time of a Minimum Expected Delay Routing System / J. S. C. Lui, R. R. Muntz // ACM SIGMETRICS Performance Evaluation Review. 1992. Vol. 20, Is. 1. Pp. 140–151. DOI: 10.1145/149439.133099.

15. Саати, Т. Л. Элементы теории массового обслуживания и ее приложения = Elements of Queueing Theory: With Applications / Пер. с англ. Е. Г. Коваленко; под ред. И. Н. Коваленко. — 2-е изд. — Москва: Советское радио, 1971. — 520 с.

16. Гончаренко, В. А. Анализ адаптивных алгоритмов диспетчеризации заданий в кластерах информационно-вычислительных сетей // Сборник алгоритмов и программ типовых задач / Под ред. И. А. Кудряшова. Вып. 24. — Москва: Министерство обороны РФ, 2006. — С. 222–233.

17. Рыжиков, Ю. И. Алгоритмический подход к задачам массового обслуживания: Монография. — Санкт-Петербург: ВКА им. А. Ф. Можайского, 2013. — 496 с.

DOI: 10.24412/2413-2527-2022-331-46-51

The Choice of Structures of Heterogeneous Information-Computer Systems Based on the Apparatus of Genetic Algorithms

PhD I. V. Zakharov, PhD A. O. Shushakov, S. S. Zykova
Mozhaisky Military Space Academy
Saint Petersburg, Russia
x.vano-z80@yandex.ru, shushakovaleksei@mail.ru

Abstract. **Sufficiently adequate dynamic models of the functioning of complex systems are characterized by high computational complexity, which leads to a significant complexity of optimization procedures. Therefore, the solution of the problem of combinatorial optimization by a complete search of possible solutions in practice is unacceptable.**

The advantages of evolutionary search as a method of combinatorial optimization of the structure of the information and computing system is the possibility of various ways of setting the target function and types of optimization variables, as well as in the use of probabilistic, rather than deterministic rules for finding solutions. A method of formalizing the structure of a heterogeneous information and computing system is proposed, which takes into account its hierarchical-network structure.

The presented approach allows by selecting rational parameters of the genetic algorithm and using the stochastic fitness function with a variable coefficient of variation to achieve a satisfactory speed of its convergence with a large dimension of the task. Examples of application of this method to the search for a rational structure of the computing system are given.

Keywords: **computer system, evolution search, genetic algorithm, stochastic fitness-function.**

Introduction

Modern information-computer systems (ICS) have complex heterogeneous hierarchical-network structures, with a large number of elements, often reconfigurable. Given the large number of possible configurations of computer systems — structures, parameters of the functioning of their elements and the composition of the tasks they solve — a mathematical apparatus of structural and parametric synthesis of hierarchical ICS is necessary.

The presence of such an apparatus will allow on the basis of modeling of ICS to form a rational structure adapted to solve specific computational tasks in various conditions. Therefore, the issues of improving the structure of the ICS come to the fore. The fundamental features of the systems of this class require taking into account the diversity of elements, the variety of their possible states, the heterogeneity of the connections in the system, the conditions of functioning. At the same time, traditional approaches to solving problems of structural-parametric synthesis are often reduced to a consistent choice of system architecture based on a qualitative analysis of existing options, formalization of the selected type of structure and parametric synthesis of its components using known optimization methods. However, sufficiently adequate dynamic models of the functioning of complex ICS, as a rule, are characterized by high computational complexity, which leads to a significant complexity of optimization procedures. Therefore, the solution of the problem of combinatorial optimization by a complete search of possible solutions in practice is unacceptable, which is primarily due to the following circumstances:

– significant number of the set of solutions, depending on the allowable number of elements in the structure and the considered number of their types;

– the labor-intensity of accurate calculation of the objective function (OF) on the basis of simulation and analytical models;

– time spent on obtaining statistical estimates of the quality of the solution;

– the need to repeatedly solve the noted problem in practice when varying the initial data, which, as a rule, is required in applied problems in order to implement a scenario approach that removes uncertainties in the conditions of functioning.

The numerical method of setting the OF and the alleged presence of several local extremes leads to serious difficulties in using well-known mathematical methods and requires special approaches based on a significant limitation of the many solutions under consideration.

A powerful tool for the approximate solution of complex combinatorial optimization problems are the methods of evolutionary search, among which genetic algorithms (GA) stand out. They are ways to solve optimization problems on the basis of evolutionary modeling, based on the use of analogies with natural processes of natural selection [1–4]. The advantages of optimization methods based on GA in comparison with classical ones consist primarily in the possibility of various ways of specifying OF and types of optimization variables, as well as in the use of probabilistic, rather than deterministic rules for finding solutions [5].

The theory of GA is currently quite developed, but significant non-trivial issues in specific cases are, firstly, the construction of the so-called «fitness function» (FF) to assess the quality of solution options and the formalization of optimization variables in the space of possible solutions, which in this case requires taking into account the peculiarities of coding a hierarchical-network heterogeneous structure, and, secondly, the justification for adjusting the parameters of the GA that ensure obtaining satisfactory according to the specified criterion of the decision in a limited time [6–8].

The traditional approach involves adjusting the parameters of the algorithm to obtain a satisfactory result by its repeated implementation with a pre-built FF. In the case of limited time to solve the problem and its significant labor intensity, as noted above, such a path does not seem quite appropriate. Therefore, to solve the problem of stochastic combinatorial optimization,

it is proposed to use a GA with a stochastic FF with a variable (changing in the search process) coefficient of variation, conducting a preliminary selection of rational parameters for adjusting the algorithm.

FORMULATION OF THE PROBLEM OF CHOOSING THE STRUCTURE OF AN INHOMOGENEOUS INFORMATION-COMPUTER SYSTEM

The essence of the task of choosing the structure of the ICS is as follows. There is a nomenclature of electronic component base (ECB), which can be used as elements of ICS. An imitation-analytical model of ICS is given, correlating its structure with the achieved indicator of the target effect, acting as a OF. It is necessary to find the structure of the ICS that provides the highest value of the mathematical expectation of the OF indicator. The time for solving the problem is limited, which is due to the laboriousness of obtaining estimates of the values of the OF by repeatedly implementing simulation modeling of the work of the ICS in the study of many solutions.

With a sufficiently general statement of the problem, we will assume that the:

1. Nomenclature of ECB for building ICS many available types of processor $\mathcal{B}^p = \{\mathcal{b}_i^{pr}\}$ and interface elements (modules) $\mathcal{B}^{if} = \{\mathcal{b}_i^{if}\}, \mathcal{b}_i =< \vec{f}_i^{fu}, \vec{f}_i^{re}, \vec{f}_i^{mg} >, \mathcal{b}_i \in \mathcal{B}, \mathcal{B} = \mathcal{B}^{pr} \cup \mathcal{B}^{if}$, having functional parameters $\vec{f}_i^{fu}$, resource parameters $\vec{f}_i^{re}$, weight and size parameters $\vec{f}_i^{mg}$, defining their particular quality indicators $\vec{d}(\mathcal{b}_i) =< d_j(\vec{f}_i^{fu}, \vec{f}_i^{re}, \vec{f}_i^{mg}) >$. Particular quality indicators are understood, for example, performance, failure rate, power consumption, etc.

2. A model of the functioning of the ICS, which allows to assess the value $\widetilde{\Psi}(S)$ OF under conditions of deterministic and random internal and external factors, where S — hierarchically-the network structure of the ICS, given by the parameters of the elements and the matrix of their adjacency. The OF can be, for example, the productivity of the ICS at a given time interval for a given set of tasks.

3. Vector-function of technical and operational indicators of ICS quality: $W(S_i) =< w_j(S_i) >$.

Restrictions:

1. Set of permissible values of technical and operational indicators of ICS quality $\mathcal{W}^{dir}$.

2. Computational labor intensity of solving the problem, represented in the limit number Υ implementations of the OF $\widetilde{\Psi}(S)$.

Find: On a variety of possible structures $\mathbb{S}$ ICS the structure S^*, providing the maximum expected target effect:

$$S^* = arg \max_{S_i \in \mathbb{S}} M[\widetilde{\Psi}(S_i)], W(S_i) \in \mathcal{W}^{dir}, S_i \in \mathbb{S}.$$

Given the high complexity of solving the problem in practice, it is advisable to build an algorithm for finding its approximate solution. Let's assume that the value of the OF lies in the range $0 \le \widetilde{\Psi} \le 1$, and the value of the OF equal to 1 corresponds to the ICS with the optimal structure.

The essence of GA is stated, for example, in [4, 5]. However, the effectiveness of GA searches ultimately depends on the «tuning» of the research-use ratio determined by the parameters of genetic operators.

ALGORITHM FOR CHOOSING THE STRUCTURE OF HETEROGENEOUS INFORMATION-COMPUTER SYSTEM

The developed method includes as the main stages: filtration of ECB and formalization of possible structures of ICS; adjustment of rational parameters of GA using the deterministic variant of FF; substantiation of the initial coefficient of variation of stochastic FF, taking into account the limit on the maximum number of OF implementations; implementation of the constructed GA. The implementation of this method can be described as a sequence of the following steps.

Step 1. Narrowing of the set of ECB is carried out according to the Pareto principle by allocating subsets $\mathcal{B}^{pr*}$ and $\mathcal{B}^{if*}$, containing respectively options for building computational (CM) and interface modules (IM) with non-improving characteristics.

Step 2. Formalization of the structure $S \to X$ (Fig. 1). Depending on the available element base, the characteristics of subscribers interacting with the ICS, the requirements for interfaces, etc., a structure template is formed that corresponds to the maximum possible composition of elements and a variety of connections within the capabilities of the implemented architecture: the number u of levels, the number m_k, $m_u = 1$ of blocks on each level; each block i-level combines l_i lower nodes: $m_k = l_k m_{k-1}$, herewith $n_{cm} \le m_0 \le n_{cm}^2$, here is n_{cm} — the maximum possible number CM in structure. Structure option S is encoded as a «chromosome»

$$X(S) =< x_j^{(k)}(S) | k = 0, \dots, u, \ j = 1, \dots, m_k >,$$

here is «gene» $x_j^{(k)}(S)$ — processor module implementation number for $k = 0$ or implementing an interface module for $k = 1, \dots, u$; $x_j^{(k)}(S) = 0$ means the absence of an element at a given position of the structure.

Step 3. To build a FF $\widetilde{\Psi}_N(S_i)$ based on OF $\widehat{\Psi}$ the Bayes — Laplace criterion is used, taking into account the penalty function of constraints: $\widetilde{\Psi}_N(S_i) = M[\widehat{\Psi}] \times \varpi(S_i)$, here is $\varpi(S_i)$=0, $\vec{W}(S_i) \notin \mathcal{W}^{dir}$; $\varpi(S_i)$=1, $\vec{W}(S_i) \in \mathcal{W}^{dir}$; $M[\widehat{\Psi}] = \frac{\sum_{k=1}^{N} \Psi_i}{N}$, Ψ_k — implementation of OF in k-th from N experiences.

In order to adjust the rational parameters of GA, a deterministic version of FF is built, which has low computational laborintensity, heuristically or, for example, using estimates of mathematical expectation of parameters of random factors $\hat{Z}$: $\Phi(X_i) = \Psi(S_i | M[\hat{Z}])$. Based on $\Phi(X_i)$ a stochastic variant of FF is constructed, having a normal (Gauss) distribution and coefficient of variation V:

$$\Phi_V(X_i) = \left(1 + V \times \sqrt{-2 \ln \hat{\varrho}_1} \times \cos 2\pi \hat{\varrho}_2\right) \times \Phi(X_i),$$

where $\hat{\varrho}_1, \hat{\varrho}_2 \in [0; 1]$ — uniformly distributed random variables.

Step 4. For some similar structure of the ICS (typical structure, possible construction option, analogues of the desired ICS, etc.), an estimate of the coefficient of variation of the OF is determined:

$$V_0 = V[\widehat{\Psi}] = \frac{1}{\widetilde{\Psi}_N} \sqrt{\frac{\sum_{k=1}^{N} (\Psi_k - \widetilde{\Psi}_N)^2}{N}}, \ N > 10^5.$$

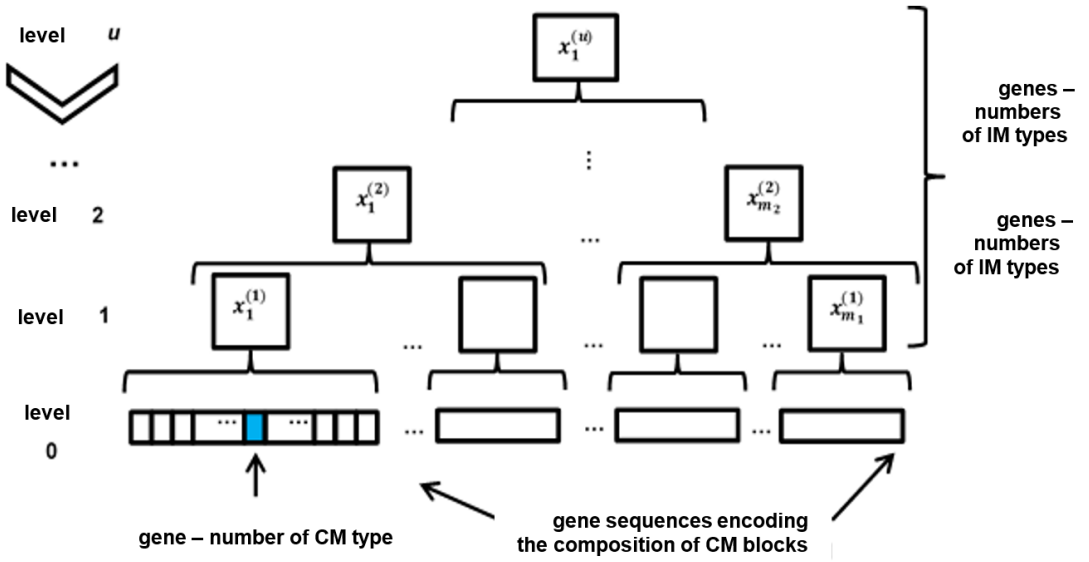

Fig. 1. Scheme of coding variants of the ICS structure

Step 5. Finding an option $X_0 = \arg\max_{X_i} \Phi(X_i)$, providing the maximum value of the deterministic FF, for example by enumeration, which is provided with low computational labor-intensity $\Phi(X_i)$.

Step 6. The initial (for the first generation) coefficient of variation V of the stochastic FF is selected from the range corresponding to $-3 \le \lg V \le -0.5$.

Step 7. GA parameters are selected $\mathcal{G} =< M, K_1, K_2 >$ (discussed below when describing GA operators, recommended ranges M = 0.5, … , 0.05; K_1, K_2 = 0.1, … , 0.9).

Step 8. Population size is selected Ω, $2 \le \Omega \le \Upsilon/2\,(V/V_0)^2$, and the number of generations is determined N_p, corresponding to the limit on the number of calculations of the OF Υ:

$$N_p(V,\Omega)\colon \Upsilon \ge \sum_{i=1}^{N_p} \left(\frac{V_0}{V_i}\right)^2,$$

$$V_i = V - \frac{N_p}{i}(V - V_\Delta),$$

here is V_Δ — a given coefficient of variation of FF for the latest generation, which determines the accuracy of the solution.

Step 9. By series N_γ GA launches with selected parameters $\mathcal{G}$, N_p , Ω indicators are determined

$$\varphi = \frac{\sum_{j=1}^{N_\gamma} \Phi_{V_i}(X_j^*)}{n \times \Phi(X_0)},$$

here is $\Phi_{V_i}(X_j^*)$ — result *j-th* run GA for N_p generations, and

$$P_\gamma = \frac{N\{\Phi_{V_i}(X_j^*)/\Phi(X_0) \ge \gamma\}}{N_\gamma},$$

here is $N\{\Phi_{V_i}(X_j^*)/\Phi(X_0) \ge \gamma\}$ — number of «successful» launches, when $\Phi_{V_i}(X_j^*)/\Phi(X_0) \ge \gamma$. Meaning φ consists in the average relative proximity of the resulting solution to the optimal value of FF, a P_γ it makes sense to have the probability of «success» of the launch of the GA, when $\varphi \ge \gamma$, where is γ — the value of the acceptable deviation set in advance. Statistically expedient number of launches of GA to assess its effectiveness based on the variance of the binomial distribution and the «three sigma» rule we will count $N_\gamma \approx \frac{9P_\gamma(1-P_\gamma)}{(1-\alpha)^2}$, α — confidence probability. So, for example, $N_\gamma = 360 \ldots 900$ for $\alpha = 0.05$, $P_\gamma = 0.5, \ldots, 0.9$. From these considerations, we obtain an indicator of the relative effectiveness of the GA with the selected parameters $\psi = \varphi \times P_\gamma$.

Step 10. Repeating steps 8–9 and varying population size Ω, get the dependence of relative effectiveness ψ from Ω at fixed V and $\mathcal{G}$, which is easily interpolated, which allows you to find $\Omega^* = \arg\max_{\Omega} \psi(V, \mathcal{G}, \Omega)$ and get $\psi^*(V, \mathcal{G}) = \psi(V, \mathcal{G}, \Omega^*)$.

Step 11. By repeating steps 7–10 and varying the parameters $\mathcal{G}$ heuristically the selection is carried out

$$\mathcal{G}^* = \arg\max_{\mathcal{G}} \psi^*(V, \mathcal{G})$$

and receiving $\psi^{**}(V) = \psi^*(V, \mathcal{G}^*)$.

Step 12. Repeat steps 6–11 for different V and similar to the step10 is being sought $V_{opt} = \arg\max_{V} \psi^{**}(V)$.

Step 13. Define

$\mathcal{G}_{opt} = \arg\max_{\mathcal{G}} \psi^*(V_{opt}, \mathcal{G})$,

$\Omega_{opt} = \arg\max_{\Omega} \psi(V_{opt}, \mathcal{G}_{opt}, \Omega)$.

Step 14. In the case of a low assessment of the performance of the GA, what should be considered $\psi(V_{opt}, \mathcal{G}_{opt}, \Omega_{opt}) < \gamma$, it is necessary, returning to step 1, to reduce the numbers of the sets $\mathcal{B}^{pr}, \mathcal{B}^{if}$ due to more strict filtering of the element base and (or), returning to step 2, facilitate structure parameters $u, m_k, k = 1, \ldots u$ by reducing the permissible complexity and diversity of structures.

Step 15. Launch of GA with FF for *i*-th generation

$$\tilde{\Psi}_{N_i},\ N_i = \frac{V_0^2}{V_i^2},\ V_i = V_{opt} - \frac{N_p}{i}(V_{opt} - V_\Delta)$$

and parameters: mutations and crossing overs respectively $< M, K_1, K_2 >= \mathcal{G}_{opt}$, population volume Ω_{opt}, number of generations $N_p(V_{opt}, \mathcal{G}_{opt}, \Omega_{opt})$, which leads to the desired solution $X^* \to S^*$.

FEATURES OF THE FORMATION OF GENETIC OPERATORS IN THE OPTIMIZATION OF THE STRUCTURE OF INFORMATION-COMPUTER SYSTEM

Let denote through F some FF and describe the main operators of GA, justified as a result of the research.

Mutation operator $X' = \mathrm{Mut}(X, M)$ assumes for mutating with probability M gene equally-likeable choice of alleles from the set of possible: with probability M

$$x_j'^{(0)} = [\hat{\varrho} \times \mathrm{card}(\mathcal{B}^{pr*}) + 1)],\ j = 1, \dots, m_0;$$

$$x_j'^{(k)} = [\hat{\varrho} \times \mathrm{card}(\mathcal{B}^{if*}) + 1)], j = 1, \dots, m_k, k = 1, \dots, u.$$

Here and below, the square brackets mean rounding to the integer at the bottom.

Selection operator assumes selection for the next generation (range Sel) chromosomes with a FF value not lower than the average in the generation:

$$Sel = \{i: F(X_i) \geq \bar{F}, i = 1 \dots \Omega\} =$$

$$= \{sel_k | k = 1, \dots, \mathrm{card}(Sel)\},\ \bar{F} = \frac{1}{\Omega}\sum_{i=1}^{\Omega} F(X_i);$$

$$\overline{Sel} = \{i: i \notin Sel, i = 1, \dots, \Omega\} = \{\overline{sel}_k | k = 1, \dots, \Omega - \mathrm{card}(Sel)\},$$

here is Ω — number of chromosomes in generation.

Crossing over operator $X' = \mathrm{Kross}(X, Y, K_1, K_2)$ involves the exchange of genes on the epymous positions of the parent chromosomes X, Y with probability K: if $F(X) \geq F(Y)$, $\hat{\varrho} > K$; or $F(X) \leq F(Y)$, $\hat{\varrho} < K$: $x_j'^{(k)} = y_j^{(k)}$, else $x_j'^{(k)} = x_j^{(k)}$, where is $\hat{\varrho} \in [0; 1]$ — played uniformly distributed random variable, $K = K_1 + (K_2 - K_1) \times {}^{k}/_{u}$, K_1, K_2 — crossing over parameters for 0-th и (u-1)-th hierarchical level respectively, $i, j = 1, \dots, m_k$, $k = 0, \dots, u$. This construction of the crossing over takes into account the expediency of different intensity of exchange of structural units of different levels of hierarchy between parent variants that shown by researches.

Implementation of the GA includes the following steps:

1. Formation of a random initial population

$$X(i) = \mathrm{Mut}(X(i), 1), i = 1, \dots, \Omega.$$

2. Equally probable chromosome selection from the entire generation $i_k = [\Omega\hat{\varrho}] + 1$, equally likely choice of chromosome from among those selected for the next generation:

$$j_k = sel_{[card(Sel)\cdot\hat{\varrho}]+1},$$

and the use of a crossing over operator for them

$$X(k) = \mathrm{Kross}(X(i_k), X(j_k), K_1, K_2), k \in \overline{Sel}.$$

This method has the advantage that such an operator does not change the total number of chromosomes in the generation, forming a new one instead of the one being removed, and reduces the likelihood of losing successful search directions during selection, leaving the possibility of participation of «rejected» chromosomes in the formation of the next generation.

3. Application of the mutation operator for all chromosomes in the generation except the best FF value:

$$X(i) = \mathrm{Mut}(X(i), M), i = 1, \dots \Omega, i \neq \arg\max_{i=1,\dots\Omega} F[X(i)],$$

that ensures that the best possible solution is preserved.

4. Iterative execution of steps 2–3 and formation of the result $X^* = \arg\max_{i=1,\dots\Omega} F[X(i)]$ while achieve the number of generations N_p.

EXAMPLE OF SEARCH THE STRUCTURE OF INFORMATION-COMPUTER SYSTEM

Let us give an example of the application of the developed approach [9, 10] to the choice of the rational structure of a multi-module computing system. For this purpose, an appropriate software package was developed and used [11]. At the same time, the calculation time of a single implementation of the OF was about 300 ms for 100 integration steps (a PC based on Intel i5 with a clock frequency of 3 GHz in the MATLAB 7), the required standard deviation of the OF score of 1 %, which corresponds to 10^4 experiments (the initial coefficient of variation of the OF is close to 1). With the total allowable number of elements of the three-level structure $N = 22$, the estimate of the total number of its formalized variants of the structure was $1{,}9 \times 10^9$, taking into account the resource constraints on the composition of the elements - about 4.3×10^4.

On Figure 2 shows the experimental dependence of FF on the number of generations and the coefficient of variation with rational parameters justified by the above $KV = 0.95$; $KN = 0.8$; $M = 0.3$; $\Omega = 10$.

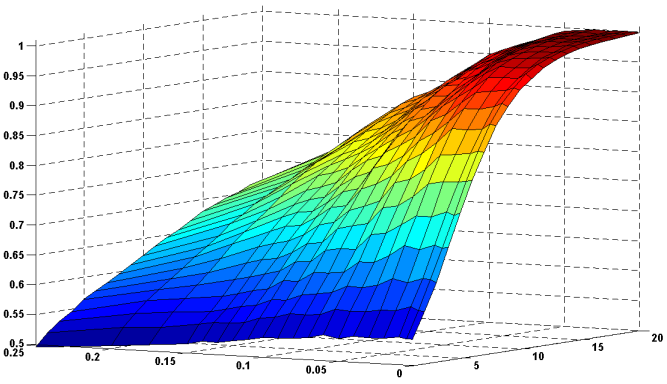

Fig. 2. Example of dependence of the objective function on the number of generations of GA and the coefficient of variation of FF

On Figure 3 shows the experimental dependence of the performance indicator ψ on the number of OF calculations for various methods of constructing the GA (heuristic «directional-random» setting corresponds to $KV = 0.9$; $KN = 0.6$; $M = 0.1$; $\Omega = 12$; «random-directional» — $KV = 0.6$; $KN = 0.5$; $M = 0.4$; $\Omega = 4$).

For example, a search time limit of 36 hours was set, which corresponds to the enumeration of 48 variants of structures in the traditional way. In this case, the use of the developed method made it possible to find a quasi-optimal solution that provides a value of the OF estimate of about 0.99. At the same time, for the traditional method of using GA (constant coefficient of variation of FF) with rational parameters, this value was 0.56, i. e. the increment of the OF was 77 %. On the other hand, fixing the required value of the OF assessment, it is possible to estimate the gain by the time of its achievement: about 8 times for the value of 0.95, 12 times — for 0.9.

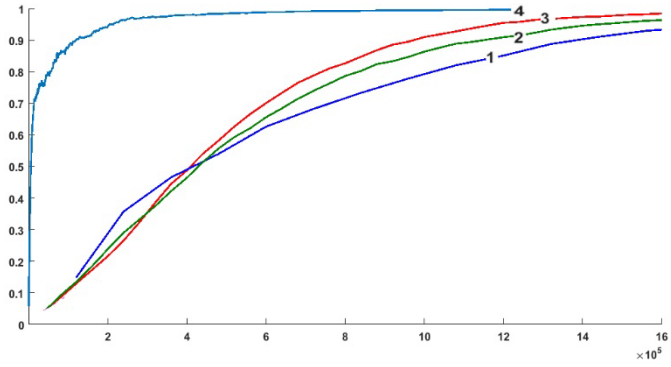

Fig. 3. Comparative analysis of the effectiveness of GA in the search for the optimal structure of the ICS
1 — heuristic «directional-random» setting;
2 — heuristic «random-directional» setting;
3 — setting with a constant coefficient of variation of FF;
4 — setting with variable coefficient of variation FF

In practice, the choice of the option of constructing a complex technical object is, as a rule, multi-criteria, and the final decision is made as an optimal compromise from some subset of possible solutions. The fact is that, firstly, a significant impact on the decision is exerted by difficult to formalize factors associated, for example, with the labor-intensity of the implementation of a particular option with existing developments, etc. Secondly, it is necessary to eliminate uncertainties in the formation of initial data, for example, sets of computational tasks. Therefore, the application of the developed method is most expedient in cases where it is required in a limited time to form a set of quasi-optimal solutions in the scope of possible structures, thereby providing a reasonable optimal-compromise choice of the desired structure.

Conclusion

A method of heuristic solution of a complex combinatorial optimization problem of large dimension with a stochastic target function through the use of an evolutionary modeling apparatus – genetic algorithms is proposed. The problem of structural-parametric optimization of complex heterogeneous hierarchical-network ICS fully belongs to this class. This is due, on the one hand, to the many possible options for their construction, taking into account the variety of types of elements and their places. in the structure, and on the other hand, the complexity of the models of functioning and the corresponding complexity of their computational implementation. Therefore, despite the capabilities of modern computer technology, the proposed method should be used to develop proposals for substantiating the technical appearance of complex ICS. Statistical research has shown the possibility of significantly reducing the time spent on finding rational options for building ICS.

References

1. Holland J. H. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence. Cambridge (MA), MIT Press, 1992, 228 p.

2. Balashov E. P. Evolyutsionnyy sintez system [Evolutionary synthesis of systems]. Moscow, Radio i Svyaz Publishing House, 1985, 328 p. (In Russian)

3. Burakov M. V. Geneticheskiy algoritm: teoriya i praktika: Uchebnoe posobie [Genetic algorithm: theory and practice: Study guide]. Saint Petersburg, Saint Petersburg State University of Aerospace Instrumentation, 2008, 164 p. (In Russian)

4. Goldberg D. E. Genetic Algorithms in Search, Optimization and Machine Learning. Thirteenth Edition. New York, Addison-Wesley Professional, 1989, 432 p.

5. Kureychik V. V., Kureychik V. M., Rodzin S. I. Teoriya evolyutsionnykh vychisleniy: Monografiya [Theory of evolutionary computing: Monograph]. Moscow, Fizmatlit Publishing House, 2012, 260 p. (In Russian)

6. Rodzin S. I. O problemnykh voprosakh teorii geneticheskikh algoritmov [On problematic issues of the theory of genetic algorithms], *Izvestiya Taganrogskogo gosudarstvennogo radiotekhnicheskogo universiteta [Izvestiya TSURE]*, 2006, No. 8 (63), Pp. 51–56. (In Russian)

7. Gladkov L. A., Kureychik V. V., Kureychik V. M. Geneticheskie algoritmy: Uchebnoe posobie [Genetic algorithms: Study guide]. Moscow, Fizmatlit Publishing House, 2006, 320 p. (in Russian)

8. Jones T., Forrest S. Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms. In: *Eshelman L. J. (eds.) Proceedings of the 6th International Conference on Genetic Algorithms (ICGA 1995), Pittsburgh, PA, USA, July 15–19, 1995*. San Francisco (CA), Morgan Kaufmann, 1995, Pp. 184–192.

9. Zakharov I. V. Ratsionalnyy vybor struktur i konfiguratsiy neodnorodnykh vychislitelnykh sistem pri pomoshchi evolyutsionnogo poiska [The Rational Choice of Structures and Configurations of the Heterogeneous Computer Systems Through the Evolutionary Search], *Vestnik Rossiyskogo novogo universiteta. Seriya «Slozhnye sistemy: modeli, analiz i upravlenie» [Vestnik of Russian New University. Series «Complex Systems: Models, Analysis, Management»]*, 2018, Is. 1, Pp. 85–90. DOI: 10.25586/RNU.V9187.18.04.P.85.

10. Zakharov I. V., Zabuzov V. S., Sokolovsky A. N. Programmnyy kompleks dlya ratsionalnogo vybora neodnorodnoy ierarkhicheskoy struktury vychislitelnoy sistemy i ee konfiguratsiy na osnove geneticheskikh algoritmov [Software Package for Rational Choice of Heterogeneous Hierarchical Structure of the Computer System and Its Configurations Based on Genetic Algorithms]. Certificate of State registration of a computer program RU No. 2018614368, published at April 04, 2018, 1 p. (In Russian)

11. Zakharov I. V., Terekhov V. G. Primenenie geneticheskikh algoritmov k zadacham postroeniya bortovykh vychislitelnykh sistem i upravleniya ikh funktsionirovaniem [Application of Genetic Algorithms for Tasks of Designing and Management of On-Board Computer Systems, *Estestvennye i tekhnicheskie nauki [Estestvennye i tekhnicheskie nauki]*, 2017, No. 8 (110), Pp. 92–95 (In Russian)

DOI: 10.24412/2413-2527-2022-331-46-51

Выбор структур неоднородных информационно-вычислительных систем на основе аппарата генетических алгоритмов

к.т.н. И. В. Захаров, к.т.н. А. О. Шушаков, С. С. Зыкова
Военно-космическая академия имени А. Ф. Можайского
Санкт-Петербург, Россия
x.vano-z80@yandex.ru, shushakovaleksei@mail.ru

Аннотация. **Достаточно адекватные динамические модели функционирования сложных систем характеризуются высокой вычислительной сложностью, что ведет к существенной трудоемкости оптимизационных процедур. Поэтому решение задачи комбинаторной оптимизации путем полного перебора возможных решений на практике оказывается неприемлемым.**

Преимуществами эволюционного поиска как метода комбинаторной оптимизации структуры информационно-вычислительной системы является наличие возможности различных способов задания целевой функции и типов переменных оптимизации, а также в использовании вероятностных, а не детерминированных правил поиска решений. Предложен способ формализации структуры гетерогенной информационно-вычислительной системы, который учитывает ее иерархически-сетевую структуру.

Представленный подход позволяет посредством выбора рациональных параметров генетического алгоритма и использования стохастической фитнесс-функции с переменным коэффициентом вариации достигать удовлетворительной скорости его сходимости при большой размерности задачи. Приведены примеры приложения указанного метода к поиску рациональной структуры вычислительной системы.

Ключевые слова: **вычислительная система, эволюционный поиск, генетический алгоритм, стохастическая фитнесс-функция.**

Литература

1. Holland, J. H. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence.— Cambridge (MA): MIT Press, 1992. — 228 p.

2. Балашов, Е. П. Эволюционный синтез систем. — Москва: Радио и связь, 1985. — 328 с.

3. Бураков, М. В. Генетический алгоритм: теория и практика: Учебное пособие. — Санкт-Петербург: ГУАП, 2008. — 164 с.

4. Goldberg, D. E. Genetic Algorithms in Search, Optimization and Machine Learning. Thirteenth Edition. — New York: Addison-Wesley Professional, 1989. — 432 p.

5. Курейчик, В. В. Теория эволюционных вычислений: Монография / В. В. Курейчик, В. М. Курейчик, С. И. Родзин. — Москва: Физматлит, 2012. — 260 с.

6. Родзин, С. И. О проблемных вопросах теории генетических алгоритмов // Известия Таганрогского государственного радиотехнического университета. 2006. № 8 (63). С. 51–56.

7. Гладков, Л. А. Генетические алгоритмы: Учебное пособие / Л. А. Гладков, В. В. Курейчик, В. М. Курейчик ; под ред. В. М. Курейчика. — Изд. 2-е, испр. и доп. — Москва: Физматлит, 2006. — 320 с.

8. Jones, T. Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms / T. Jones, S. Forrest // Proceedings of the 6th International Conference on Genetic Algorithms (ICGA 1995), (Pittsburgh, PA, USA, 15–19 July 1995) / L. J. Eshelman (eds.). — San Francisco (CA): Morgan Kaufmann, 1995. — Pp. 184–192.

9. Захаров, И. В. Рациональный выбор структур и конфигураций неоднородных вычислительных систем при помощи эволюционного поиска // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2018. Вып. 1. С. 85–90.
DOI: 10.25586/RNU.V9187.18.04.P.85.

10. Свидетельство о государственной регистрации программы для ЭВМ № 2018614368 Российская Федерация. Программный комплекс для рационального выбора неоднородной иерархической структуры вычислительной системы и ее конфигураций на основе генетических алгоритмов: № 2018610087: заявл. 09.01.2018: опубл. 04.04.2018 / Захаров И. В., Забузов В. С., Соколовский А. Н.; заявитель Захаров И. В. — 1 с.

11. Захаров, И. В. Применение генетических алгоритмов к задачам построения бортовых вычислительных систем и управления их функционированием / И. В. Захаров, В. Г. Терехов // Естественные и технические науки. 2017. № 8 (110). С. 92–95.

DOI: 10.24412/2413-2527-2022-331-52-60

Water Transport Information Security Trainer Concept

M. V. Kardakova, Grand PhD A. P. Nyrkov, Yu. V. Tsymay
Admiral Makarov State University of Maritime and Inland Shipping
Saint Petersburg, Russia
m.v.kardakova@ya.ru, apnyrkow@mail.ru, m-walua@yandex.ru

Abstract. **The article discusses the possibility of using game simulators in training personnel or students. The relevance of the use of simulators in teaching information security students in water transport is considered. The relevance of such simulators is emphasized by the possibility of their use in distance learning. The use of a recognizable genre, which is familiar and popular among students in the virtual world, allows you to acquire positive emotions, increase interest and motivation for the development of cognitive activity. The game mechanics used in the simulator are described. An example of calculating the risk of an emergency on a ship is considered, options for using the calculations in the simulator are proposed. The concept of the simulator itself is described and the concept of the interface is developed. The use of simulators in teaching allows you to teach students the practical skills of working with something. As a rule, such an approach to teaching requires the student to know the theory of the subject, as well as to actively participate in the work. Thus, the student has the opportunity to demonstrate his knowledge in practice and improve his skills. Of course, no simulator will replace a real device or a real situation, but such an approach in teaching will help to consolidate the material and in the future the student, seeing a real object, will not be lost and will be able to calmly work with it. The use of information security simulators for educational purposes ensures the consolidation of theoretical knowledge and their application in practice. It is very important that the training simulators, on the one hand, be close to real conditions, on the other hand, work on them was interesting and convenient for the student.**

Keywords: **transport, water transport, vessel, ship, training apparatus, the game, education, risk.**

INTRODUCTION

The use of information security simulators for educational purposes ensures the consolidation of theoretical knowledge and their application in practice. It is very important that the training simulators, on the one hand, be close to real conditions, on the other hand, work on them was interesting and convenient for the student [1].

Simulators created in the form of a game allow students to master new competencies with interest. Such technologies can become an effective tool in the educational process for teaching students new things or consolidating the material already passed [2]. Also, the passage of such a game can be set as homework, for the student's independent work. The relevance of such simulators is also emphasized by the possibility of using them in distance learning [3].

The use of a recognizable genre, which is familiar and popular among students in the virtual world, allows you to acquire positive emotions, increase interest and motivation for the development of cognitive activity [4].

MATERIALS AND METHODS

The use of simulators in teaching allows you to teach students the practical skills of working with something. As a rule, such an approach to teaching requires the student to know the theory of the subject, as well as to actively participate in the work. Thus, the student has the opportunity to demonstrate his knowledge in practice and improve his skills. Of course, no simulator will replace a real device or a real situation, but such an approach in teaching will help to consolidate the material and in the future the student, seeing a real object, will not be lost and will be able to calmly work with it.

For example, «Virtual complex «Protection of an object from leaks of information through technical channels» TZI-VIRT», the simulator is aimed at obtaining primary experience in preparing a room for certification, consolidating in practice knowledge on the physics of the formation of channels of leakage of technical information, familiarization with modern protective devices and methods of their installation, depending on the characteristics of the virtual office. This software package helps in the study of methods of technical protection of information. The program allows you to design premises, simulate threats and arrange protection against them.

Also interesting as an example is the software package «Virtual complex «Detection of embedded devices and hidden video cameras» RAM-VIDEO-VIRT». The simulator is designed to gain knowledge about methods for detecting embedded devices and hidden video cameras, as well as to gain experience with modern security tools designed to detect embedded devices and hidden video cameras.

An interesting option is to use the game as a simulator, when a student memorizes new material while playing and masters new competencies.

The use of game techniques in the lessons allows the student to acquire the following skills [5]:

- make quick decisions and be responsible for them;
- use the experience gained in performing professional tasks, determine the best solutions, evaluate their quality and effectiveness;
- communicate with each other and foster teamwork;
- take responsibility for the work of all team members, for the result of assignments.

To date, many interesting educational games have been created that help to consolidate this or that material, but they are not structured into a single course on information security and are suitable only as an auxiliary environment for studying certain topics.

Using game technologies in pedagogy, it is necessary to go through several main stages, which include drawing up and planning goals, defining tasks and plans, and, finally, complet-

ing the assigned tasks. It is necessary to discuss and analyze with students the whole process of work and the results obtained.

Any activity makes sense when these conditions are as close as possible to real life. The student must have a voice, discretion and responsibility. Only if these requirements are met, a person fully defines himself as a specialist.

The classification of educational gaming technologies is quite large. So gaming educational technologies are usually classified by the field of activity, by the nature of the educational process, by the method of the game, by the subject area, by the game environment, etc. Within the framework of this article, an intellectual, educational, simulation computer game on information security in water transport will be considered.

The classification of computer games is also varied, but we will pay special attention to adventure games, strategy games, puzzles, computer simulations and educational games. All these genres can be used when teaching students and staff any new competencies.

THE CONCEPT OF THE SIMULATOR

The purpose of the simulator is to create a learning environment where cognitive and training activities will be available in a playful manner that is interesting for the student. The simulator should consist of a plausible small boat model. The student will have to be able to choose protection equipment and install it on the object. There should also be access to the ship's computer interface for «installation» and operation of the software [6, 7]. You can also add the passage of mini-games to the main program for learning, mastering or repeating certain competencies. These mini games can be used individually or as a quest.

For the correct placement of protection equipment and installation of software, the student will need to have knowledge of the legal framework, as well as to calculate the risks [8, 9].

The game will contain the following objects:

- components of the security system;
- ship cabins;
- captain's bridge;
- equipment of the vessel.

This plot has the following attractive features:

- the theme of a real ship is used, but students dream of getting a real experience;
- the variety of ship equipment, work situations and work responsibilities of an information security specialist creates a lot of potential for the development of the plot;
- you can add secret levels for those who want to further study the material on the game.

Table 1 below presents the interpretation of the elements of the mechanics in the context of the game plot and setting.

Table 1

Interpretation of elements of mechanics

Element of mechanics	
Points for solving problems	Points in individual player ratings and points in student ratings on the course. The ratings are divided according to the competencies that are studied at one level or another.
Points for finding items	Points in individual ratings and points in student ratings by competency.
Points for the correct use of items	Points in individual ratings and points in student ratings by competency.
Ranking among all students	Ratings between students and the rating of the player himself.
Levels	Divided into competencies, according to complexity
Helping the student to complete the assignment if the assignment takes too long	The task can be completed no longer than two hours or lesson, depending on the level. If the player completes the task for a long time and cannot solve the problem, but also does not take hints, the gameplay prompts him to the correct way out of the situation.
Limiting lesson time	When the time ends and if the task was not done, then the student receives a «reprimand» card, if the student accumulates a certain number of such cards, then he will need to go through the game again. If the task was done, but with errors, then the student is invited to redo the work. If everything is done correctly, «the manager gives a new task to the employees».
Penultimate level reached	It is proposed to pass the last level, which will be a practice-oriented examination task. In this task, the student will need to use all the knowledge gained earlier. This level should correspond to all competences of the discipline.
Reached the last level	The student receives a full report on what competencies he has learned and where he made mistakes. It is also possible to view such reports after passing each level.
Performing three tasks at once quickly and without errors and without asking for hints	The player receives a bonus item that gives additional opportunities when solving tasks.
Completed the level on your own without asking for hints	The player receives a part of the «magic» item, if the player collects the «magic» item, he can exchange it for points in the rating or for additional opportunities.
All levels are completed independently, without prompting for tips and correctly	Additional points are awarded in the rating.
The player did not complete the levels at the end of the game time	The player is asked to play the game first before completing the additional session. 3 additional levels are also added
The minimum number of points or below the minimum at the time of the end of the game	The player is invited to go through three additional levels.
Not a single task was completed correctly	The player is invited to go through some levels again, and it is also proposed to go through three additional levels.
All tasks are completed with tips	The player is invited to go through three additional levels, without the possibility of using hints.

Thus, the following game mechanics will be used in the work [2]:

1. Mechanic «Achievement». This mechanic is based on a material or virtual expression of the result of performing an action. The results can be seen on their own or as a reward. Anything can be a reward.

2. Avoidance mechanic. With this mechanic, the player is motivated not by a reward, but by the fact that he can avoid punishment. This usually helps to keep the activity level consistently according to the schedule provided by the developer.

3. The Reward for Effort mechanic. The idea behind the mechanics is that when you play, you feel more joy from work than from rest.

4. Mechanics «Chain of events». In this mechanic, the reward is used as a link in a chain of related events. In many cases, players see these events as separate elements. When a link in a chain is unlocked, the player perceives this as a reward for the actions taken.

5. Mechanics «Countdown». Applying the presented mechanics, it is necessary to create situations when a limited time is allotted for solving problems or overcoming obstacles. Using this method, you can increase the activity of the players and increase its performance compared to the original. However, it is important to understand that activity only increases for a limited period of time.

6. Mechanics «Combined rating of winners». Mechanics are used when there is a need to use a common scoring system for a number of game scenarios, and they can be completely dissimilar and unrelated to each other.

7. Mechanic «Reward for a specific sequence of actions». Players are rewarded not for one, but for several consecutive actions. At first, the technique can reduce the activity of the participants, since the action does not bring a reward, but later it increases, because the time for receiving the reward is approaching.

8. Loyalty mechanic. The mechanic of loyalty is to establish a non-verbal relationship between the participant in the game and the reality of the game. This connection is achieved by introducing a person to his participation in the game world (for example, he may have real estate in the game), and then this is enhanced by visual images that other players can see.

9. Meta-game mechanics. The meaning of this mechanic is that an additional one is built into the main game. During the game, the participant can find an additional game, because the author does not advertise its presence, so as not to create confusion or increase interest in the game. But the developer's advantage is that when players find meta games, they are usually very happy because it creates a pleasant surprise effect.

10. Mechanics «Micro-competition». Used when working with mini-games. It allows you to form mini-scores and is most suitable for games with several game mechanisms, as well as for situations with a large number of mini-competitions. It is easy to increase the loyalty of game participants with the help of a varied system of rewarding winners in mini-games.

11. Mechanic «Modifiers». In the mechanics under consideration, a certain object or artifact is used during the game, the use of which affects the results of tasks. This is called a modifier. Often the player earns it by completing a series of tasks or performing a series of important actions.

12. Mechanic «Pride». Based on positive feelings from the achieved result.

13. User progress mechanic. It is a mechanism that reflects the progress of the player in completing the tasks assigned to him by the game.

14. Instant reward mechanic. The player receives all information about his position in real time. As a result, he instantly reacts emotionally.

15. Real Prize Distribution mechanics. When using this mechanic, the player receives a reward that is of real, tangible value to the player for the results of the game. Each player can receive such a reward if they meet certain requirements or receive exceptional results.

Plot arc: the protagonist is a specialist in an information security services firm. He is tasked with improving the security system on the ship. Passing each level, he must solve information security problems close to real conditions. It is necessary that after passing the game the student masters certain competencies.

It is also interesting that the teacher can change the sequence of levels; the teacher can even remove some levels if he considers them not suitable in his course. To do this, the teacher will need to go to the «level constructor».

Level example

For example, consider the level of definition of the risk of an emergency, which is planned to be implemented in the near future.

At the initial stage of the game, students can be asked to calculate the possibility of an emergency on a given vessel. You can also give students the opportunity to choose a level: «Easy», «Intermediate», «Difficult». The selected level will determine the maximum number of points that a student can score when passing the level.

When choosing any level of difficulty, the student will be asked to take a training test, and if the test is successfully completed, the task will gradually open. When choosing an «easy» level, the student will need to choose from several proposed forms with the correct formulas and substitute parameters there based on the route of the vessel, the cargo being transported, the type of vessel, etc. When choosing a level of «medium» complexity, the student will need to draw up a form with formulas himself, but he will have the opportunity to refer to a short synopsis of theoretical information. And when choosing the «Difficult» level, the student will not have the opportunity to refer to theoretical materials. At the «intermediate» and «difficult» level, it is worth adding a «check» button, when the wrong decisions of the student will be highlighted during pressing, but it is worth limiting the number of clicks on this button.

The risk assessment of emergencies in this example is determined using statistics, route, etc. The model, which will describe the state of the ship's safety using risk assessment, is based on the statistics of emergencies in water transport. Statistics allow you to determine the type of flow of extreme events, as well as the trend equation [10–12].

To solve this problem, let us consider a statistical model of safety in water transport, taking into account the limited time interval τ. Based on Poisson's law, consider the probability that k events will occur:

$$P\{X(t,\tau)=k\}=\frac{a^k e^{-a}}{k!},$$

where P is the probability of an event occurring;
$X(t, \tau)$ is a function of the number of random hazardous events;
k is the number of random hazardous events during the time τ;
a is a parameter depending on τ and k; $a = \alpha\tau$, where λ is the intensity of the flow of dangerous events.

We can consider the general flow of random hazardous events in water transport as private flows of random events connected together for a number of causal factors. The sum of independent Poisson flows is also a Poisson flow [13, 14].

It was also taken into account that the occurrence of emergency situations in water transport can be defined as a sequence of incompatible events A_j^l and joint events B_i. Incompatible events A_j^l may include technogenic threats of the type (j = 1, 2, ..., J) on in the navigation area l (l = 1, 2, ..., L), for example, failure or breakdown of navigation systems, such a threat could be the cause of an emergency such as grounding, collision with another vessel or coastline, deviation from the route, flooding, etc. It should be noted that an emergency on a ship can be the result of only one of the many considered joint events A_j^l [15, 16]. The type of consequences i can be different, for example, flooding, grounding, fire, crash, collision.

Thus, if we take the average, from the ratio of the number of ships that have suffered corresponding accidents due to the onset of a man-made threat of the j-th type with damage i in the considered navigation area l to the total number of ships, we can determine the probability of a man-made threat of type j causing damage i.

$$P\left(A_{ji}^l\right)=\frac{\sum_j\sum_i\sum_l N_{jil}}{\sum_l N_l}.$$

The weighted estimates of failures can be calculated using the formula:

$$\omega\left(A_{ji}^l\right)=\frac{\sum_l N_{jil}}{\sum_j\sum_l\sum_i N_{jil}}.$$

It should be borne in mind that the causes of an accident taken from statistical data can be subjective, since decisions about the specific cause of the emergency are based on expert judgment. An expert opinion is accepted during an official investigation of an accident in water transport [17, 18]. Therefore, we can consider the «subjective» a priori probabilities of the event $P(A_j^l)$. Thus, events A_j^l and events B_i will be considered as hypotheses. Based on Bayes' theorem, it is possible to determine the probability of the «hypotheses» that caused the event B_i [19–21].

$$P\left(A_j^l|B_i\right)=\frac{P\left(A_j^l\right)\times P\left(B_i|A_j^l\right)}{\sum_{j=1}^{J}\sum_{t=1}^{L}P\left(A_j^l\right)\times P\left(B_i|A_j^l\right)},$$

where $P(A_j^l)$ are probabilistic hypotheses A_j^l;
$P\left(B_i|A_j^l\right)$ — conditional probabilities of the event B_i under the hypothesis $P(A_j^l)$.

To predict and calculate the intensity of emergency situations due to natural disasters, unpredictable events, unfavorable weather conditions in the time interval T for region l, statistical data are required [22]. Thus, the intensity of emergency situations can be calculated by the formula:

$$\gamma_{\left(B_i|A_{jl}\right)}=\frac{\sum_j\sum_l\sum_i N_{jil}}{T_l}.$$

If we consider one ship whose route passes through region l:

$$\gamma^{\sum S}_{\left(B_i|A_{jl}\right)}=\frac{\sum_j\sum_l\sum_i N_{jil}}{T_l\times\sum_l S_{lT}},$$

where $\sum_l S_{lT}$ is the number of ships passing through region l during time T.

Thus, we will calculate the level of emergency risk:

$$R\left(B_i|A_j^l\right)=1-\exp\left(-\gamma_{B_i|A_j^l}T_l\right);$$

$$R(B_i)=\sum_{j=1}^{J}P\left(A_j^l\right)\times R\left(B_i|A_j^l\right);$$

$$R(B)=\sum_{i=1}^{3}\sum_{j=1}^{J}P\left(A_j^l\right)\times R\left(B_i|A_j^l\right).$$

The cost of the risk of an emergency R_i^l can be calculated by determining the amount of possible damage $\int W_i^l dl$.

$$R_i^l=P\left(B_i|A_j^l\right)\int W_i^l dl.$$

The assessment of the risks of emergencies in water transport along the given routes can be presented as a sum of risks [23, 24]:

$$R_\omega=R_1+R_2+R_3,$$

where R_1 is the risk of a natural disaster (natural threat);
R_2 is the risk of realization of man-made threats;
R_3 is the risk of realization of anthropogenic threats.

All of the above can be used as theoretical information for calculations. You can also consider several ways to calculate risk.

Interface concept

After loading the game, the user must have access to the menu: continue the game, select the game, select the level of the game, select the mini-game, view reports, ratings, settings, view controls, information about the game. Also, the user should be able to exit the game and return to the main menu and exit the main menu.

The game screen should display the time remaining until the completion of the level, place in the rating, level number, location (depending on the task), equipment (depending on the task).

Consider the use cases in the game menu presented in Figure 1.

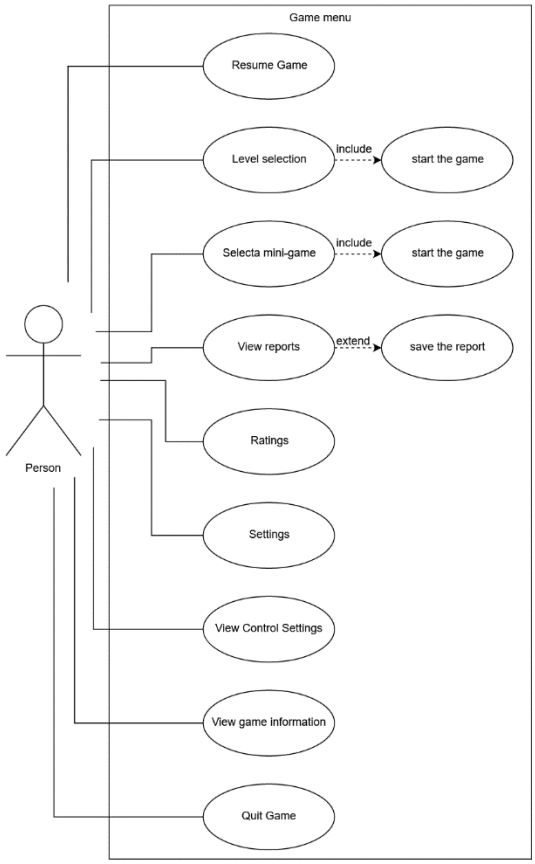

Fig. 1. Diagram of use cases in the game menu

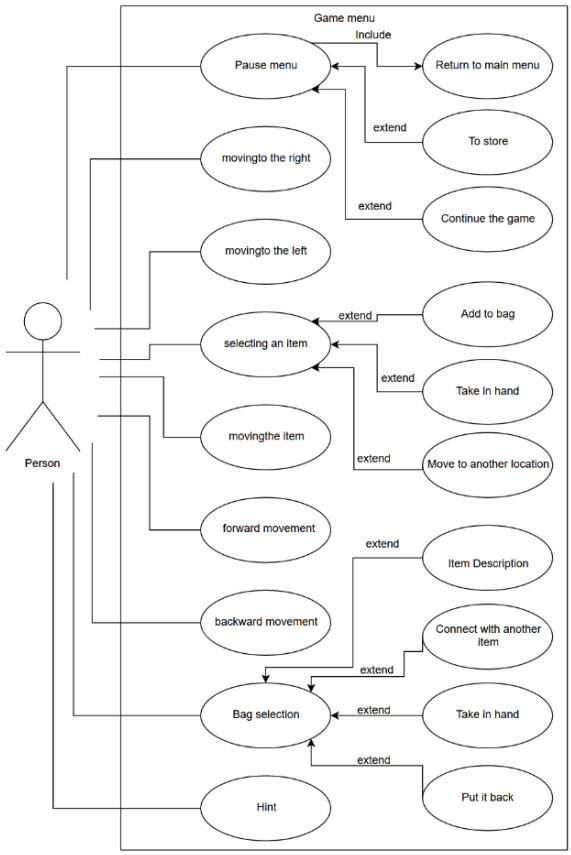

Fig. 2. Diagram of use cases in a game scene

Resume Game: The user can return to a saved game.

Level selection: the user can switch to a map with game levels, then he must select a level and start the game.

Select a mini-game: the user can switch to a map with mini-games, after which he must select a mini-game and start the game.

View reports: the user can view reports on the game, see what mistakes he made, which competencies he mastered better and which ones were worse. It is possible to save the report.

Ratings: User can see their rating in a group or among all students on the course.

Settings: the user can enter the settings menu, mute the sound, or select the volume. Turn off music. Choose a language.

View Control Settings: The user can perform a Control View to move around the playing field and manipulate objects.

View game information: the user can view information about the game, goals and objectives of the game.

Quit Game: The user can log out of the game.

Next, consider the use cases in the game scene, presented in Figure 2.

Pause menu: the player can pause the gameplay (if after 15 minutes the player does not return, the main menu opens). Then he has to resume it or go to the main menu.

Controls: the user can control the behavior of the character by moving to the right, left, selecting an item and moving the item.

Bag selection: the player can look at the items that he has in the inventory. They can be connected and modified, or taken in hand.

Hint: the player can ask the game for a hint.

The player is responsible for moving the character around the playing field and manipulating objects. The character must move on the floor, not bump into obstacles in the form of furniture, take various objects. The level is considered not passed in the following cases:

- playing time ended, and the player was unable to solve the tasks assigned to him;
- the player solved the problems with errors.

Consider the activity diagram (Fig. 3). This diagram shows how the structure of a typical problem in the game will look like, i. e. structural element of the quest. Any task can be completed in at least three ways: the path where the user can make a mistake, the path where the user can complete the quest, and the path, and the path where the user does not have time to complete the quest.

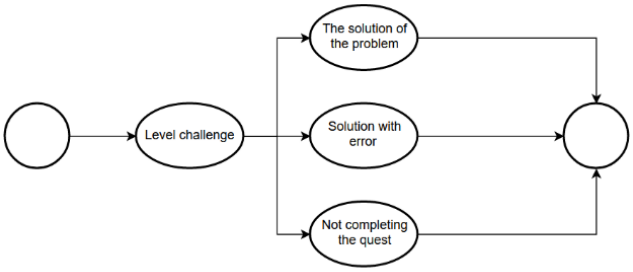

Fig. 3. Diagram of activity of the general case

A non-linear sequence of related tasks defines the plot [25]. The possibility of three endings for all tasks performed is considered. The student's performance of the tasks depends on the available ending options. If the student completes all the quests correctly, then he gets the maximum mark for the game. If a student completes the tasks incorrectly, then the student receives an average grade and two options, either to stay with the grade he earned, or to go through additional quests to increase his grade. And if the student completed all the quests with gross errors, or did not complete them at all, then the student is invited to go through the game again, or be left with an unsatisfactory mark for the game.

The game is located only in one location - the ship, for which it is necessary to choose the means and methods of information security. The player sees the ship only from the inside.

Some quests will be implemented by the mechanics of dialogue, when it will be necessary to write letters to employees in compliance with the rules of business correspondence, where at each stage there is a choice of three options, each of which has its own consequences. If a student makes gross mistakes when solving a problem, then he will have to go through the quest again in order to get points for him.

Conclusion

The use of game mechanics in staff training is a fairly effective way of presenting information, which is aimed at increasing students' interest in their chosen profession. Simulators help students to try their hand at practice. By combining these two educational technologies, you can get a simulator on which the student will be interested in practicing and improving his knowledge.

Acknowledgements

We should also single out Sergey Kolesnichenko for his quality advice when working on an article. We should also thank Alexey Astapkovich for his help in checking the text and creating a working environment.

References

1. Sokolov S. S., Mitrofanova A. V., Glebov N. B., et al. Formation of the Electronic Educational Trajectory for Maritime Students, *Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS): Proceedings of 2018 IEEE International Conference, Saint Petersburg, Russia, September 24–28, 2018*. Institute of Electrical and Electronics Engineers, 2018, Pp. 150–153.
DOI: 10.1109/ITMQIS.2018.8524913.

2. Vishtak O. V., Abushaev I. R. Trebovaniya k konstruktoru veb-kvestov [Requirements Designer Web Quests], *Sovremennye web-tekhnologii v tsifrovom obrazovanii: znachenie, vozmozhnosti, realizatsiya: Sbornik statey uchastnikov V Mezhdunarodnoy nauchno-prakticheskoy konferentsii [Modern Web Technologies in Digital Education: Meaning, Opportunities, Realization: Collection of Research Articles of the V International Scientific and Practical Conference], Arzamas, Russia, May 17–18, 2019*. Arzamas, Arzamas Branch of Lobachevsky University, 2019, Pp. 74–77. (In Russian)

3. Kovalnogova N. M., Sokolov S. S., Chernyi S. G., et al. Applying of E-learning and Distance Learning Technologies in Electronic Informational-Educational Environment in Modern University Complex, *Proceedings of the XIX International Conference on Soft Computing and Measurements (SCM`2016), Saint Petersburg, Russia, May 25–27, 2016*. Institute of Electrical and Electronics Engineers, 2016, Pp. 446–448. DOI: 10.1109/SCM.2016.7519809.

4. Small G. W., Vorgan G. iBrain: Surviving the Technological Alteration of the Modern Mind. HarperCollins Publishers, 2008, 256 p.

5. Skalozub I. S. Kompyuternye igry kak sredstvo razvitiya kommunikativnykh i lichnostnykh osobennostey podrostkov [Computer Games as a Means of Developing Communicative and Personal Characteristics of Adolescents], *Molodoy uchenyy [Young Scientist]*, 2015, No. 4 (84), Pp. 674–677. (In Russian)

6. Ilchenko L. M., Sokolov S. S., Egorova K. V., et al. Analysis of the Main Types of Automated Information Systems in the Transport Sector, Functioning Mainly in Urban Agglomerations, *Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Moscow, Russia, January 26–28, 2021*. Institute of Electrical and Electronics Engineers, 2021, Pp. 401–406.
DOI: 10.1109/ElConRus51938.2021.9396630.

7. Egorova K. V., Ilchenko L. M., Sokolov S. S., et al. Analysis of Automated Transport Information Systems in the Context of Critical Information Infrastructure of the Russian Federation, *Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Moscow, Russia, January 26–28, 2021*. Institute of Electrical and Electronics Engineers, 2021, Pp. 310–313.
DOI: 10.1109/ElConRus51938.2021.9396083.

8. Kardakova M. V., Shipunov I. S., Nyrkov A. P., Knysh T. P. Cyber Security on Sea Transport. In: *Murgul V., Pasetti M. (eds.) Energy Management of Municipal Facilities and Sustainable Energy Technologies (EMMFT 2018): Proceedings of the 20th International Scientific Conference, Voronezh, Samara, Russia, December 10–13, 2018. Volume 1*. Cham, Springer Nature, 2019, Pp. 481–490. (Advances in Intelligent Systems and Computing, Vol. 982).
DOI: 10.1007/978-3-030-19756-8_46.

9. Sokolov S. S., Saveleva M. N., Mitrofanova A. V., et al. Implementation of Training Programs Using Digital Distance Education Technologies for Seafarers, *Proceedings of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Moscow, Russia, January 27–30, 2020*. Institute of Electrical and Electronics Engineers, 2020, Pp. 521–525.
DOI: 10.1109/EIConRus49466.2020.9039034.

10. Shipunov I. S., Voevodskiy K. S., Nyrkov A. P., et al. About the Problems of Ensuring Information Security on Unmanned Ships, *Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Moscow, Russia, January 28–31, 2019*. Institute of Electrical and Electronics Engineers, 2019, Pp. 339–343.
DOI: 10.1109/EIConRus.2019.8657219.

11. Shipunov I. S., Voevodskiy K. S., Nyrkov A. P., et al. Trusted Transport Telemetry by Using Distributed Databases, *Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Moscow, Russia, January 28–31,*

2019. Institute of Electrical and Electronics Engineers, 2019, Pp. 344–347. DOI: 10.1109/EIConRus.2019.8657215.

12. Sokolov S. S., Alimov O. M., Nekrashevich P. S., et al. Security Issues and IoT Integration for in Russian Industry, *Proceedings of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Moscow, Russia, January 27–30, 2020*. Institute of Electrical and Electronics Engineers, 2020, Pp. 517–520. DOI: 10.1109/EIConRus49466.2020.9039213.

13. Kotenko I. V., Saenko I. B., Lauta O. S. Analytical Modeling and Assessment of Cyber Resilience on the Base of Stochastic Networks Conversion, *Proceedings of the 10th International Workshop on Resilient Networks Design and Modeling (RNDM 2018), Longyearbyen, Norway, August 27–29, 2018*. Institute of Electrical and Electronics Engineers, 2018, Pp. 1–8. DOI: 10.1109/RNDM.2018.8489830.

14. Faustova O. G. Metodika otsenki riskov vozniknoveniya chrezvychaynykh situatsiy v multimodalnykh perevozkakh [Technique of an Assessment of Risks of Emergency Situations in Multimodal Transportations], *Vestnik Astrakhanskogo gosudarstvennogo tekhnicheskogo universiteta. Seriya: Morskaya tekhnika i tekhnologiya [Vestnik of Astrakhan State Technical University. Series: Marine Engineering and Technologies]*, 2014, No. 1, Pp. 109–116. (In Russian)

15. Veselkov V. V., Vikhrov N. M., Nyrkov A. P., et al. Development of Methods to Identify Risks to Build Up the Automated Diagnosis Systems, *Proceedings of the 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Moscow, Russia, February 01–03, 2017*. Institute of Electrical and Electronics Engineers, 2017, Pp. 598–601.
DOI: 10.1109/EIConRus.2017.7910625.

16. Saenko I. B., Ageev S. A., Kotenko I. V. Detection of Traffic Anomalies in Multi-Service Networks Based on a Fuzzy Logical Inference. In: *Badica C., et al. (eds.) Intelligent Distributed Computing X (IDC`2016): Proceedings of the 10th International Symposium on Intelligent Distributed Computing Paris, France, October 10–12, 2016*. Cham, Springer Nature, 2017, Pp. 79–88. (Studies in Computational Intelligence, Vol. 678). DOI: 10.1007/978-3-319-48829-5_8.

17. Sokolov S. S., Alimov O. M., Golubeva M. G. The Automating Process of Information Security Management, *Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Moscow, Saint Petersburg, Russia, January 29–February 01, 2018*. Institute of Electrical and Electronics Engineers, 2018, Pp. 124–127. DOI: 10.1109/EIConRus.2018.8317045.

18. Sokolov S. S., Glebov N. B., Novoselov R. O., Chernyi S. G. Database Problems of Maritime Transport Industry on High Load Platform, TransSiberia 2018: *Proceedings of the Siberian Transport Forum, Novosibirsk, Russia, May 16–19, 2018. MATEC Web of Conferences*, 2018, Vol. 239, Art. No. 03004, 6 p. DOI: 10.1051/matecconf/201823903004.

19. Anardovich S. S., Rush E. A. Metody otsenki riskov i prognozirovanie stsenariev razvitiya chrezvychaynykh situatsiy pri zheleznodorozhnykh perevozkakh [Risk Assessment Methods and Forecasting Emergency Development Scenarios for Railway Transportation], S*ovremennye tekhnologii. Sistemnyy analiz. Modelirovanie [Modern Technologies. System Analysis. Modeling]*, 2020, Vol. 65, No. 1, Pp. 66–75.
DOI: 10.26731/1813-9108.2020.1(65).66-75. (In Russian)

20. Egorova K. V., Ilchenko L. M., Sokolov S. S., Katorin Yu. F. Analysis of Automated Transport Information Systems in the Context of Critical Information Infrastructure of the Russian Federation, *Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Moscow, Russia, January 26–28, 2021*. Institute of Electrical and Electronics Engineers, 2021, Pp. 310–313.
DOI: 10.1109/ElConRus51938.2021.9396083.

21. Sokolov S. S., Gaskarov V. D., Knysh T. P., Sagitova A. S. IoT Security: Threats, Risks, Attacks. In: *Mottaeva A. (ed.) Proceedings of the XIII International Scientific Conference on Architecture and Construction (ISCAC 2020), Novosibirsk, Russia, September 22–24, 2020*. Singapore, Springer Nature Singapore, 2020, Pp. 47–56. (Lecture Notes in Civil Engineering, Vol. 130). DOI: 10.1007/978-981-33-6208-6_6.

22. Burlov V. G., Grachev M. I. Management Model in Digital Ecosystems, *Germany and Russia: Ecosystems Without Borders: Proceedings of the International Conference Within the Framework of the VIII Annual International Baltic Maritime Forum, Kaliningrad, Russia, 05–10 October 2020. IOP Conference Series: Earth and Environmental Science*, 2021, Vol. 689, Art. No. 012010, 7 p.
DOI: 10.1088/1755-1315/689/1/012010.

23. Sokolov S. S., Zhilenkov A. A., Chernyi S. G., et al. Hybrid Neural Networks in Cyber Physical System Interface Control Systems, *Bulletin of Electrical Engineering and Informatics*, 2020, Vol. 9, No. 3, Pp. 1268–1275.
DOI: 10.11591/eei.v9i3.1293.

24. Sokolov S. S., Saenko I. B., Mitrofanov M. A., et al. Analytical Modeling of Computer Attacks on Intelligent Transport Systems Based on the Transformation of Stochastic Networks. In: *Kovalev S., et al. (eds.) Intelligent Information Technologies for Industry (IITI`21): Proceedings of the Fifth International Scientific Conference Sirius, Russia, September 30–October 04, 2021*. Cham, Springer Nature, 2021, Pp. 489–498. (Lecture Notes in Networks and Systems, Vol. 330).
DOI: 10.1007/978-3-030-87178-9_49.

25. Glebov N. B., Zhilenkov A. A., Chernyi S. G., Sokolov S. S. Process of the Positioning Complex Modeling Objects with Elements of Intellectual Analysis, *Procedia Computer Science*, 2019, Vol. 150, Pp. 609–615.
DOI: 10.1016/j.procs.2019.02.094.

DOI: 10.24412/2413-2527-2022-331-52-60

Концепция игрового тренажера по защите информации на водном транспорте

М. В. Кардакова, д.т.н. А. П. Ныркov, Ю. В. Цымай
Государственный университет морского и речного флота имени адмирала С. О. Макарова
Санкт-Петербург, Россия
m.v.kardakova@ya.ru, apnyrkow@mail.ru, m-walua@yandex.ru

Аннотация. **В статье рассматривается возможность применения игровых тренажеров при обучении персонала или студентов. Рассмотрена актуальность применения тренажеров при обучении студентов информационной безопасности на водном транспорте. Актуальность таких тренажеров подчеркивается возможностью их использования в дистанционном обучении. Использование узнаваемого жанра, знакомого и популярного среди школьников в виртуальном мире, позволяет получить положительные эмоции, повысить интерес и мотивацию к развитию познавательной деятельности. Описаны игровые механики, применяемые в тренажере. Рассмотрен пример расчета риска возникновения чрезвычайной ситуации на судне, предложены варианты использования расчетов в тренажере. Описана концепция самого тренажера и разработана концепция интерфейса. Использование тренажеров в обучении позволяет научить студентов практическим навыкам работы с чем-либо. Как правило, такой подход к обучению требует от студента знания теории предмета, а также активного участия в работе. Таким образом, у студента есть возможность продемонстрировать свои знания на практике и повысить квалификацию. Конечно, никакой тренажер не заменит реальный прибор или реальную ситуацию, но такой подход в обучении поможет закрепить материал и в будущем ученик, увидев реальный объект, не потеряется и сможет спокойно работать с ним. Использование тренажеров информационной безопасности в образовательных целях обеспечивает закрепление теоретических знаний и их применение на практике. Очень важно, чтобы тренажеры, с одной стороны, были приближены к реальным условиям, а с другой, чтобы работа на них была интересной и удобной для обучающегося.**

Ключевые слова: **транспорт, водный транспорт, судно, корабль, тренажер, игра, образование, риск.**

Литература

1. Formation of the Electronic Educational Trajectory for Maritime Students / S. S. Sokolov, A. V. Mitrofanova, N. B. Glebov, [et al.] // Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS): Proceedings of 2018 IEEE International Conference (Saint Petersburg, Russia, 24–28 September 2018). — Institute of Electrical and Electronics Engineers, 2018. — Pp. 150–153.
DOI: 10.1109/ITMQIS.2018.8524913.

2. Виштак, О. В. Требования к конструктору вебквестов / О. В. Виштак, И. Р. Абушаев // Современные web-технологии в цифровом образовании: значение, возможности, реализация: Сборник статей участников V Международной научно-практической конференции (Арзамас, Россия, 17–18 мая 2019 г.). — Арзамас: Арзамасский филиал ННГУ им. Лобачевского, 2019. — С. 74–77.

3. Applying of E-learning and Distance Learning Technologies in Electronic Informational-Educational Environment in Modern University Complex / N. M. Kovalnogova, S. S. Sokolov, S. G. Chernyi, [et al.] // Proceedings of the XIX International Conference on Soft Computing and Measurements (SCM`2016), (Saint Petersburg, Russia, 25–27 May 2016). — Institute of Electrical and Electronics Engineers, 2016. — Pp. 446–448. DOI: 10.1109/SCM.2016.7519809.

4. Small, G. W. iBrain: Surviving the Technological Alteration of the Modern Mind / G. W. Small, G. Vorgan. — HarperCollins Publishers, 2008. — 256 p.

5. Скалозуб, И. С. Компьютерные игры как средство развития коммуникативных и личностных особенностей подростков // Молодой ученый. 2015. № 4 (84). С. 674–677.

6. Analysis of the Main Types of Automated Information Systems in the Transport Sector, Functioning Mainly in Urban Agglomerations / L. M. Ilchenko, S. S. Sokolov, K. V. Egorova, [et al.] // Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), (Saint Petersburg, Moscow, Russia, 26–28 January 2021). — Institute of Electrical and Electronics Engineers, 2021. — Pp. 401–406.
DOI: 10.1109/ElConRus51938.2021.9396630.

7. Analysis of Automated Transport Information Systems in the Context of Critical Information Infrastructure of the Russian Federation / K. V. Egorova, L. M. Ilchenko, S. S. Sokolov, [et al.] // Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), (Saint Petersburg, Moscow, Russia, 26–28 January 2021). — Institute of Electrical and Electronics Engineers, 2021. — Pp. 310–313.
DOI: 10.1109/ElConRus51938.2021.9396083.

8. Cyber Security on Sea Transport / M. V. Kardakova, I. S. Shipunov, A. P. Nyrkov, T. P. Knysh // Energy Management of Municipal Facilities and Sustainable Energy Technologies (EMMFT 2018): Proceedings of the 20th International Scientific Conference (Voronezh, Samara, Russia, 10–13 December 2018). Volume 1 / V. Murgul, M. Pasetti (eds.). — Cham: Springer Nature, 2019. — Pp. 481–490. — (Advances in Intelligent Systems and Computing. Vol. 982).
DOI: 10.1007/978-3-030-19756-8_46.

9. Implementation of Training Programs Using Digital Distance Education Technologies for Seafarers / S. S. Sokolov, M. N. Saveleva, A. V. Mitrofanova, [et al.] // Proceedings of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), (Saint Petersburg, Moscow, Russia, 27–30 January 2020). — Institute of Electrical and Electronics Engineers, 2020. — Pp. 521–525. DOI: 10.1109/EIConRus49466.2020.9039034.

10. About the Problems of Ensuring Information Security on Unmanned Ships / I. S. Shipunov, K. S. Voevodskiy,

A. P. Nyrkov, [et al.] // Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), (Saint Petersburg, Moscow, Russia, 28–31 January 2019). — Institute of Electrical and Electronics Engineers, 2019. — Pp. 339–343.
DOI: 10.1109/EIConRus.2019.8657219.

11. Trusted Transport Telemetry by Using Distributed Databases / I. S. Shipunov, K. S. Voevodskiy, A. P. Nyrkov, [et al.] // Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), (Saint Petersburg, Moscow, Russia, 28–31 January 2019). — Institute of Electrical and Electronics Engineers, 2019. — Pp. 344–347. DOI: 10.1109/EIConRus.2019.8657215.

12. Security Issues and IoT Integration for in Russian Industry / S. S. Sokolov, O. M. Alimov, P. S. Nekrashevich, [et al.] // Proceedings of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), (Saint Petersburg, Moscow, Russia, 27–30 January 2020). — Institute of Electrical and Electronics Engineers, 2020. — Pp. 517–520. DOI: 10.1109/EIConRus49466.2020.9039213.

13. Kotenko, I. V. Analytical Modeling and Assessment of Cyber Resilience on the Base of Stochastic Networks Conversion / I. V. Kotenko, I. B. Saenko, O. S. Lauta // Proceedings of the 10th International Workshop on Resilient Networks Design and Modeling (RNDM 2018), (Longyearbyen, Norway, 27–29 August 2018). — Institute of Electrical and Electronics Engineers, 2018. — Pp. 1–8.
DOI: 10.1109/RNDM.2018.8489830.

14. Фаустова, О. Г. Методика оценки рисков возникновения чрезвычайных ситуаций в мультимодальных перевозках // Вестник Астраханского государственного технического университета. Серия: Морская техника и технология. 2014. № 1. С. 109–116.

15. Development of Methods to Identify Risks to Build Up the Automated Diagnosis Systems / V. V. Veselkov, N. M. Vikhrov, A. P. Nyrkov, [et al.] // Proceedings of the 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), (Saint Petersburg, Moscow, Russia, 01–03 February 2017). — Institute of Electrical and Electronics Engineers, 2017. — Pp. 598–601.
DOI: 10.1109/EIConRus.2017.7910625.

16. Saenko, I. B. Detection of Traffic Anomalies in Multi-Service Networks Based on a Fuzzy Logical Inference / I. B. Saenko, S. A. Ageev, I. V. Kotenko // Intelligent Distributed Computing X (IDC`2016): Proceedings of the 10th International Symposium on Intelligent Distributed Computing (Paris, France, 10–12 October 2016) / C. Badica, [et al.] (eds.). — Cham: Springer Nature, 2017. — Pp. 79–88. — (Studies in Computational Intelligence. Vol. 678).
DOI: 10.1007/978-3-319-48829-5_8.

17. The Automating Process of Information Security Management / S. S. Sokolov, O. M. Alimov, M. G. Golubeva // Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), (Moscow, Saint Petersburg, Russia, 29 January–01 February 2018). — Institute of Electrical and Electronics Engineers, 2018. — Pp. 124–127.
DOI: 10.1109/EIConRus.2018.8317045.

18. Database Problems of Maritime Transport Industry on High Load Platform / S. S. Sokolov, N. B. Glebov, R. O. Novoselov, S. G. Chernyi // TransSiberia 2018: Proceedings of the Siberian Transport Forum (Novosibirsk, Russia, 16–19 May 2018). MATEC Web of Conferences. 2018. Vol. 239. Art. No. 03004. 6 p. DOI: 10.1051/matecconf/201823903004.

19. Анардович, С. С. Методы оценки рисков и прогнозирование сценариев развития чрезвычайных ситуаций при железнодорожных перевозках / С. С. Анардович, Е. А. Руш // Современные технологии. Системный анализ. Моделирование. 2020. Т. 65, № 1. С. 66–75.
DOI: 10.26731/1813-9108.2020.1(65).66-75.

20. Analysis of Automated Transport Information Systems in the Context of Critical Information Infrastructure of the Russian Federation / K. V. Egorova, L. M. Ilchenko, S. S. Sokolov, Yu. F. Katorin // Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), (Saint Petersburg, Moscow, Russia, 26–28 January 2021). — Institute of Electrical and Electronics Engineers, 2021. — Pp. 310–313.
DOI: 10.1109/ElConRus51938.2021.9396083.

21. IoT Security: Threats, Risks, Attacks / S. S. Sokolov, V. D. Gaskarov, T. P. Knysh, A. S. Sagitova // Proceedings of the XIII International Scientific Conference on Architecture and Construction (ISCAC 2020), (Novosibirsk, Russia, 22–24 September 2020) / A. Mottaeva (ed.). — Singapore: Springer Nature Singapore, 2020. — Pp. 47–56. — (Lecture Notes in Civil Engineering. Vol. 130). DOI: 10.1007/978-981-33-6208-6_6.

22. Burlov, V. G. Management Model in Digital Ecosystems / V. G. Burlov, M. I. Grachev // Germany and Russia: Ecosystems Without Borders: Proceedings of the International Conference Within the Framework of the VIII Annual International Baltic Maritime Forum (Kaliningrad, Russia, 05–10 October 2020). IOP Conference Series: Earth and Environmental Science. 2021. Vol. 689. Art. No. 012010. 7 p.
DOI: 10.1088/1755-1315/689/1/012010.

23. Hybrid Neural Networks in Cyber Physical System Interface Control Systems / S. S. Sokolov, A. A. Zhilenkov, S. G. Chernyi, [et al.] // Bulletin of Electrical Engineering and Informatics. 2020. Vol. 9, No. 3. Pp. 1268–1275.
DOI: 10.11591/eei.v9i3.1293.

24. Analytical Modeling of Computer Attacks on Intelligent Transport Systems Based on the Transformation of Stochastic Networks / S. S. Sokolov, I. B. Saenko, M. A. Mitrofanov, [et al.] // Intelligent Information Technologies for Industry (IITI`21): Proceedings of the Fifth International Scientific Conference (Sirius, Russia, 30 September–04 October 2021) / S. Kovalev, [et al.] (eds.). — Cham: Springer Nature, 2021. — Pp. 489–498. — (Lecture Notes in Networks and Systems. Vol. 330). DOI: 10.1007/978-3-030-87178-9_49.

25. Process of the Positioning Complex Modeling Objects with Elements of Intellectual Analysis / N. B. Glebov, A. A. Zhilenkov, S. G. Chernyi, S. S. Sokolov // Procedia Computer Science. 2019. Vol. 150. Pp. 609–615.
DOI: 10.1016/j.procs.2019.02.094.

DOI: 10.24412/2413-2527-2022-331-61-67

Analysis Method of the Stability of the Combined Labeling of Digital Audio Signals

Grand PhD A. A. Kornienko, PhD M. V. Gofman, PhD S. V. Kornienko
Emperor Alexander I St. Petersburg State Transport University
Saint Petersburg, Russia
kaa.pgups@ya.ru, offmail3000@mail.ru, sv.diass99@ya.ru

Abstract. **The article analyzes the stability of the method of combined marking of digital audio signals. The following are the main characteristics of labelling quality and stability: auditory transparency (inaudible) marker, lossy compression resistance and resistance to stegoanalysis, airway transmission immunity. The article contains quantitative estimates of indicators for the specified characteristics which received analytically from simulations and field experiments. To assess auditory transparency (inaudible) of acoustic artifacts that appear as a result of the use of the marker analytical method is used, which is based on the signal-marker ratio metric. Lossy compression stability is applied to MP3 transformations under different compression modes. Resistance to stegoanalysis is evaluated using a universal method, which uses the ratio of number of units to zeros in the least significant digitized audio signal reference grades. Evaluation of audio immunity over the air audio channel by correct/incorrect rate of marker rectoration based on the results of field experiments which were implemented in laboratory settings.**

Keywords: **steganography, digital audio signal, combined labeling, quality and sustainability evaluation, air audio channel, immunity transmission.**

INTRODUCTION

One of the main problems with steganographic protection of audio information and copyright is the problem of quality assurance and sustainability of digital labeling to various transition, to stegoanalysis, to the action of noise and liability in audio channel. This is particularly important when the required sustainability in the marking of digital audio signals which are embedded in audiostego systems with multiple (spatial) inputs and multiple (spatial) outputs is ensured. Such audio stegosystem have several acoustic dynamics (transmitters) and several microphones (receivers) and allows to implement a label in digital audio signal settings to spatial, frequency and time areas, it means that is to do a combination labeling. The spatial multichannel of such audio stegosystem also makes it possible to use methods MIMO-technologies to achieve the required immunity.

Quality analysis of labeling digital audio signals is primarily related to evaluation of auditory transparency (inaudible) or audibility of acoustic artifacts resulting from the implementation of the label that defines thresholds for the insertion force of a label.

Audibility (non-audibility) evaluation methods of acoustic artifacts are divided into analytical methods which use different metrics or indicators [1–3], automatic methods [4–8], based on the use of various psychoacoustic models of human hearing organs, manual methods of hearing evaluation of markers [9], most of which are test adaptations that used to assess the quality of audio information.

Evaluation of the sustainability of digital audio signal labeling involves the use of methods to assess sustainability to digital transformation, to stegoanalysis, acoustic noise/liability resistance (immunity) in the channel, for example, in the air acoustic channel, sustainability to other impacts.

Assessing the resistance of labeling methods to transformations is reduced to assessing the resistance to various types of attacks on labelled audio information and audio signals. The article [10] focusses on the sustainability of digital data labelling. Attacks on digital data labelling systems can be divided into the following: active attacks, passive attacks, introduction attacks, collusion attacks, simple attacks, ambiguity attacks, removal attacks, synchronization attacks and other attacks. Simple attacks are the most commonly used type of attacks. The article [11] describes Stirmark software that allows you to perform simple attacks. Simple attacks on digital audio signal, for example, include a lossy compression attack.

An overview of digital audio stegoanalysis methods is presented in article [12]. Methods of stegoanalysis of digital audio signals can be divided into two classes. A class of methods aimed at stegoanalysis of audio signals presented in uncompressed format (for example, FLAC or WAV), and a class of methods aimed at stegoanalysis of audio signals presented in compressed format (for example, MP3 or AAC). In turn, the first class of methods can be divided into two subclasses. The first subclass includes methods focused on predefined implementation algorithms — these are the so-called «targeted» methods of stegoanalysis. The second subclass includes methods in which it is either not expected to have knowledge about the implementation algorithm in advance, or a certain set of possible or assumed implementation algorithms are known — these are the so-called «universal» methods of stegoanalysis. The class of universal methods can also be divided into two subclasses. The first subclass is methods in which the decision on the presence of a marker is made only on the basis of the properties of the stegoanalysed audio signal. The second subclass is methods in which the solution is made on the basis of comparing the properties of a stegoanalysed audio signal with the properties of some other audio signal (usually obtained from a stegoanalysed signal), regarding which it is precisely known whether it is a stegoaudio signal or just an audio signal. Among the universal methods of stegoanalysis, the most common method based on extraction and analysis of the least significant bit, or LSB stegoanalysis [13].

Evaluation of the noise immunity of marked audio signals when transmitted over an air channel can be obtained based on the results of full-scale experiments and simulation results. Full-scale experiments on the transmission of a marked audio signal can be broadcasting of marked audio signals using

acoustic speakers while recording of the broadcast audio signal with an acoustic microphone to subsequently check whether the marker has been preserved in the recorded audio signal recorded by the microphone or not. Simulation modelling is based on the use of some model of air audio channel, which simulates the impact on the marked audio signal with subsequent assessment of the safety of the marker in the digital audio signal received at the output of the channel.

This article analyses and evaluates the main indicators of quality and stability of the method of combined digital marking of audio signals proposed and developed in articles [14–16] for audio stegosystems with multiple input and multiple output. The quality of the method is characterized by hearing transparency (inaudible). To assess the auditory transparency (inaudibleness) of acoustic artefacts that appear as a result of the introduction of a marker, the article uses an analytical method based on the signal-to-marker relationship metric.

Resistance to lossy compression attack is carried out in relation to MP3 transformations by simulation. Stegoanalysis resistance is assessed using the universal least significant bit method (LSB method), which uses the ratio of the number of units to zeros in the least significant bits of the marked digital audio signal. The assessment of noise immunity of the transmission of marked audio signals through an air audio channel is carried out on the basis of the results of full-scale experiments performed in the laboratory.

A software and hardware model were developed for simulation modelling of digital transformations (attacks). MATLAB was used as a software development environment for the model.

100 musical compositions included in the collection «100 Greatest 00s: The Best Songs from the Decade» were used as the original digital audio signals for labelling. Each of the original digital audio signals had the following parameters: bit depth — 16 bits/countdown, sampling rate — 44 100 countdown/sec, two audio tracks in each digital audio signal (stereo signals). Thus, the digital speed of all audio signals used was the same and equal: $16 \times 44\,100 \times 2 = 1\,411\,200$ bit/s.

Analysis and results of auditory transparency assessment

The auditory transparency assessment of the marker was carried out using the analytical method of checking the audibility of artefacts from the introduction of the marker. The verification method was based on the value of the signal-to-marker ratio

$$\mathrm{SWR} = 10 \lg \frac{\sum_n x_n^2}{\sum_n (x_n - y_n)^2},$$

where x_n and y_n — are the references of the original and labelled digital audio signals, respectively. The evaluated labelling method refers to a type of methods that mask the marker in the frequency region of the Fourier spectrum. For such methods, a value of 20 dB is usually used as a threshold for the signal-to-marker ratio. Thus, if for the analyzed marked digital audio signal, the value of this ratio is not less than this threshold, it is highly likely that the human ear will not hear acoustic artefacts resulting from the introduction of the marker.

As sequences $\boldsymbol{\alpha}_i = (\alpha_1\, \alpha_2 \cdots \alpha_{N_\alpha})$ and $\boldsymbol{\beta}_i = (\beta_1\, \beta_2 \cdots \beta_{N_\beta})$, used in the process of converting information according to the marking methods described in the works [14–16], Kasami and Gold sequences were used, respectively.

Sequences $\boldsymbol{\gamma}_i = (\gamma_1\, \gamma_2 \cdots \gamma_{N_\gamma})$, taking into account the similarity of the amplitude spectra of adjacent blocks, length N_{block}, readings of a separate audio track of the marked audio signal, were designed in accordance with the rule

$$\gamma_i = \varphi_i \otimes \mathrm{RZcode},$$

where $\otimes$ — is Kronecker's product, φ_i — is Kasami sequences of length 15, and $\mathrm{RZcode} = (1 \quad -1)$.

Hearing transparency was assessed with different sets of parameters $\{N_\alpha, N_\beta, N_\gamma, N_{block}\}$, but with a fixed embedding force of 0.1.

- At $N_\alpha = 1023, N_\beta = 31, N_\gamma = 30, N_{block} = 64$ and the force of embedding 0.1 of the marker element, the average value of the signal-to-marker ratio calculated for all 100 analysed musical compositions was 24.80661 dB, with the dispersion of this ratio equal to 2.382539 dB.
- At $N_\alpha = 1023, N_\beta = 63, N_\gamma = 30, N_{block} = 128$ and embedding force 0.1, the average value of the signal-to-marker ratio calculated from all 100 analysed musical compositions was 23.00086 dB, with the dispersion of this ratio equal to 1.074242 dB.
- At $N_\alpha = 1023, N_\beta = 127, N_\gamma = 30, N_{block} = 256$ and embedding force 0.1, the average value of the signal-to-marker ratio calculated for all 100 analysed musical compositions was 21.70835 dB, with the dispersion of this ratio equal to 0.439844 dB.
- At $N_\alpha = 1023, N_\beta = 511, N_\gamma = 30, N_{block} = 1024$ and embedding force 0.1, the average value of the signal-to-marker ratio calculated for all 100 musical compositions analysed was 20.68482 dB, with the dispersion of this ratio equal to 0.0446 dB.

Based on these results, it can be concluded that at fixed values N_α and N_γ, the increase in N_β and N_{block} leads to a decrease in the average value and variance of the signal-to-marker ratio. Thus, the analysed marking method provides auditory transparency to the marker when assessing audibility using an analytical method based on the value of the signal-to-marker ratio, when a value of 20 dB is used as a threshold.

Loss resilience assessment analyses and results

The evaluation of the resistance of the analyzed labelling method to a lossy compression attack using MP3 conversion was performed for different MP3 conversion modes and different bit compression speeds. Resistance to MP3 conversion in stereo, mono and joint stereo modes was checked. Combinations of modes and compression speeds were selected so that the original sampling rate of the compressed digital audio signal was maintained.

The MP3 attack procedure consisted of the following sequence of actions. Initially, the original digital audio signal stored in WAVE format was labelled using the developed labelling method. Further, the marked digital audio signal was subjected to an MP3 attack by compression in a certain mode and at a certain bit rate of compression. After that, the MP3 file was decompressed into a WAV file. Finally, it was checked whether the digital audio signal received at the MP3 decoder output retained the embedded token or not. SOUND FORGE Pro 14.0 Suite was used to perform an MP3 attack.

MP3 compression parameters in stereo mode were as follows: compression quality — Fastest encode; bit rate — 96 kbit/s; sampling rate — 44 100 Hz. MP3 decompression parameters: format — PCM (uncompressed); sampling rate — 44 100 Hz; number of bits to count — 16; number of audio tracks — 2 (stereo). The used bit rate of 96 kbit/s is the minimum allowed in this audio processing program at a sampling rate of 44 100 Hz in stereo mode. After the MP3 compression attack in stereo mode, all markers were detected without errors and the information bits encoded in them were restored without errors.

MP3 compression parameters in mono mode were as follows: compression quality — Fastest encode; bit rate — 48 kbit/s; sampling rate — 44 100 Hz. MP3 compression parameters: format — PCM (uncompressed); sampling rate — 44 100 Hz; number of bits to count — 16; number of audio tracks — 1 (mono). The used bit speed of 48 kbit/s is the minimum allowable audio processing program at a sampling rate of 44 100 Hz in mono mode. After an MP3 compression attack in mono mode, all markers without errors were detected and the information bits encoded in them were restored without errors.

MP3 compression parameters in joint stereo mode were as follows: compression quality — Fastest encode; bit rate — 96 kbit/s; sampling rate — 44 100 Hz; Mid/side stereo and Intensity stereo are used. MP3 decompression parameters: format — PCM (uncompressed); sampling rate — 44 100 Hz; number of bits to count — 16, number of audio tracks — 2 (stereo). The used bit speed of 48 kbit/s is the minimum allowable audio processing program at a sampling rate of 44 100 Hz in mono mode. The used bit rate of 96 kbit/s is the minimum allowed in this audio processing program at a sampling rate of 44 100 Hz in joint stereo mode. The joint stereo mode leads to the fact that after the reverse conversion to WAVE Stereo format, each of the two audio tracks will include not only for the most part (since after compression with loss some data is irretrievably lost) the original audio track, but also partially another audio track, which means that two markers can be found in each of the audio tracks. Usually, the lower the bit rate, the more audio tracks «mix» together and the more likely it is to detect a marker from the neighboring audio track in the analyzed audio track. After an MP3 attack in joint stereo mode at a bit speed of 96 kbit/s after analysing 200 audio tracks (100 stereo audio signals) in which each of the two markers was searched (thus, 400 checks were performed: two markers were searched in each of the 200 audio tracks after decompression, due to the possible introduction of a marker from the:

- 14 times out of 400 times a marker brought from the neighboring audio track was not found (this indicator is affected by bit speed);
- 0 times out of 400 times the original marker was not found on the corresponding audio track;
- 11 times out of 400 times there was an error in establishing synchronization with the introduced token (i. e. the introduced token was detected, but its beginning was not correctly defined);
- 3 times out of 400 times there was an error in establishing synchronization with the original token (i. e. the original marker of the corresponding track was detected, but its beginning was not correctly determined).

However, when the bit rate was increased to 320 kbit/s, no synchronization errors were found with the original token — all original tokens were detected and all information bits were correctly restored. It should be noted that at a speed of 320 kbit/s, the mixing of audio tracks as a result of compression was so small that markers introduced from neighboring audio tracks were not found in audio tracks after decompression.

Analysis and results of stegoanalysis resistance evaluation

Using the computer model of the audio stegosystem, the stability of the analyzed marking method to the least significant bit method (LSB method) was assessed. This method of stegoanalysis is a universal method, as it does not require knowledge of the marking method. The method of stegoanalysis of marked digital audio signals using the least significant bit method is based on the assumption that if the ratio of the number of units to zeros in binary bits of the analyzed binary plane of the digital signal reference values exceeds some threshold, it is likely that there is a hidden message in this bit plane. Usually, the threshold is 1.1. It is taken as a threshold. All 100 original digital audio signals have been labelled. All bit planes of reference of marked digital audio signals were analyzed. The sustainability of the developed marking method to stegoanalysis was assessed under different sets of parameters $\{N_\alpha, N_\beta, N_\gamma, N_{block}\}$:

- At $N_\alpha = 1023, N_\beta = 31, N_\gamma = 30, N_{block} = 64$ and embedding force 0.1, the average value of the ratio of units to zeros in binary bits along all 16 bit planes (musical compositions were stored in files with 16 bit depth of reference values) of all 100 analyzed musical compositions was 0.993058, with the dispersion of this ratio equal to 0.00080572.
- At $N_\alpha = 1023, N_\beta = 63, N_\gamma = 30, N_{block} = 128$ and embedding force 0.1, the average value of the ratio of units to zeros in binary bits along all 16-bit planes (musical compositions were stored in files with 16-bit depth of reference values) of all 100 analyzed musical compositions was 0.997205415, with the dispersion of this ratio equal to 0.000459767.

Thus, the developed marking method is resistant to LSB stegoanalysis, providing average values of the number of units to zero ratios for all bit planes are much closer to 1 than to threshold 1.1, while the variance of the average value is less than 10^{-3}.

Analysis and results of the assessment of noise immunity of the transmission of marked audio signals through the air audio channel

The evaluation of the stability of the developed method of marking digital audio signals to the interference of the air audio channel was carried out based on the results of full-scale experiments using a software and hardware laboratory unit. The software part of the laboratory installation was realized in the MATLAB environment. The hardware of the laboratory unit was as follows: the AKG Lyra 4-capsule microphone in Tight Stereo mode was used as an acoustic microphone, and two pairs of Edifier R1280DB speakers were used as acoustic speakers.

Four Kasami sequences of the same length $N_\alpha = 1023$ were used as vectors $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\boldsymbol{\alpha}_3$ and $\boldsymbol{\alpha}_4$ and four vectors $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$ and $\boldsymbol{\beta}_4$ as Gold sequences of the same length $N_\beta = 63$ were used. When constructing vectors $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$, $\boldsymbol{\gamma}_3$ and $\boldsymbol{\gamma}_4$ as vectors $\boldsymbol{\varphi}_1$, $\boldsymbol{\varphi}_2$, $\boldsymbol{\varphi}_3$ and $\boldsymbol{\varphi}_4$ four different Kasami sequences of the same length 15 were used, and the vector $(1 \quad -1)$ was used

as a RZcode word; as a result, $N_\gamma = 30$. The size of the reference block was equal to $N_{block} = 2(N_\beta + 1) = 128$. At such values of marker parameters and sampling rate of F_s = 44 100 Hz, the transmission of one marker takes

$$\frac{N_\alpha N_{block} N_\gamma}{F_s} \cong 89.1 \text{ s}.$$

The value for the amount (determining the size of the analysis window) was taken to 8 093.

Two types of indoor field experiments were carried out to assess the resistance to noise and interference of the air audio channel. The first type of experiment included transmitting marked stereo audio signals using a pair of speakers and stereo recording of broadcast audio signals using a microphone located in line of sight at a distance of 3 meters. The second type of experiment included the transmission of two marked stereo audio signals with two pairs of speakers and stereo recording of broadcast audio signals using a microphone, also located in line of sight at a distance of 3 meters. The second type of experiment simulated a possible situation in practice when there is a significant extraneous acoustic noise during broadcasting.

During the first type of experiment, each of the 100 source stereo audio signals was labelled. Two markers were introduced into each of the two audio tracks of the original digital audio signal. If the duration of the audio signal was not enough to introduce two markers, the audio signal was artificially extended by cyclic repetition from the beginning. At the same time, markers in one audio track were separated from each other in time by a protective interval of 10 % of their own duration. Transmission (broadcasting) through the air audio channel of different marked stereo audio signals was carried out separately. Thus, 400 markers were transmitted through the air audio channel. Speaker amplifiers were set to 4 out of 50 divisions — this gain was minimal, at which the broadcast audio signal was barely audible. The results of the first type of experiment showed that the developed marking method is resistant to the influences of an air audio channel when using Edifier R1280DB speakers and an AKG Lyra microphone located at a distance of 3 meters in line of sight. Thus, all 400 markers were found and the information bits in them encoded were correctly restored.

For the second type of experiment, 50 digital stereo audio signals were randomly selected from 100 original stereo audio signals, which were labelled. Two markers were introduced into each of the two audio tracks. If the duration of the audio signal was not enough to introduce two markers, the audio signal was artificially extended by cyclic repetition from the beginning. At the same time, markers in one audio track were separated from each other in time by a protective interval of 10 % of their own duration. After marking, 25 pairs were randomly formed from 50 marked audio signals. Each stereo audio signal from the pair was broadcast to the air audio channel simultaneously with another stereo audio signal from the same pair using four speakers. This broadcasting of quadro-audio signals, consisting of four audio tracks, simulated the situation of significant extraneous acoustic noise. Thus, 200 markers were transmitted through the air audio channel. Speaker amplifiers were installed on 10 out of 50 divisions. The second type of field experiments showed that the developed marking method is quite resistant to the influences of the air audio channel when using Edifier R1280DB speakers and the AKG Lyra microphone, located at a distance of 3 meters in line of sight. As a result of the transfer, only three markers out of 200 could not detect and restore information.

Conclusion

The article analyzed and evaluated various indicators of quality and stability of the combined method of marking digital audio signals, developed in articles [14–16]. The results of hearing transparency analysis showed that the developed marking method provides a signal-to-marker ratio of at least 20 dB, which is sufficient to prevent acoustic artefacts that inevitably appear as a result of the introduction of markers into digital audio signal. The results of the lossy compression resistance assessment showed that in the case of MP3 compression in mono and stereo modes, the embedded marker can also be detected and the information transferred by it can be restored. When MP3 compression is used in joint stereo mode, then, due to the mixing of stereo audio tracks during compression, synchronization errors may occur, but when the bit rate is 320 kbit/s, then there are no synchronization errors. The results of the lossy compression resistance assessment showed that in the case of MP3 compression in mono and stereo modes, the embedded marker can also be detected and the information transferred by it can be restored. When MP3 compression is used in joint stereo mode, then, due to the mixing of stereo audio tracks during compression, synchronization errors may occur, but when the bit rate is 320 kbit/s, then there are no synchronization errors. The results of the assessment of stegoanalysis resistance using the universal least significant bit method (LSB method) showed that changes in digital audio signals are such that the one-to-zero ratios in individual bit planes of the marked digital audio signal do not exceed threshold 1.1, which in the research article [13] is designated as an indicator of the presence of a marker in digital audio signal. The results of the experimental noise resistance assessment of the air audio channel showed that it is possible to steadily transmit marked stereo audio signals. Under the specified conditions of the experiment, it was possible to detect and restore information sequences of 197 markers out of 200.

References

1. Hu Y., Loizou P. C. Evaluation of Objective Quality Measures for Speech Enhancement, *IEEE Transactions on Audio, Speech, and Language Processing*, 2008, Vol. 16, No. 1, Pp. 229–238. DOI: 10.1109/TASL.2007.911054.

2. Özer H., Sankur B., Memon N., Avcıbaş İ. Detection of Audio Covert Channels Using Statistical Footprints of Hidden Messages, *Digital Signal Processing*, 2006, Vol. 16, Is. 4, Pp. 389–401. DOI: 10.1016/j.dsp.2005.12.001.

3. Kutter M., Petitcolas F. A. P. Fair Benchmark for Image Watermarking Systems. In: *Wong P. W., Delp III E. J. (eds.) Proceedings of the SPIE/IS&T Electronic Imaging 1999. Security and Watermarking of Multimedia Contents, San Jose, CA, USA, January 23–29, 1999. Proceedings of SPIE*, 1999, Vol. 3657, Pp. 226–239. DOI: 10.1117/12.344672.

4. GOST R 54711-2011. Zvukovoe veshchanie tsifrovoe. Kodirovanie signalov tsifrovogo veshchaniya s sokrashcheniem izbytochnosti dlya peredachi po tsifrovym kanalam svyazi. MPEG-1 chast III (MPEG-1 audio) [Digital sound broadcasting. Coding of sound broadcasting signals with redundancy reduction for transfer on digital communication

channels. MPEG-1 part III (MPEG-1 audio)]. Effective from July 01, 2013. Moscow, StandartInform Publishing House, 2014, 172 p. (In Russian)

5. GOST R 54712-2011. Zvukovoe veshchanie tsifrovoe. Kodirovanie signalov tsifrovogo veshchaniya s sokrashcheniem izbytochnosti dlya peredachi po tsifrovym kanalam svyazi. MPEG-2 chast III (MPEG-2 audio) [Digital sound broadcasting. Coding of sound broadcasting signals with redundancy reduction for transfer on digital communication channels. MPEG-2 part III (MPEG-2 audio)]. Effective from July 01, 2013. Moscow, StandartInform Publishing House, 2014, 109 p. (In Russian)

6. Recommendation ITU-R BS.1196-8. Audio coding for digital broadcasting. Approved in 14.10.2019. International Telecommunication Union, 2019, 43 p. Available at: http://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1196-8-201910-I!!PDF-E.pdf (accessed 08 Nov 2021).

7. Recommendation ITU-R BS.1387-1. Method for objective measurements of perceived audio quality. Approved in 24.11.2001. International Telecommunication Union, 2001, 100 p. Available at: http://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1387-1-200111-I!!PDF-E.pdf (accessed 08 Nov 2021).

8. Tiede T., Treurniet W., Bitto R., et al. PEAQ — The ITU Standard for Objective Measurement of Perceived Audio Quality, *Journal of the Audio Engineering Society*, 2000, Vol. 48, No. 1, Pp. 3–29.

9. Hua G., Huang G., Shi Y. Q., et al. Twenty Years of Digital Audio Watermarking — A Comprehensive Review, *Signal Processing*, 2016, Vol. 128, Pp. 222–242.
DOI: 10.1016/j.sigpro.2016.04.005.

10. Agarwal N., Singh A. K., Singh P. K. Survey of Robust and Imperceptible Watermarking, *Multimedia Tools and Applications*, 2019, Vol. 78, Is. 7, Pp. 8603–8633.
DOI: 10.1007/s11042-018-7128-5.

11. Steinebach M., Petitcolas F. A. P., Raynal F., et al. StirMark Benchmark: Audio Watermarking Attacks, *Information Technology: Coding and Computing (ITCC 2001): Proceedings of the II International Conference, Las Vegas, NV, USA, April 02–04, 2001*. Institute of Electrical and Electronics Engineers, 2001, Pp. 49–54.
DOI: 10.1109/ITCC.2001.918764.

12. Ghasemzadeh H., Kayvanrad M. H. Comprehensive Review of Audio Steganalysis Methods, *IET Signal Processing*, 2018, Vol. 12, Is. 6, Pp. 673–687.
DOI: 10.1049/iet-spr.2016.0651.

13. Dittmann J., Hesse D. Network Based Intrusion Detection to Detect Steganographic Communication Channels: On the Example of Audio Data, *Proceedings of the IEEE 6th Workshop on Multimedia Signal Processing, Siena, Italy, September 29–October 01, 2014*. Institute of Electrical and Electronics Engineers, 2004, Pp. 343–346.
DOI: 10.1109/MMSP.2004.1436563.

14. Gofman M. V. Razrabotka modeli mnogokanalnoy audiostegosistemy na osnove markirovaniya tsifrovykh audiosignalov [Development of a Model of a Multi-Channel Audio Stegosystem Based on Digital Audio Watermarking], *Sovremennaya nauka: aktualnye problemy teorii i praktiki. Seriya: Estestvennye i Tekhnicheskie Nauki [Modern Science: Actual Problems of Theory and Practice. Series: Natural and Technical Sciences]*, 2021, No. 6, Pp. 78–83.
DOI: 10.37882/2223–2966.2021.06.11. (In Russian)

15. Gofman M. V. Obnaruzhenie markera v tsifrovom audiosignale avtorizovannym poluchatelem [Detection of Marker in Digital Audio Signal by Authorized Recipient], *Sovremennaya nauka: aktualnye problemy teorii i praktiki. Seriya: Estestvennye i Tekhnicheskie Nauki [Modern Science: Actual Problems of Theory and Practice. Series: Natural and Technical Sciences]*, 2021, No. 2, Pp. 45–50.
DOI: 10.37882/2223-2966.2021.02.09. (In Russian)

16. Gofman M. V., Kornienko A. A., Mironchikov E. T., Nikitin A. B. Tsifrovoe markirovanie audiosignalov dlya robastnoy skrytoy akusticheskoy svyazi cherez vozdushnyy audiokanal [Digital Watermarking of Audio Signals for Robust Hidden Audio Communication via Air Audio Channel], *Trudy SPIIRAN [SPIIRAS Proceedings]*, 2017, Is. 6 (55), Pp. 185–215. DOI: 10.15622/sp.55.8. (In Russian)

DOI: 10.24412/2413-2527-2022-331-61-67

Анализ устойчивости метода комбинированного маркирования цифровых аудиосигналов

д.т.н. А. А. Корниенко, к.т.н. М. В. Гофман, к.т.н. С. В. Корниенко
Петербургский государственный университет путей сообщения Императора Александра I
Санкт-Петербург, Россия
kaa.pgups@ya.ru, offmail3000@mail.ru, sv.diass99@ya.ru

Аннотация. **В статье анализируется устойчивость метода комбинированного маркирования цифровых аудиосигналов к преобразованиям и помехам, применяемого в аудиостего-системах с множественным (пространственным) входом и множественным (пространственным) выходом. В качестве основных характеристик устойчивости используются следующие: слуховая транспарентность (неслышимость маркера), устойчивость маркирования к сжатию с потерями (к MP3-преобразованиям), помехоустойчивость передачи через воздушный аудиоканал. Статья содержит количественные оценки показателей для указанных характеристик, полученные в результате натурных экспериментов и имитационного моделирования.**

Ключевые слова: **стеганография, цифровой аудиосигнал, комбинированное маркирование, оценка качества и устойчивости, воздушный аудиоканал, помехоустойчивость передачи.**

Литература

1. Hu, Y. Evaluation of Objective Quality Measures for Speech Enhancement / Y. Hu, P. C. Loizou // IEEE Transactions on Audio, Speech, and Language Processing. 2008. Vol. 16, No. 1. Pp. 229–238. DOI: 10.1109/TASL.2007.911054.

2. Detection of Audio Covert Channels Using Statistical Footprints of Hidden Messages / H. Özer, B. Sankur, N. Memon, İ. Avcıbaş // Digital Signal Processing. 2006. Vol. 16, Is. 4. Pp. 389–401. DOI: 10.1016/j.dsp.2005.12.001.

3. Kutter, M. Fair Benchmark for Image Watermarking Systems / M. Kutter, F. A. P. Petitcolas // Proceedings of the SPIE/IS&T Electronic Imaging 1999. Security and Watermarking of Multimedia Contents (San Jose, CA, USA, 23–29 January 1999) / P. W. Wong, E. J. Delp III (eds.). Proceedings of SPIE. 1999. Vol. 3657. Pp. 226–239.
DOI: 10.1117/12.344672.

4. ГОСТ Р 54711-2011. Звуковое вещание цифровое. Кодирование сигналов цифрового вещания с сокращением избыточности для передачи по цифровым каналам связи. MPEG-1 часть III (MPEG-1 audio) = Digital sound broadcasting. Coding of sound broadcasting signals with redundancy reduction for transfer on digital communication channels. MPEG-1 part III (MPEG-1 audio): национальный стандарт Российской Федерации: утвержден и введен в действие приказом Федерального агентства по техническому регулированию и метрологии от 13 декабря 2011 г. № 872-ст: дата введения 2013–07–01. — Москва: Стандартинформ, 2014. — 172 с.

5. ГОСТ Р 54712-2011. Звуковое вещание цифровое. Кодирование сигналов цифрового вещания с сокращением избыточности для передачи по цифровым каналам связи. MPEG-2 часть III (MPEG-2 audio) = Digital sound broadcasting. Coding of sound broadcasting signals with redundancy reduction for transfer on digital communication channels. MPEG-2 part III (MPEG-2 audio): национальный стандарт Российской Федерации: утвержден и введен в действие приказом Федерального агентства по техническому регулированию и метрологии от 13 декабря 2011 г. № 873-ст: дата введения 2013–07–01. — Москва: Стандартинформ, 2014. — 109 с.

6. Recommendation ITU-R BS.1196-8. Audio coding for digital broadcasting. Approved in 14.10.2019. — International Telecommunication Union, 2019. — 43 p.
URL: http://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1196-8-201910-I!!PDF-E.pdf (дата обращения 08.11.2021).

7. Recommendation ITU-R BS.1387-1. Method for objective measurements of perceived audio quality. Approved in 24.11.2001. — International Telecommunication Union, 2001. — 100 p. URL: http://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1387-1-200111-I!!PDF-E.pdf (дата обращения 08.11.2021).

8. PEAQ — The ITU Standard for Objective Measurement of Perceived Audio Quality / T. Tiede, W. Treurniet, R. Bitto, [et al.] // Journal of the Audio Engineering Society. 2000. Vol. 48, No. 1. Pp. 3–29.

9. Twenty Years of Digital Audio Watermarking — A Comprehensive Review / G. Hua, G. Huang, Y. Q. Shi, [et al.] // Signal Processing. 2016. Vol. 128. Pp. 222–242.
DOI: 10.1016/j.sigpro.2016.04.005.

10. Agarwal, N. Survey of Robust and Imperceptible Watermarking / N. Agarwal, A. K. Singh, P. K. Singh // Multimedia Tools and Applications. 2019. Vol. 78, Is. 7. Pp. 8603–8633. DOI: 10.1007/s11042-018-7128-5.

11. StirMark Benchmark: Audio Watermarking Attacks / M. Steinebach, F. A. P. Petitcolas, F. Raynal, [et al.] // Information Technology: Coding and Computing (ITCC 2001): Proceedings of the II International Conference (Las Vegas, NV, USA, 02–04 April 2001). — Institute of Electrical and Electronics Engineers, 2001. — Pp. 49–54.
DOI: 10.1109/ITCC.2001.918764.

12. Ghasemzadeh, H. Comprehensive Review of Audio Steganalysis Methods / H. Ghasemzadeh, M. H. Kayvanrad // IET Signal Processing. 2018. Vol. 12, Is. 6. Pp. 673–687. DOI: 10.1049/iet-spr.2016.0651.

13. Dittmann, J. Network Based Intrusion Detection to Detect Steganographic Communication Channels: On the Example of Audio Data / J. Dittmann, D. Hesse // Proceedings of the IEEE 6th Workshop on Multimedia Signal Processing (Siena, Italy, 29 September–01 October 2014). — Institute of Electrical and Electronics Engineers, 2004. — Pp. 343–346. DOI: 10.1109/MMSP.2004.1436563.

14. Гофман, М. В. Разработка модели многоканальной аудиостегосистемы на основе маркирования цифровых аудиосигналов // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и Технические Науки. 2021. № 6. С. 78–83.
DOI: 10.37882/2223-2966.2021.06.11.

15. Гофман, М. В. Обнаружение маркера в цифровом аудиосигнале авторизованным получателем // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и Технические Науки. 2021. № 2. С. 45–50. DOI: 10.37882/2223-2966.2021.02.09.

16. Цифровое маркирование аудиосигналов для робастной скрытой акустической связи через воздушный аудиоканал / М. В. Гофман, А. А. Корниенко, Е. Т. Мирончиков, А. Б. Никитин // Труды СПИИРАН. 2017. Вып. 6 (55). С. 185–215. DOI: 10.15622/sp.55.8.

DOI: 10.24412/2413-2527-2022-331-68-73

Features of Text Preprocessing for Performing Sentiment Analysis

N. E. Kosykh, I. A. Molodkin
Emperor Alexander I St. Petersburg
State Transport University
Saint Petersburg, Russia
nikitosagi@mail.ru, molodkin@pgups.ru

Grand PhD A. D. Khomonenko
Emperor Alexander I St. Petersburg
State Transport University,
Mozhaisky Military Space Academy
Saint Petersburg, Russia
khomon@mail.ru

Abstract. **The object of the research is the analysis of the sentiment of the Russian-language corpus of texts. The subject of the research is a comparison of the effectiveness of the approaches of preliminary text cleaning before sentiment analysis. The aim of the research is to develop a generalized method for preliminary data cleaning to create a neural network model. A distinctive feature of the proposed solutions is the use of modern and lightweight libraries for the possibility of preliminary preparation of a text for training with a neural network, and the hypothesis of using a truncated dictionary based on the assumption of data redundancy has been tested. The results obtained show the usefulness of the developed algorithm in terms of obtaining improved results in the learning process and indicate that, due to its versatility, it can be extrapolated for further use on other text data.**

Keywords: **mining, data analysis, sentiment analysis, neural networks, text processing.**

Introduction

Websites, social networks, micro-blogs are sources of a huge amount of user information for analysts, entrepreneurs and scientists trying to extract useful information from a large amount of user-generated text. Thanks to the feelings, emotions, images contained in the texts, one can draw certain conclusions to improve the quality of service or the quality of the services provided. For example, moderators of large Internet services and platforms can find solutions for timely tracking of violators by analyzing comments and user messages. For these reasons, a vast area of neural network technologies is aimed at developing systems for the automatic classification of texts (for example, aspect extraction, opinion analysis, sentiment analysis).

Recently, significant progress has been made in the development of various methods, starting with a rule base and machine learning to deep learning.

Despite the impressive results, the developed systems work exclusively with the specifically considered language, since the studied text corpora can have different dimensions, and also include non-linguistic elements that impede the analysis of data.

On the other hand, the nature of the content generated by users in social networks requires more careful preprocessing [1] before the stage of neural network analysis. In the field of tweet sentiment analysis, most scientific articles have used the approach of learning from scratch on corpuses consisting exclusively of thematic content, so that the models can eventually better understand local jargon with a special lexical and syntactic structure [2], as well as using abbreviations, punctuation marks and emoticons, etc. The use of such approaches imposes two restrictions: the first requires a large body of text for training in Russian, and the second — the need for significant resources and time to train models from scratch.

The approach described in this article allows testing the developed algorithm on a truncated corpus of texts and, if satisfactory results are obtained, use it when obtaining results on an integral corpus.

Vector transformation methods

The mathematical algorithms of future neural networks work with numbers, therefore, in order to use the mathematical apparatus for processing and analyzing text data, it is necessary first of all to transform words and sentences into numerical vectors, and preferably without losing semantic connection.

There are two main approaches to transforming a text document corpus [3] into a vector space:

1. Thematic modeling allows you to find statistical hidden patterns in the body of documents: latent semantic analysis (PLSA), latent Dirichlet distribution (LDA).

2. Distributive hypothesis — linguistic elements with similar distributions have similar meanings. This approach includes word embedding (WB) — a method that converts text tokens, most often represented by words, into dense small-sized vector representations trained on large unlabeled text corpora, where each token in the system is linked to another within the context. The Word2Vec model [4] proposed by Mikolov was an implementation of the word embedding method. The algorithm can encode text using two main approaches: skip-gramm or bag-of-word model [5]. The former predicts words in the surrounding context starting from the current word, while the latter predicts words based on the words surrounding it in the context. Also, the Global Vectors approach, is a distributed computing model, allowing you to code words faster on a large amount of data.

Ready-made vector representations obtained as a result of transformations allow you to compare texts and search for relationships, to perform classification and clustering of texts.

Natural language processing

Natural language processing is a branch of data science that deals with textual data that is used to analyze and solve business problems. Before using the collected data for the analysis and forecasting of user values, it is necessary to ensure their suitability, their preliminary processing is important [6].

The ultimate goal of processing data in natural language is to transfer this data to artificial intelligence (AI), which is capable of understanding human language. From a practical point of view, processing can be thought of as a set of methods and algorithms for extracting some useful information from text data.

Today, there is a fairly large range of practical tasks where natural language processing is required:

- machine translation of texts from foreign languages;
- automatic annotation of texts;
- classification of texts by categories (sentiment analysis, spam filter, rubrication);
- chat bots;
- recognition of given entities.

TEXT PRECLEANING PROCEDURE

Raw data obtained using the official API (http://developer.twitter.com/en/docs/api-reference-index) or taken from ready-made datasets contains heterogeneous information due to the nature of colloquial speech, for example: emoticons, hashtags, URLs, phone numbers, numbers without context, dates, etc. An example of some noisy messages is shown in Table 1.

Table 1

Example of background information

Интернет уже 5 день не работает :(
RT @supreme_fm: Хочу тату :(http://t.co/2LuWQRS5H3
Почему многим нравится что я пишу? Я же ничего крутого не записываю :(А тем временем бабушка оладушки жарит :3 но это уже другая история
Мне нравится готовить, но не нравится, что я часто режусь ножом или чистилкой для картошки :(#печальбеда
Объелась я золотистой жареной картошки и квашеной капусты с красным луком под маслом. Вкусно, но зачем так много ;(
Я скоро уничтожу свой телефон полностью, опять его в лужу грохнула (

This work uses a set of procedures to transform chaotic information into a more traditional and understandable form. The sequence of actions for each procedure does not depend on specific language features [7]. Actions are usually based on linguistic rules contained in conversion tables or regular expressions.

The data must be preprocessed to perform mining tasks. First, you need to perform preprocessing of the text, which includes:

1. Removing punctuation marks and specials characters. Punctuation marks and other special characters should be removed to eliminate inaccuracies in determining polarity. In some cases, signs can stick together with words and be incorrectly identified in dictionaries, which makes them inaccessible for analysis.

2. Removing URL's. Typically, URLs are not used to analyze sentiment in informal texts. For example, consider the following sentence «I went to the site funny-jokes.rf, because I'm bored», which is negative, however, due to the presence of the word «funny» it can become neutral, and this is a false-positive result. To avoid this kind of collision, you need to remove the URLs beforehand.

3. Removing stop words: conjunctions, interjections. Words like what, like, like, etc. will not affect the polarity of expressions; to reduce the computational complexity, it is worth getting rid of them. Python contains stoplists for different languages in the NLTK module.

4. Removing forwarded references. Such references are usually marked with the name of the original author plus the letter combination RT. In a thematic analysis, additional information for statistics can be obtained from this data, but in sentiment analysis it is redundant (Fig. 1).

5. Breakdown of offers into tokens. Typically, this is the process of separating individual words into an array, separated by punctuation marks and whitespace.

6. Normalization of words. Normalization is the process of bringing words to their standard morphological form, which is implemented by the stemming method — the method of eliminating the endings of words, or lemmatization — bringing a word to its normal (dictionary) form («is» — «to be», «written» — «to write») [8].

7. Removing number sequences. The use of numbers in the analysis can increase the size of the studied vocabulary, which will complicate the process of training the model on unnecessary data.

To solve the problem, a universal pipeline (method) was developed, which can be used for preliminary text cleaning in any of the projects under study. It includes all of the above stages of processing textual data written in the form of regular expressions, as well as splitting into tokens and applying lemmatization or stemming methods to choose from (Fig. 1).

```
def preprocess_text(text):
    text = text.lower().replace("ё", "е")
    text = re.sub('((www\.[^\s]+)|(https?://[^\s]+))', 'url', text)
    text = re.sub('@[^\s]+', '', text)
    text = re.sub('(?:(?:\d+,?)+(?:\.?\d+)?)',' ',text)
    text = re.sub('[^a-zA-Za-яА-Я1-9]+', ' ', text)
    text = re.sub(' +', ' ', text)
    text = re.sub('[a-zA-Z]+', '', text)
    text = re.sub(" \d+", " ", text)

    tokens = []
    for token in text.split():
        if token and token not in stop:
            token = token.strip()
            token = stemmer.stem(token)
            token = morph.normal_forms(token)[0]
            tokens.append(token)

    text = ' '.join(tokens)

    return text.strip()
```

Fig. 1. Data preprocessing pipeline

It is also worth noting that the processing time by truncating the endings is several times faster than searching for a dictionary form. If you sacrifice quality for speed you can opt for the Snowball algorithm otherwise it is better to use a lemmatizer.

COMPARISON OF TEXT NORMALIZATION ALGORITHMS

There are two types of stemmers in the NLTK library — Porter's stemmer and Snowball [9], the second, in turn, is an improved version of the first, and we will test it on an array of sentences. After processing as a basis, we get a lot of non-dictionary words. Unlike the previous approach, lemmatization shortens a word to its dictionary form and in most cases is semantically complete (Table 2).

Table 2

Comparison of text processing by stemming and lemmatization

Stemming (word processing)	Lemmatization (word processing)
'получа осторожн котлетк перекладыва др посуд'	['получаться осторожно котлетка перекладывать др посуда'
'единствен расстраива бал невозможн сход кинотеатр друз интересн фильм'	'единственный расстраивать бали невозможность сходить кинотеатр друг интересный фильм'
'базарн сегодн спрашива суш парен вес кг сказа пор лол'	'базарнов сегодня спрашивать сушиться парень весить кг сказать пора лола'
'маловат получа незлобин'	'маловатый получаться незлобин',
'любл всем тво истерик солнышк'	'любить весь твой истерик солнышко']

The next section displays experimental data run through the NLTK Snowball and MorphAnalyzer modules. The strengths and weaknesses of each approach are identified.

Vector text representation

In most mining tasks, after the text cleanup stage, we create a dictionary of index mappings, so that frequently occurring words are assigned a lower index, and then the word appears less frequently in the text corpus. One trivial solution would be to use a combination of the split () and strip () methods, splitting the input data stream into individual words and writing them to an array. Then, using the Tokenize module, we encode a random array of words into a vector, where the total TF-IDF coefficient is calculated for each word, and then we update the dictionary based on the list of received tokens by calling the fit_on_texts () method.

Results of statistical tests of algorithms for one dataset

Additionally, let us analyze the length of the records loaded into the array after the stage of splitting into tokens (Fig. 2). The figure shows a histogram that shows the distribution between the number of records and their corresponding length (by words).

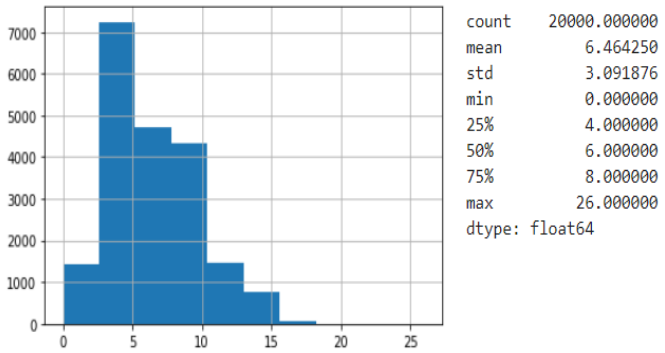

Fig. 2. Sampling before deleting unnecessary records

The average length of a message in the corpus is 6 words. A number of records have zero length. The presence of these values in the data array will not have any effect on the learning process of the model [10], it will only become redundant information. We also make a hypothesis that in the presence of sentences from 1 token there is a possibility of incorrect interpretation of the polarity, it follows that at the stage of preliminary data preparation, a check for the length of the loaded records should be added.

Based on the condition specified below, we check the initial data set and write into the new array only those sentences that, after the preprocessing stage, have more than two words:

```
sentences1 = []
labels1 = []
for id,item in enumerate(sentences):
  if len(item.split()) >= 2:
     sentences1.append(item)
     labels1.append(labels[id])

assert len(sentences1) == len(labels1)
```

After preparing a new dataset, let's build a histogram (Fig. 3) to check the condition's performance.

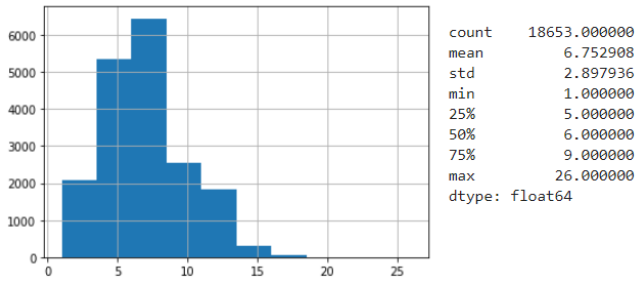

Fig. 3. Fetch after deleting unnecessary records

As you can see from the graph, words and their corresponding labels were removed, where there were no tokens and tokens less than three. Within a small number of test records, this may seem like an unnecessary step, but on a large body of texts, this can lead to dirty statistics. A sample of 20 000 records was taken as a basis, after deleting records that do not meet the condition, 18 653 remained. It is necessary to establish whether there is a difference between a more extensive dictionary or a better quality.

The next step involves building a neural network model based on long-term neural memory LSTM [11]. Its advantages include fast scanning and acceptable learning speed. However, it is not suitable for repeated training on different parameters due to its peculiarity to remember the states in past sessions. As a temporary solution to reset the weights, it is necessary to compile the trainable model each time.

The results of cross-validation in terms of accuracy for models trained using different text processing techniques are shown in Table 3. In the first case, the source text was divided into tokens, each of which was reduced to a stem or to a dictionary form, depending on the user's choice. These tokens became the vocabulary for future network learning. In another case, all the same stages except for those *n*-grams that contained less than 3 words.

Table 3

Results of cross-validation in training models

Data set 20000 words	Epoch No. 6	Epoch No. 7	Epoch No. 8	Epoch No. 9	Epoch No. 10
***n*-gramma >=0+lemmatization**	0.7601	0.7962	0.8192	0.8369	0.8569
***n*-gramma >=0+stemming**	0.7734	0.8035	0.8215	0.8443	0.8582
***n*-gramma >2+lemmatization**	0.7852	0.8097	0.8278	0.8497	0.8638
***n*-gramma >2+stemming**	0.7844	0.7844	0.8309	0.8480	0.8566

By the tenth epoch of learning, an approach that includes skipping *n*-grams (where $n <= 2$) and lemmatization at the preprocessing stage gives an accuracy advantage over the rest of the techniques under consideration.

Conclusion

Our studies have shown that training a neural network on a training dataset containing sentences longer than 2 tokens and processed by the lemmatization method gives a small increase in the training accuracy at higher epochs compared to the original (unprocessed) dataset, and also has a more intensively decreasing function losses (Fig. 4).

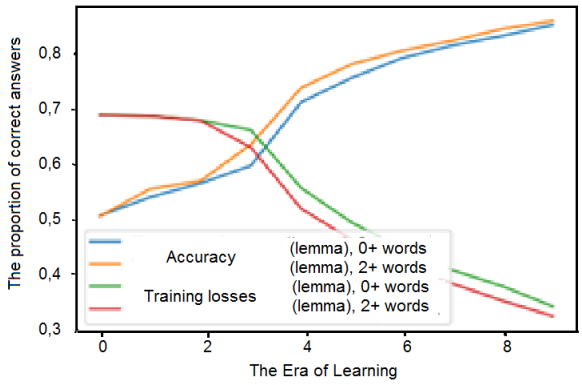

Fig. 4. Data processing + lemmatization

This difference is due to a more accurate distribution of weight coefficients [12] between neurons in the process of training the network. If we compare the use of an abbreviated or full vocabulary on stemming, then the difference becomes insignificant towards the achievement of the later eras of learning (Fig. 5).

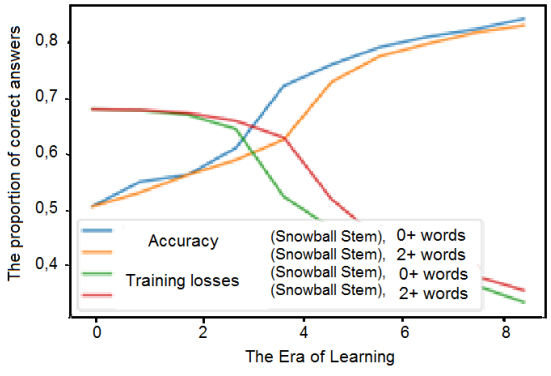

Fig. 5. Data processing + stemming

It is proposed to further study natural language processing algorithms on an existing dataset with subsequent *n*-gram network training, as well as improve the preliminary text cleaning algorithm for the following: automatic replacement of toxic words or selection of appropriate synonyms, replacement of double letters with single letters, replacement of jargon words with commonly used ones, as well as encoding emoticons and converting them to a common number format to improve accuracy and reduce losses in the process of training the model.

References

1. Gan N. Application of Algorithms for Natural Language Processing in IT-Monitoring with Python Libraries, *Towards Data Science*. Published online at 14 August 2021. Available at: http://towardsdatascience.com/application-of-algorithms-for-natural-language-processing-in-it-monitoring-with-python-libraries-57eddc45c4ac (accessed 04 Nov 2021).

2. Hemalatha I., Varma G. P. S., Govardhan A. Preprocessing the Informal Text for Efficient Sentiment Analysis, *International Journal of Emerging Trends and Technology in Computer Science*, 2012, Vol. 1, Is. 2, Pp. 58–61.

3. Lychenko N. M., Sorokovaja A. V. Sravnenie effektivnosti metodov vektornogo predstavleniya slov dlya opredeleniya tonalnosti tekstov [Comparison of Effectiveness of Word Representations Methods in Vector Space for the Text Sentiment Analysis], *Matematicheskie struktury i modelirovanie [Mathematical Structures and Modeling]*, 2019, No. 4 (52), Pp. 97–110. DOI: 10.24147/2222-8772.2019.4.97-110. (In Russian)

4. Church K. W. Word2Vec, *Natural Language Engineering*, 2017, Vol. 23, Is. 1, Pp. 155–162. DOI: 10.1017/S1351324916000334.

5. Paramonov I. Yu., Smagin V. A., Kosykh N. E., Khomonenko A. D. Metody i modeli issledovaniya slozhnykh sistem i obrabotki bolshikh dannykh: Monografiya [Methods and models for studying of complex systems and big data processing: Monograph]. Saint Petersburg, LAN Publishing House, 2020, 236 p. (In Russian)

6. Krivaltsevich E. V. Obrabotka estestvennogo yazyka (Natural Language Processing) pri ispolzovanii tekhnologii NLTK (Natural Language Toolkit) na baze yazyka programmirovaniya Python [Natural Language Processing using NLTK (Natural Language Toolkit) technology based on the Python programming language], *Ideas. Search. Decisions: Collection of Articles of the VII International Scientific and Practical Conference, Minsk, Belarus, November 25, 2014*. Minsk, Belarusian State University, 2015, Pp. 41–51. (In Russian)

7. Predvaritelnaya obrabotka dannykh v Python [Data preprocessing in Python], Machinelearningmastery.ru — Mashinnoe obuchenie, neyronnye seti, iskusstvennyy intellect [Machinelearningmastery.ru — Machine Learning, Neural Networks, Artificial Intelligence]. Published online at August 23, 2019. URL: http://www.machinelearningmastery.ru/data-preprocessing-in-python-b52b652e37d5 (accessed 05 Nov 2021). (In Russian)

8. Gupta I., Joshi N. Tweet Normalization: A Knowledge Based Approach, *Proceedings of the 2017 IEEE International Conference on Infocom Technologies and Unmanned Systems (ICTUS) (Trends and Future Directions), Dubai, United Arab Emirates, December 18–20, 2017*. Institute of Electrical and Electronics Engineers, 2017, Pp. 157–162. DOI: 10.1109/ICTUS.2017.8285996.

9. Malik U. Python for NLP: Tokenization, Stemming, and Lemmatization with SpaCy Library, *Stack Abuse*. Available at: http://stackabuse.com/python-for-nlp-tokenization-stemming-and-lemmatization-with-spacy-library (accessed 10 Nov 2021).

10. Morrissey M., Wasser L., Diaz J., Palomino J. Analyze the Sentiment of Tweets from Twitter Data and Tweepy in Python, Earth Data Science — Earth Lab. Last update 11 September 2020. Available at: http://www.earthdatascience.org/courses/use-data-open-source-python/intro-to-apis/analyze-tweet-sentiment-in-python (accessed 29 Nov 2021).

11. Huang Z., Hu W., Yu K. Bidirectional LSTM-CRF Models for Sequence Tagging, *arXiv*, 2015, Vol. 1508.01991, 10 p. DOI: 10.48550/arXiv.1508.01991.

12. Korobov V. B. Sravnitelnyy analiz metodov opredeleniya vesovykh koeffitsientov «vliyayushchikh faktorov» [Comparative Analysis of Methods for Determining the Weight Coefficients of «Influencing Factors»], *Sotsiologiya: metodologiya, metody, matematicheskie modeli (Sotsiologiya: 4M) [Sociology: Methodology, Methods, Mathematical Modeling (Sociology: 4M)]*, 2005, No. 20, Pp. 54–73.

DOI: 10.24412/2413-2527-2022-331-68-73

Особенности предварительной обработки текстовых данных при анализе тональности текстов

Н. Е. Косых, И. А. Молодкин
Петербургский государственный университет
путей сообщения Императора Александра I
Санкт-Петербург, Россия
nikitosagi@mail.ru, molodkin@pgups.ru

д.т.н. А. Д. Хомоненко
Петербургский государственный университет
путей сообщения Императора Александра I,
Военно-космическая академия имени А. Ф. Можайского
Санкт-Петербург, Россия
khomon@mail.ru

Аннотация. **Объект исследования — анализ тональности русскоязычного корпуса текстов. Предмет исследования — сравнение эффективности подходов предварительной очистки текста перед анализом тональности. Цель исследования — разработка обобщенного метода предварительной очистки данных для создания модели нейросети. Отличительной чертой предложенных решений является использование современных и легковесных библиотек для возможности предварительной подготовке текста к обучению нейросетью; также апробирована гипотеза использования усеченного словаря на основе предположения об избыточности данных. Полученные результаты показывают полезность разработанного алгоритма с точки зрения получения улучшенных результатов в процессе обучения и указывают на то, что благодаря своей универсальности он может быть экстраполирован для дальнейшего использования на других текстовых данных.**

Ключевые слова: **интеллектуальный анализ, анализ данных, сентиментный анализ, нейронные сети, обработка текста.**

Литература

1. Gan, N. Application of Algorithms for Natural Language Processing in IT-Monitoring with Python Libraries // Towards Data Science. — 2021. — 14 August. URL: http://towardsdatascience.com/application-of-algorithms-for-natural-language-processing-in-it-monitoring-with-python-libraries-57eddc45c4ac (дата обращения 04.11.2021).

2. Hemalatha, I. Preprocessing the Informal Text for Efficient Sentiment Analysis / I. Hemalatha, G. P. S. Varma, A. Govardhan // International Journal of Emerging Trends and Technology in Computer Science. 2012. Vol. 1, Is. 2. Pp. 58–61.

3. Лыченко, Н. М. Сравнение эффективности методов векторного представления слов для определения тональности текстов / Н. М. Лыченко, А. В. Сороковая // Математические структуры и моделирование. 2019. № 4 (52). С. 97–110. DOI: 10.24147/2222-8772.2019.4.97-110.

4. Church, K. W. Word2Vec // Natural Language Engineering. 2017. Vol. 23, Is. 1. Pp. 155–162. DOI: 10.1017/S1351324916000334.

5. Методы и модели исследования сложных систем и обработки больших данных: Монография / И. Ю. Парамонов, В. А. Смагин, Н. Е. Косых, А. Д. Хомоненко; под ред. В. А. Смагина, А. Д. Хомоненко. — Санкт-Петербург: Лань, 2020. — 236 с. — (Учебники для вызов. Специальная литература).

6. Кривальцевич, Е. В. Обработка естественного языка (Natural Language Processing) при использовании технологии NLTK (Natural Language Toolkit) на базе языка программирования Python // Идеи. Поиски. Решения: Сборник статей VII Международной научно-практической конференции (Минск, Беларусь, 25 ноября 2014 г.). — Минск: Белорусский гос. ун-т, 2015. — С. 41–51.

7. Предварительная обработка данных в Python // Machinelearningmastery.ru — Машинное обучение, нейронные сети, искусственный интеллект. — 2019. — 23 августа. URL: http://www.machinelearningmastery.ru/data-preprocessing-in-python-b52b652e37d5 (дата обращения 05.11.2021).

8. Gupta, I. Tweet Normalization: A Knowledge Based Approach / I. Gupta, N. Joshi // Proceedings of the 2017 International Conference on Infocom Technologies and Unmanned Systems (ICTUS) (Trends and Future Directions), (Dubai, United Arab Emirates, 18–20 December 2017). — Institute of Electrical and Electronics Engineers, 2017. — Pp. 157–162. DOI: 10.1109/ICTUS.2017.8285996.

9. Malik, U. Python for NLP: Tokenization, Stemming, and Lemmatization with SpaCy Library // Stack Abuse. URL: http://stackabuse.com/python-for-nlp-tokenization-stemming-and-lemmatization-with-spacy-library (дата обращения 10.11.2021).

10. Analyze the Sentiment of Tweets from Twitter Data and Tweepy in Python / M. Morrissey, L. Wasser, J. Diaz, J. Palomino // Earth Data Science — Earth Lab. Last update 11 September 2020. URL: http://www.earthdatascience.org/courses/use-data-open-source-python/intro-to-apis/analyze-tweet-sentiment-in-python (дата обращения: 29.11.2021).

11. Huang, Z. Bidirectional LSTM-CRF Models for Sequence Tagging / Z. Huang, W. Hu, K. Yu // arXiv. 2015. Vol. 1508.01991. 10 p. DOI: 10.48550/arXiv.1508.01991.

12. Коробов, В. Б. Сравнительный анализ методов определения весовых коэффициентов «влияющих факторов» // Социология: методология, методы, математические модели (Социология: 4М). 2005. № 20. С. 54–73.

DOI: 10.24412/2413-2527-2022-331-74-81

Russian version of the article © V. A. Smagin, V. P. Bubnov is published
in *Automation on Transport*, 2021, Vol. 7, No. 4, Pp. 617–630.
DOI: 10.20295/2412-9186-2021-7-4-617-630.

A Few Remarks About the Most Important Element of Metrology — Person

Grand PhD V. A. Smagin
International Informatization Academy
Saint Petersburg, Russia
va_smagin@mail.ru

Grand PhD V. P. Bubnov
Emperor Alexander I St. Petersburg State Transport University
Saint Petersburg, Russia
bubnov1950@yandex.ru

Abstract. **The article considers a person as an element of metrology. His functional duties and specific actions are not taken into account. The metrologist is presented as a two-phase system, including two stages of the life cycle, the first cycle of which is a phase of concentration - work for effect, the second cycle - a phase of chaos, consisting in the restoration of spent energy to continue the first phase. A formal model of a person-metrologist is presented. The duration of the inter-verification period, optimal in terms of the availability factor, and the average number of object repairs per one year have been determined. With the help of the real model the average age of the person operator in terms of maximum availability, with and without preventive periods has been determined. In the formal and real person-operator models, the distribution of person lifespan is determined by the statistically extreme Weibull distribution law. From the formal point of view the chaos environment is characterized by a probability distribution function opposite to the concentration environment distribution function according to P. Levy.**

Keywords: **metrologist, metrology, two-phase system, concentration environment, chaos environment, average failure rate, concentration function, concentration assurance.**

INTRODUCTION

Metrology traces its history back to ancient times, but only in the XX century did it become one of the main fundamental sciences. Metrology is divided into three main sections. Theoretical, or fundamental — considers general theoretical problems (development of the theory and problems of measuring physical quantities, their units, measurement methods). Applied — studies the issues of practical application of theoretical metrology developments. She is in charge of all issues of metrological support. Legislative — establishes mandatory technical and legal requirements for the use of units of physical quantities, methods and measuring instruments.

It is appropriate to raise the question of the main element of the science of metrology - the metrologist. In foreign and domestic literature, many works are devoted to the person operator [1–10]. Basically, they deal with aspects of the interaction of a person and a team with hardware and software systems. However, in our opinion, studying the elements of metrology (such as standards, measuring instruments) and their practical significance, we have the right to consider the most modern metrologist from a technical point of view as a metrological element, and in a broader sense - as a living metrological system.

The aim of the article is to study the person metrologist as a metrological element.

FORMAL MODEL OF A PERSON METROLOGIST

First, let us consider a prototype of a model — a technical model — using the example of «average failure rate and availability factor of a measuring device, taking into account its metrological checks» [11]. To assess the reliability of restored objects, the reliability indicator is used — the average failure rate [12].

In this article, the average failure rate of recoverable objects is considered under the condition that periodic preventive maintenance is carried out at the facilities. It is assumed that during restoration and prevention, the object is restored completely to its original state.

The article posed the task of determining the average failure rate of an object on which state checks can be periodically carried out. With them, the object could be in a functional state, but requiring updating, for example, by adjusting its parameters. When a failure was detected, the object was replaced with a new one. An integral equation was derived for the corresponding average failure rate of the object, and its properties were investigated. The purpose of the article was to establish the first connection between the reliability indicators of hardware and software objects with metrological indicators that make up a necessary part of ensuring the quality of objects.

The following designations were adopted: $\omega(t)$ — average failure rate; $a(t)$ — probability density of time to failure; $Q(t)$, $P(t)$ — likelihood of failure and the likelihood of failure-free operation; $U(t)$ — function of time distribution of the beginning of verification; $v(t)$ — density of the probability of the duration of the verification and adjustment of the parameters of the object; $g(t)$ — distribution density of the recovery time of an object after a failure; τ — moment of the appointment of the first verification; θ — moment before the first failure occurs.

The average failure rate was determined by the sum of three components corresponding to the following inconsistent events:

- there was exactly one refusal of the object in time t, provided that verification was not scheduled during this time;
- there were several object failures within t time, provided that the first failure occurred before the appointment of the first verification;

- there were several object failures within t time, provided that the first verification was scheduled before the first failure occurred.

Then the expression for the average failure rate took the form:

$$\omega(t) = a(t) \times [1 - U(t)] + \\ + \int_0^t [1 - U(t)] \times a(\tau) \times \omega(t - \tau) d\tau + \\ + \int_0^t [1 - Q(\tau)] \times \int_0^{t-\tau} v(\theta) \times \omega(t - \tau - \theta) d\theta dU(\tau). \quad (1)$$

Expression (1) is obtained under the condition that after refusal and verification, the object is replaced by a serviceable (new) one instantly. The control over the state of the object's elements is perfect. For (1), the Laplace transform of the mean frequency is determined:

$$\omega^*(s) = \frac{a^*(s)}{1 - a^*(s) - b^*(s)v^*(s)}, \quad (2)$$

where

$$a^*(s) = \int_0^\infty a(z)[1 - U(z)]e^{-sz} dz;$$

$$b^*(s) = \int_0^\infty [1 - Q(z)]e^{-sz} dU(z). \quad (3)$$

The image of the probability density of the duration of the verification and adjustment of the object parameters was represented by the sum of two random components, therefore the Laplace image is equal to:

$$v^*(s) = u^*(s) \times r^*(s), \quad (4)$$

where $u^*(s)$, $r^*(s)$ — are the images of the time densities of verification and adjustment. In practice, checks on objects are carried out regularly, so it makes sense to consider a degenerate distribution as $U(t)$, i. e.

$$U(t) = \begin{cases} 0, t \lhd T \\ 1, t \geq T \end{cases}, \quad (5)$$

where T — the period between adjacent verifications.

From expression (2), taking into account formulas (3)–(5), under the condition of long-term operation of the object, a steady-state value of the average failure rate was obtained:

$$\omega(\infty, T) = \frac{Q(T)}{\int_0^T P(z)dz + t_{ur} P(T)}, \quad (6)$$

where t_{ur} — average duration of one verification with parameter adjustment.

Reasoning in a similar way, we got $\omega(t)$, $\omega^*(s)$, $\omega(\infty, T)$ for a situation when the restoration of an object was not performed instantly, but after a random time:

$$\omega(t) = a(t)[1 - U(t)] + \\ + \int_0^t [1 - U(\tau)] \int_0^{t-\tau} g(\theta)\omega(t - \tau - \theta) d\theta a(\tau) d\tau + \\ + \int_0^t [1 - Q(\tau)] \int_0^{t-\tau} v(\theta)\omega(t - \tau - \theta) d\theta dU(\tau); \quad (7)$$

$$\omega^*(s) = \frac{a^*(s)}{1 - a^*(s)g^*(s) - b^*(s)v^*(s)};$$

$$\omega(\infty, T) = \frac{Q(T)}{\int_0^T P(z)dz + t_r Q(T) + t_{ur} P(T)},$$

where t_r — is the average recovery time of the object.

From expressions (6) and (7), provided that verification is not performed ($T \to \infty$), there follow the known special cases of stationary values of the average frequency:

$$\omega(\infty, \infty) = \frac{1}{T_{av}},$$

$$\omega(\infty, \infty) = \frac{1}{t_{av} + t_r},$$

where t_{av} — is the average time of no-failure operation of the object.

USE CASE $\omega(\infty, T)$

It is required to determine the optimal in terms of availability factor the duration of the calibration period T_0 and the average number of repairs n_p of the object per one year, if the distribution law of the object's operation time to failure is Weibull's law with the parameter values:

$$\lambda_0 = 1 \times 10^{-5} \text{ h}^{-k}, \; k = 2.5.$$

The average time to repair an object after a failure is $t_r = 10$ h, and the average duration of verification and adjustment of the object's parameters is $t_{ur} = 2$ h. It is easy to verify that the stationary value of the facility availability factor is:

$$K_G = K_G(\infty, T) = \frac{\int_0^T P(z)dz}{\int_0^T P(z)dz + t_v Q(T) + t_{ur} P(T)}, \quad (8)$$

where $P(t) = e^{-\lambda_0 t^k}$, $Q(t) = 1 - P(t)$.

The quantity T_0, leading to the maximum (7) satisfies the equation:

$$\frac{t_r}{t_r - t_{ur}} = \lambda(T_0) \int_0^{T_0} P(z)dz + P(T_0), \quad (9)$$

in which $\lambda(t)$ — is the failure rate of the object.

It should be noted that expressions (8) and (9) coincide with expressions obtained in [13] in a different way.

The results of calculations according to formulas (7) and (8) are shown in Figure 1. The maximum value of $K_G = 0{,}95$ is achieved at $T_0 \approx 50$ h. The average failure rate of $T_0 \approx 50$ h corresponds to the average frequency of failures $\omega(\infty, T_0) \approx 0.00351$ per hour. The average time of no-failure operation of the facility without verification is $T_{av} \approx 89$ h, and with their verification is $T_{av} \approx 285$ h. The average expected number of repairs of the facility during the year without carrying out checks $n_p \approx 100$, and with their carrying out $n_p \approx 31$. The total operating time of the facility during the year increases by an average of half a month.

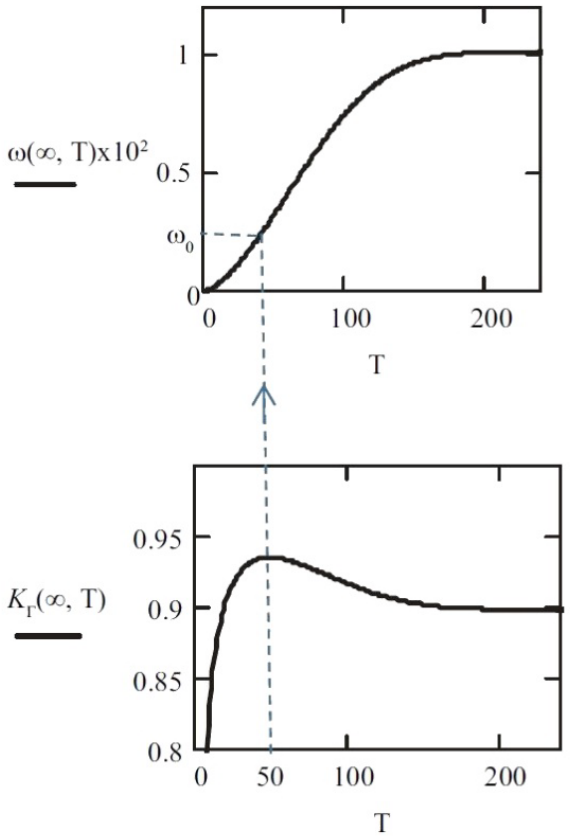

Fig. 1. Dependences of the average failure rate ω and the availability factor of the object K_G on the duration of the period T between checks

Let us consider a rather small inverse problem of metrology: how, for a given availability factor, to determine the requirement for the value of the average duration of verification and regulation of an object? For this, from expression (8) we find the value t_{ur}. It will be represented by the expression:

$$t_{ur} = \frac{\int_0^T P(z)dz}{P(T)} \times \frac{1 - K_G}{K_G} - t_r \frac{Q(T)}{P(T)}. \qquad (10)$$

It is also possible to represent the solution of equation (8) with respect to t_{ur} in the form

$$t_{ur} = t_r \left(1 - \frac{1}{\lambda(T_0) \int_0^{T_0} P(z)dz + P(T_0)}\right).$$

Graphical representation (10) is shown in Figure 2. It follows that the inflection point coordinate $t_{ur}(50) = 2.055$ h corresponds to the optimal solution.

We have considered an example of a typical operation of a complex technical system. Let's make one remark that this example is not typical for a person — a measuring device.

The remark concerns the fact that the standard deviation (RMSD) is not typical for the life of a person who is in both normal and stressful situations. The calculation shows that with an average device lifetime of 89 hours, not years, the standard deviation is 38 028 hours. Replacing the unit «hour» with the unit «year» clarifies the remark made. Therefore, based on the Weibull distribution, we will select an example that is real for a person — a measuring device.

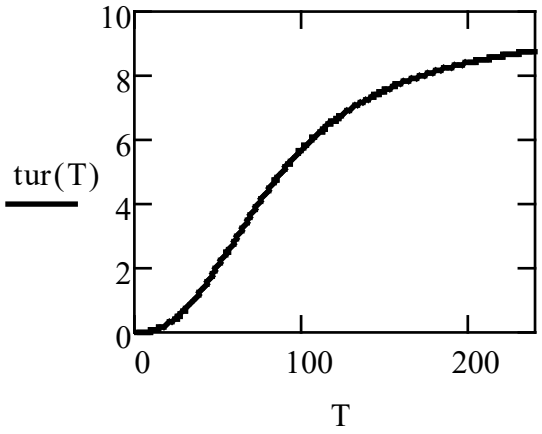

Fig. 2. Dependence of the duration of one verification with the adjustment of parameters on the duration of the period T between verifications

REAL MODEL OF PERSON METROLOGIST

So, we define the distribution of the life time of a person metrologist by the Weibull distribution law $F(x) = e^{-\lambda_0 \times x^k}$, extreme in statistics. Let's select the values of the following parameters: $\lambda_0 = 1 \times 10^{-5}$ year, $k = 2.8$. Then the average age of the metrologist is $\nu_1 = 54.4$ year; the second initial moment is $\nu_2 = 3\,307.4$ years and the standard deviation of the age is $\sigma = 21.0$ year. This means that the period of his career is in the range from 33.4 to 75.4 years. We also take the value of the average duration of recovery of a person metrologist after an illness $t_r = 3$ months and the value of the average duration of his preventive maintenance during the period of work $t_p = 1$ month.

Applying the calculations of the previous article material, we get the following graphical results. Figure 3 shows a graph of the readiness function, and Figure 4 shows a graph of the average failure rate (disease) of a person metrologist. Figure 5 shows the probability of continuous work of a metrologist during his life.

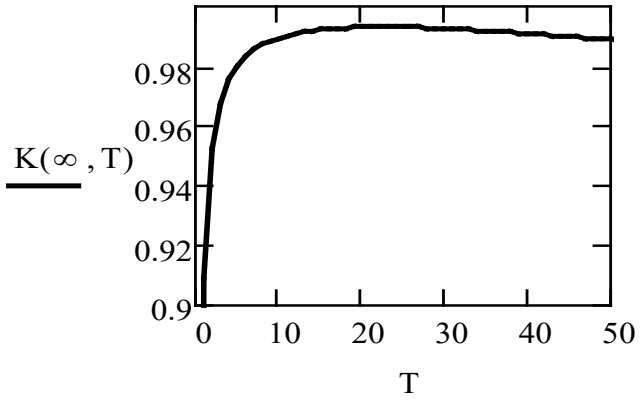

Fig. 3. Ready function

Years are indicated along the abscissas of all graphs. Based on the graph in Figure 3, it can be argued that the maximum availability factor $K_G(\infty, 25) = 0.993$ is achieved with the value of the optimal periodic preventive maintenance $T_0 = 25$ years. In this case, the average frequency of failures (diseases) is

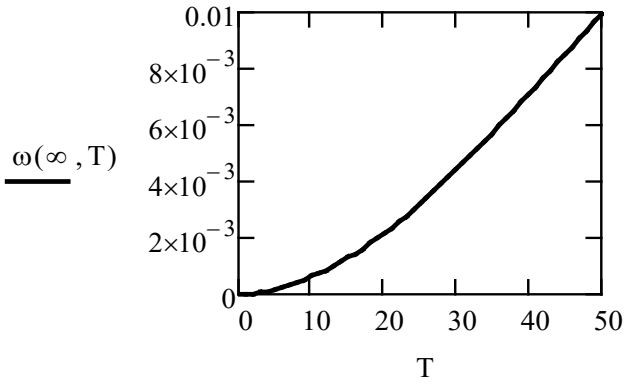

Fig. 4. Dependence of the average frequency of failures (diseases) of the metrologist on the duration of the period *T* between verifications

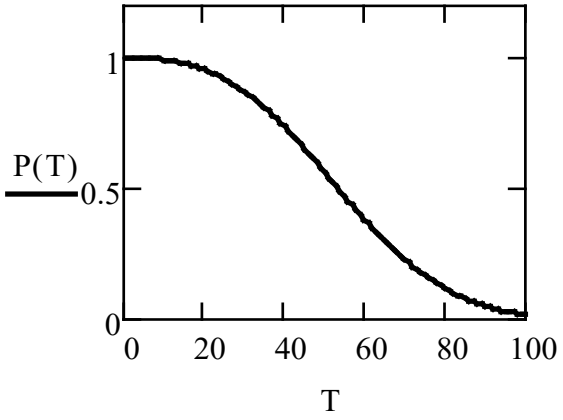

Fig. 5. Probability of continuous trouble-free operation of the metrologist depending on the lifetime

$\omega(\infty, 25) = 3{,}212 \times 10^{-3}$ year^{-1}. Figure 5 shows a graph of the probability of continuous trouble-free operation of a metrologist depending on the number of years. The results of calculations according to formulas (7) and (8) in relation to the initial data of the model of a person operator, presented in Figures 3 and 4, lead to the following additional quantitative data: if the average age of a person metrologist, excluding preventive maintenance, is $\nu_1 = 54.4$ year, of service it will be $1/\omega(\infty, 25)$ year^{-1} $= 311.333$ year. The average number of required restorations in one year will be $1/54.277 = 0.018$ once. Taking into account preventive maintenance with an optimal frequency of $1/311.333 = 3.212 \times 10^{-3}$ within one year, it will be $0.018 \times 0.25 = 4.5 \times 10^{-3}$ years. Time costs per year for treatment will be on average

$$3.212 \times 10^{-3} \times 1/12 = 2.677 \times 10^{-4} \text{ years.}$$

Thus, the gain in time to restore health per year will be on average $(4.5 \times 10^{-3})/(2.677 \times 10^{-4}) = 16.81$ times due to the periodic preventive maintenance of a person metrologist. These conclusions apply only to those setting data of the example that we have given.

On the relationship between the distribution functions of concentration and chaos

Unlike technical and software systems and their elements — measuring devices, which can function when working as intended or be restored, the measuring system «person metrologist» can also be in the two named states, but these paired states are different. If the first systems are restored in their inherent technical environment, then the second, according to their states, can be attributed to systems of a seasonal type. After they leave the state of work in a technical environment, they find themselves in a different environment that differs from the first in another property. If the systems of the first type are associated with the phenomenon of concentration of the produced product, with the receipt of a significant effect, then the systems of the second type, on the contrary, are associated with liberation from the process of concentration of the product, but on the contrary, they are removed from this need for production, go into a state of rest, relaxation, organized indifference, roughly speaking — controlled chaos.

Here, chaos is understood as the acquisition of such properties, without which it is impossible to return to systems that have the properties of the first systems. Namely, if the systems periodically change the indicated properties of the two named systems, but on the whole pass from one type to another and vice versa, then we call such systems seasonal. The seasonality of systems is the basis of the life cycle, their dialectical unity. Thanks to it, the economic growth of the state, progress in civilization, science, the welfare of the nation, its defense capability and others are possible.

An abstract example

The subject, functioning in the concentration mode with parameter t_k and duration x_k, has developed N. M. Sedyakin's resource r_k [14]. Subject then entered chaos mode with parameter t_x to restore the consumed resource. How long should it be in chaos mode, x_x, in order to restore the spent resource in concentration mode?

Formally, the modes are represented, as follows from the material presented [15–16]:

$$Q_F(x_k) = \max_{t_k}\big(F(t_k) - F(t_k + x_k + 0)\big),$$

$$W_F(x_x) = \max_{t_x}\big(F(t_x) - F(t_x + x_x + 0)\big),$$

where Q_F is the concentration distribution function, W_F is the chaos distribution function, F is the concentration resource function.

Recall that resources are defined as:

$$\begin{cases} rQ_F(x_k) = -\ln(1 - Q_F(x_k)) \\ rW_F(x_x) = -\ln(1 - W_F(x_x)). \end{cases} \quad (11)$$

Sets $[t_k], [t_x]$ are formed separately to solve the subject's problem. These sets can be used to implement various options for restoring the value of the lost resource of the metrologist's working capacity.

Taking into account the economic costs, it is possible to consider the optimal strategies of the seasonal type. Try to solve a similar problem. The question is not only how to determine the value of the restored resource, but how to make up for the loss of a part of the concentration resource. Or will it resume itself after the metrologist's preventive maintenance? But even in this

case, it is of interest how the concentration function will change. In the simplest case, using formulas (11), one can solve the following problem. Knowing the value of the spent resource in the concentration mode $rQ_F(x_k)$, , one should substitute it instead of $rW_F(x_x)$ in the chaos formula and solve the resulting equation regarding the determination of time $x_x$in the chaos mode. And then perform a reverse recalculation of this time, using the equalities of the resource values in these two modes and find a new time value x_k and the corresponding resource value in the concentration mode. Add this value of the resource with the value of the previously residual resource in the concentration mode. Draw conclusions about further changes in the regime.

Example. We use the third subsection of the article. We will express all numerical data in hours, passing from measurement in years to hours (1 year = $8{,}76 \times 10^3$ hours).

Suppose that the metrologist performed work in the environment of concentration described by the concentration function of P. Levy

$$Q_F(x_k) = \max_h\big(F(h + x_k + 0) - F(h)\big),$$

during the time $x_k = 60$ h with the set parameter $h = 30$ h and moved into the environment of chaos, described by the chaos function

$$W_F(x_k) = \max_{h_1}\big(F(h_1) - F(h_1 + x_k + 0)\big),$$

to restore his lost performance. Question: how long will it take to restore the lost operational resource and return to continue working in concentration mode? (In this case, he will add the value of the restored working capacity to the value of the residual, saved working capacity in the concentration mode.) The working capacity will be represented in the sense of N. M. Sedyakin's resource. The service life in the concentration mode will be designated as $rq(x) = -ln(1 - Q(x))$. The total potential resource will be equal to $rq(\infty) = 2.053$. The worked-out resource is equal to $rq(60) = 1.719$. The residual resource in the concentration mode is $rqo = rq(\infty) - rq(60) = 0.334$. The recovery time of the resource in the chaos mode will be determined by solving equation

$$rw(x_0) = ln(1 + W(x_0) + rq(60) = 0.$$

To do this, we will find a solution to the operator:

$$x_0 = 60$$
$$Given$$
$$ln(1 + W(x_0) + 1.719 = 0(15)$$
$$Find(x_0) = 64.514.$$

Thus, the consumed resource is equal to:

$$rw(64{,}514) = -1.719.$$

Changing the sign of the resulting number to the opposite and summing it up with the residual resource, we get:

$$1.719 + 0.334 = 2.053.$$

This is the initial potential resource.

Conclusion

An attempt is presented to consider a person metrologist as a recoverable element of the metrological system. Two stages of its life cycle as a seasonal system are investigated. The first stage consists in the direct fulfillment of the functional duties of metrology. The second stage of the cycle is associated with the restoration of the metrologist's working capacity, his health, without which it is impossible to ensure the further working capacity of the metrological system and obtaining the target effect. If this stage is associated with the implementation of the concentration function in the corresponding environment (we will call it the concentration environment), then the second stage is associated with the opposing environment (we will call it the chaos environment) associated with the provision of the first environment. From a formal point of view, the environment of chaos is characterized by a probabilistic distribution function opposite to the concentration function of P. Levy, and is called by his name [15, 16].

More complex chaos processes are not covered in the article. Within the limits of restrictions, a quantitative relationship between the stages is established and a formal way of realizing this relationship is proposed. A simple example of the interaction of a concentration environment with a chaotic environment is given.

References

1. Antonovsky A. V., Bysyuk A. S. Professionalnoe zdorovye inzhenerov-metrologov: teoreticheskie i prikladnye aspekty [Professional Health of Metrology Engineers: Theoretical and Applied Aspects]. In: *Asmakovets E. S. (ed.) Zdorovye spetsialista: problemy i puti resheniya: Materialy IV zaochnoy Mezhdunarodnoy nauchno-prakticheskoy internet-konferentsii [Specialist's Health: Problems and Solutions: Materials of IV International Scientific and Practical Internet Conference], Omsk, Russia, Plovdiv, Bulgaria, October 01–31, 2013*. Omsk, Institute of Education Development of Omsk Region, 2013, Pp. 12–25. (In Russian)

2. Alekseev G. A. Podgotovka inzhenerov v oblasti standartizatsii [Training of Engineers in the Field of Standardization], *Sovremennoe obrazovanie: soderzhanie, tekhnologii, kachestvo: Materialy XVIII Mezhdunarodnoy nauchno-metodicheskoy konferentsii [Modern Education: Content, Technology, Quality: Proceedings of the XVIII International Scientific and Methodological Conference], Saint Petersburg, Russia, April 18, 2012. Volume 1*. Saint Petersburg, Saint Petersburg Electrotechnical University «LETI», 2012, Pp. 87–89. (In Russian)

3. Smagin V. A. Matematicheskaya model nadezhnosti funktsionirovaniya kollektiva operatorov i slozhnykh programmnykh kompleksov [Mathematical Model of Reliability of Collective Operative and Complex Program Complexes Functioning], *Informatsiya i kosmos [Information and Space]*, 2007, No. 1, Pp. 75–80. (In Russian)

4. Akhmedzhanov F. M., Krymsky V. G. Algoritm otsenki nadezhnosti cheloveka-operatora na osnove modifitsirovannoy metodiki HEART [Algorithm for Assessment of Human Operator Reliability Based on Modified HEART Methodology], *Elektrotekhnicheskie i informatsionnye kompleksy i sistemy [Electrical and Data Processing Facilities and Systems]*, 2019, Vol. 15, No.1, Pp. 60–69.
DOI: 10.17122/1999-5458-2019-15-1-60-69. (In Russian)

5. Ivanov O. V., Ivanov V. O. Method for Forecasting the Reliability of an External Pilot of a Remote Piloted Aerial System, *Vestnik Natsionalnogo aviatsionnogo universiteta [Proceedings of National Aviation University]*, 2019, No. 4 (81), Pp. 29–33. DOI: 10.18372/2306-1472.81.14598.

6. Yakovlev A. V. Obobshchennyy algoritm otsenki funktsionalnogo sostoyaniya organizma cheloveka–operatora [Generalized Algorithm for Assessing the Functional State of the Operator], *Nauchnaya sessiya GUAP: Sbornik dokladov nauchnoy sessii, posvyashchennoy Vsemirnomu dnyu aviatsii i kosmonavtiki [SUAI Scientific Session: Collection of Reports of the Scientific Session Dedicated to the World Day of Aviation and Cosmonautics], Saint Petersburg, Russia, April 08–12, 2019. Volume 2.* Saint Petersburg, Saint-Petersburg State University of Aerospace Instrumentation, 2019, Pp. 288–290. (In Russian)

7. Yakovlev A. V., Matitsin V. O. Analiz primenimosti sushchestvuyushchikh metodov obrabotki dannykh dlya otsenki funktsionalnogo sostoyaniya organizma cheloveka-operatora i prognozirovaniya ego rabotosposobnosti [Analysis of the Applicability of Data Processing Methods for Assessing the Functional State of the Operator and Predicting His Performance], *Nauchnaya sessiya GUAP: Sbornik dokladov nauchnoy sessii, posvyashchennoy Vsemirnomu dnyu aviatsii i kosmonavtiki [SUAI Scientific Session: Collection of Reports of the Scientific Session Dedicated to the World Day of Aviation and Cosmonautics], Saint Petersburg, Russia, April 08–12, 2019. Volume 2.* Saint Petersburg, Saint-Petersburg State University of Aerospace Instrumentation, 2019, Pp. 291–294. (In Russian)

8. Guchuk V. V., Desova A. A., Dorofeyuk A. A., Dorofeyuk Yu. A. Strukturno-kognitivnaya metodika otsenki rabotosposobnosti cheloveka-operatora po informatsii ego pulsogrammy [Structural and Cognitive Method of Performance Assessment of Human Operator on Information of His Pulse Rate], *Kognitivnyy analiz i upravlenie razvitiem situatsiy (CASC`2011): Mezhdunarodnaya nauchno-prakticheskaya Multikonferentsiya «Upravlenie bolshimi sistemami 2011»: Trudy IX Mezhdunarodnoy konferentsii [Cognitive Analysis and Situation Control (CASC`2011): International Scientific and Practical Multi–Conference «Management of Large Systems 2011»: Proceedings of the IX International Conference], Moscow, Russia, November 14–16, 2011*. Moscow, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, 2011, Pp. 202–205. (In Russian)

9. Baranova T. I., Berlov D. N., Chiligina Yu. A., et al. Novye sposoby otsenki nadezhnosti cheloveka-operatora [New Methods of Evaluation of the Human Operator Reliability], *Aviakosmicheskaya i ekologicheskaya meditsina [Aerospace and Environmental Medicine]*, 2004, Vol. 38, No. 6, Pp. 31–36. (In Russian)

10. Smagin V. A., Bubnov V. P., Sultonov Sh. Kh. Matematicheskie modeli dlya rascheta kolichestvennykh kharakteristik optimalnogo kvantovaniya informatsii [Mathematical models for calculation the quantitative characteristics of the optimal quantization of information], *Transportnye sistemy i tekhnologii [Transportation systems and technology]*, Vol. 7, No. 1, Pp. 46–58. DOI: 10.17816/transsyst20217146-58. (In Russian)

11. Smagin V. A. Srednyaya chastota otkazov apparatury pri nenadezhnykh elementakh zameny [Average Hardware Failure Rate with Unreliable Replacement Elements], *Izvestiya Akademii nauk SSSR. Tekhnicheskaya kibernetika [Proceedings of the Academy of Sciences of the USSR. Technical Cybernetics]*, 1975, No. 3, Pp. 118–120. (In Russian)

12. Gnedenko B. V., Dinich M., Nasr Yu. O nadezhnosti dublirovannoy sistemy s vosstanovleniem i profilakticheskim obsluzhivaniem [About Reliability of Duplicated System with Restoration and Preventive Maintenance], *Izvestiya Akademii nauk SSSR. Tekhnicheskaya kibernetika [Proceedings of the Academy of Sciences of the USSR. Technical Cybernetics]*, 1975, No. 1, Pp. 66–71. (In Russian)

13. Barzilovich E. Yu., Kashtanov V. A. Nekotorye matematicheskie voprosy teorii obsluzhivaniya slozhnykh system [Some mathematical issues of the theory of service of complex systems]. Moscow, Sovetskoe Radio Publishing House, 1971, 271 p.

14. Smagin V. A. Teoreticheskoe obobshchenie fizicheskogo printsipa nadezhnosti professora N. M. Sedyakina [Theoretical Generalisation of Physical Principle of Reliability by Professor N. M. Sedyakin], *Nadezhnost [Dependability]*, 2005, No. 1 (24), Pp. 3–13. (In Russian)

15. Smagin V. A., Shurygin E. M. Application of the Concentration Function in Fuzzy Sets Theory, *Intellektualnye tekhnologii na transporte [Intellectual Technologies on Transport]*, 2020, No. 1 (21), Pp. 16–23.

16. Smagin V. A., Novikov A. N. Function of Concentration function P. Levy and Its Application in the Theory of Fuzzy Sets L. Zadeh, *Modeli, sistemy, seti v ekonomike, tekhnike, prirode i obshchestve [Models, Systems, Networks in Economics, Technology, Nature and Society]*, 2020, No. 3 (35), Pp. 102–117. DOI: 10.21685/2227-8486-2020-3-9. (In Russian)

DOI: 10.24412/2413-2527-2022-331-74-81
Русскоязычная версия статьи © В. А. Смагин, В. П. Бубнов опубликована
в журнале «Автоматика на транспорте». 2021. Т. 7, № 4. С. 617–630.
DOI: 10.20295/2412-9186-2021-7-4-617-630.

Несколько замечаний о самом важном элементе метрологии — человеке

д.т.н. В. А. Смагин
Международная академия информатизации
Санкт-Петербург, Россия
va_smagin@mail.ru

д.т.н. В. П. Бубнов
Петербургский государственный университет
путей сообщения Императора Александра I
Санкт-Петербург, Россия
bubnov1950@yandex.ru

Abstract. **В статье рассматривается человек как элемент метрологии. Функциональные обязанности и конкретные действия его не принимаются во внимание. Метролог представляется как двухфазная система, включающая два этапа жизненного цикла, первый цикл которой есть фаза концентрации — работа для получения эффекта, второй цикл — фаза хаоса, заключающаяся в восстановлении потраченных сил с целью продолжения первой фазы. Приводится формальная модель человека-метролога. Определяется оптимальная по коэффициенту готовности продолжительность межповерочного периода и среднее число ремонтов объекта за один год. При помощи реальной модели определяется средний возраст человека-оператора с точки зрения максимального коэффициента готовности, с учетом профилактических периодов и без них. В формальной и реальной моделях человека-оператора распределение времени жизни человека определяется экстремальным в статистике законом распределения Вейбулла. С формальной точки зрения среда хаоса характеризуется вероятностной функцией распределения, противоположной функции распределения среды концентрации по П. Леви. Определяется количественная связь между этапами и предлагается формальный путь реализации этой связи. Приведен простейший пример расчета восстановления исходного ресурса фазы концентрации.**

Keywords: **метролог, метрология, двухфазная система, среда концентрации, среда хаоса, средняя частота отказов, функция концентрации, обеспечение концентрации.**

Литература

1. Антоновский, А. В. Профессиональное здоровье инженеров-метрологов: теоретические и прикладные аспекты / А. В. Антоновский, А. С. Бысюк // Здоровье специалиста: проблемы и пути решения: Материалы IV заочной Международной научно-практической интернет-конференции (Омск, Россия, Пловдив, Болгария, 01–31 октября 2013 г.) / отв. ред. Е. С. Асмаковец. — Омск: Институт развития образования Омской области, 2013. — С. 12–25.

2. Алексеев, Г. А. Подготовка инженеров в области стандартизации // Современное образование: содержание, технологии, качество: Материалы XVIII Международной научно-методической конференции (Санкт-Петербург, Россия, 18 апреля 2012 г.): в 2 т. — Санкт-Петербург: ЛЭТИ, 2012. — Т. 1. — С. 87–89.

3. Смагин, В. А. Математическая модель надежности функционирования коллектива операторов и сложных программных комплексов // Информация и космос. 2007. № 1. С. 75–80.

4. Ахмеджанов, Ф. М. Алгоритм оценки надежности человека-оператора на основе модифицированной методики HEART / Ф. М. Ахмеджанов, В. Г. Крымский // Электротехнические и информационные комплексы и системы. 2019. Т. 15, № 1. С. 60–69.
DOI: 10.17122/1999-5458-2019-15-1-60-69.

5. Ivanov, O. V. Method for Forecasting the Reliability of an External Pilot of a Remote Piloted Aerial System / O. V. Ivanov, V. O. Ivanov // Вестник Национального авиационного университета. 2019. № 4 (81). С. 29–33.
DOI: 10.18372/2306-1472.81.14598.

6. Яковлев, А. В. Обобщенный алгоритм оценки функционального состояния организма человека–оператора // Научная сессия ГУАП: Сборник докладов научной сессии, посвященной Всемирному дню авиации и космонавтики (Санкт-Петербург, Россия, 08–12 апреля 2019 г.): в 3 ч. — Санкт-Петербург: ГУАП, 2019. — Ч. 2. — С. 288–290.

7. Яковлев, А. В. Анализ применимости существующих методов обработки данных для оценки функционального состояния организма человека-оператора и прогнозирования его работоспособности / А. В. Яковлев, В. О. Матыцин // Научная сессия ГУАП: Сборник докладов научной сессии, посвященной Всемирному дню авиации и космонавтики (Санкт-Петербург, Россия, 08–12 апреля 2019 г.): в 3 ч. — Санкт-Петербург: ГУАП, 2019. — Ч. 2. — С. 291–294.

8. Структурно-когнитивная методика оценки работоспособности человека-оператора по информации его пульсограммы / В. В. Гучук, А. А. Десова, А. А. Дорофеюк, Ю. А. Дорофеюк // Когнитивный анализ и управление развитием ситуаций (CASC`2011): Международная научно-практическая Мультиконференция «Управление большими системами–2011»: Труды IX Международной конференции (Москва, Россия, 14–16 ноября 2011 г.). — Москва: Институт проблем управления РАН, 2011. — С. 202–205.

9. Новые способы оценки надежности человека-оператора / Т. И. Баранова, Д. Н. Берлов, Ю. А. Чилигина, [и др.] // Авиакосмическая и экологическая медицина. 2004. Т. 38, № 6. С. 31–36.

10. Смагин, В. А. Математические модели для расчета количественных характеристик оптимального квантования информации / В. А. Смагин, В. П. Бубнов, Ш. Х. Султонов // Транспортные системы и технологии. 2021. Т. 7, № 1. С. 46–58. DOI: 10.17816/transsyst20217146-58.

11. Смагин, В. А. Средняя частота отказов аппаратуры при ненадежных элементах замены // Известия Академии наук СССР. Техническая кибернетика. 1975. № 3. С. 118–120.

12. Гнеденко, Б. В. О надежности дублированной системы с восстановлением и профилактическим обслуживанием / Б. В. Гнеденко, М. Динич, Ю. Наср // Известия Академии наук СССР. Техническая кибернетика. 1975. № 1. С. 66–71.

13. Барзилович, Е. Ю. Некоторые математические вопросы теории обслуживания сложных систем / Е. Ю. Барзилович, В. А. Каштанов. — Москва: Советское радио, 1971. — 271 с.

14. Смагин, В. А. Теоретическое обобщение физического принципа надежности профессора Н. М. Седякина // Надежность. 2005. № 1 (24). С. 3–13.

15. Smagin, V. A. Application of the Concentration Function in Fuzzy Sets Theory / V. A. Smagin, E. M. Shurygin // Интеллектуальные технологии на транспорте. 2020. № 1 (21). С. 16–23.

16. Смагин, В. А. Функция концентрации П. Леви и ее применение в теории нечетких множеств Л. Заде / В. А. Смагин, А. Н. Новиков // Модели, системы, сети в экономике, технике, природе и обществе. 2020. № 3 (35). С. 102–117. DOI: 10.21685/2227-8486-2020-3-9.