

ВЕСТНИК НОВОСИБИРСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Научный журнал
Основан в ноябре 1999 года

Серия: Информационные технологии

2022. Том 20, № 4

СОДЕРЖАНИЕ

<i>Андросов А. С., Шарый С. П.</i> IntvalPy – библиотека интервальных вычислений на языке Python	5
<i>Лисин В. А., Серый А. С., Сидорова Е. А.</i> Модель представления онтологии предметных областей на основе графовых баз данных	24
<i>Попов А. В., Пьянков О. В.</i> Исследование взаимосвязи между конфликтными свойствами и показателями эффективности организационно-технических систем на примере сети связи специального назначения	39
<i>Прокофьева А. В., Шниперов А. Н.</i> Метод стеганоанализа JPEG-изображений на основе цепей Маркова и его применение в сочетании с различными алгоритмами машинного обучения	61
<i>Тимофеев Г. А.</i> Поиск подходящей архитектуры для разработки цифрового двойника гибридных энергетических систем в изолированных от сетевых энергосистем средах с использованием ТРИЗ-эволюционного подхода	76
Информация для авторов	100

V E S T N I K

NOVOSIBIRSK STATE UNIVERSITY

Scientific Journal
Since 1999, November
In Russian

Series: Information Technologies

2022. Volume 20, № 4

CONTENTS

<i>Androsov A. S., Shary S. P.</i> IntvalPy – a python interval computation library	5
<i>Lisin V. A., Sery A. S., Sidorova E. A.</i> Domain ontology representation model based on graph databases.....	24
<i>Popov A. V., Pyankov O. V.</i> Study of the relationship between conflict properties and performance indicators of organizational and technical systems on the example of a special purpose communication network	39
<i>Prokofieva A. V., Shniperov A. N.</i> A Markov-chain-based method for JPEG image steganalysis and its application in combination with various machine learning algorithms	61
<i>Timofeev G. A.</i> Searching for an appropriate architecture for the development of digital twins for hybrid energy systems in isolated environments using the TRIZ evolutionary approach	76
Instructions for Contributors.....	100

Editor in Chief M. M. Lavrentiev

Vice-Editor A. V. Avdeev

Executive Secretary D. P. Iksanova

Editorial Board of the Series

- I. V. Bychkov*, professor, academician (Irkutsk), *B. M. Glinsky*, professor (Novosibirsk)
A. N. Gorban, professor (Lester, GB), *E. P. Gordov*, professor (Tomsk)
B. S. Dobronets, professor (Krasnoyarsk), *A. M. Elizarov*, professor (Kazan)
G. N. Erokhin, professor (Kaliningrad), *A. I. Kamysnikov*, professor (Khanty-Mansiysk)
G. P. Karev, professor (Maryland, USA), *N. A. Kolchanov*, professor, academician (Novosibirsk)
M. M. Lavrentjev, professor (Novosibirsk), *V. E. Malyshkin*, professor (Novosibirsk)
N. N. Mirenkov, professor (Aizu, Japan), *N. M. Oskorbin*, professor (Barnaul)
D. E. Palchunov, professor (Novosibirsk), *T. Pizansky*, professor (Ljubljana, Slovenia)
V. P. Potapov, professor (Kemerovo), *O. I. Potaturkin*, professor (Novosibirsk)
V. A. Serebryakov, professor (Moscow), *A. V. Starchenko*, professor (Tomsk)
S. I. Smagin, professor, corresponding member of RAS (Khabarovsk)
D. A. Tusupov, professor (Astana, Kazakhstan)
V. V. Shajdurov, professor, corresponding member of RAS (Krasnoyarsk)
Yu. I. Shokin, professor, academician (Novosibirsk)

*The journal is published quarterly in Russian since 1999
by Novosibirsk State University Press*

*The address for correspondence
Faculty of Information Technologies, Novosibirsk State University
1 Pirogov Street, Novosibirsk, 630090, Russia
Tel. +7 (383) 363 42 46*

E-mail address: inftech@vestnik.nsu.ru

On-line version: <http://elibrary.ru>

Научная статья

УДК 519.68

DOI 10.25205/1818-7900-2022-20-4-5-23

IntvalPy – библиотека интервальных вычислений на языке Python

Артём Станиславович Андросов, Сергей Петрович Шарый

Федеральный исследовательский центр информационных и вычислительных технологий
Новосибирск, Россия

artem.androsov@gmail.com, <https://orcid.org/0009-0008-9043-0157>

shary@ict.nsc.ru, <https://orcid.org/0000-0002-9454-7626>

Аннотация

В статье рассматривается библиотека IntvalPy, реализующая интервальные вычисления на языке Python. В отличие от других существующих интервальных библиотек IntvalPy предоставляет возможность работы как с классической интервальной арифметикой, так и с полной интервальной арифметикой Каухера. Кроме того, библиотека была разработана с учетом стандарта IEEE 1788–2015 на интервальные вычисления на ЭВМ, что гарантирует высокую точность результатов, а также совместимость с другими программными продуктами. Верхнеуровневая функциональность библиотеки IntvalPy реализует новейшие методы для распознавания и оценивания множеств решений интервальных линейных систем уравнений, вычисления их формальных решений, а также визуализации множеств решений интервальных уравнений и систем уравнений. В качестве примера приложения библиотеки была решена практически важная задача оценивания параметров электрохимического процесса формирования рыхлых осадков порошков металла. Кроме того, были проведены численные и качественные сравнения с другими интервальными библиотеками для демонстрации функциональных возможностей и оптимальности реализованных интервальных классов.

Ключевые слова

интервальный анализ, библиотека интервальных вычислений, Python, неточные данные, интервальные системы уравнений, визуализация

Для цитирования

Андросов А. С., Шарый С. П. IntvalPy – библиотека интервальных вычислений на языке Python // Вестник НГУ. Серия: Информационные технологии. 2022. Т. 20, № 4. С. 5–23. DOI 10.25205/1818-7900-2022-20-4-5-23

IntvalPy — a Python Interval Computation Library

Artem S. Androsov, Sergey P. Shary

Federal Research Center for Information and Computational Technologies
Novosibirsk, Russian Federation

artem.androsov@gmail.com, <https://orcid.org/0009-0008-9043-0157>

shary@ict.nsc.ru, <https://orcid.org/0000-0002-9454-7626>

Abstract

The article presents the IntvalPy library which implements interval computations in Python. Unlike other existing interval libraries, IntvalPy allows one to work with both classical interval arithmetic and complete Kaucher interval arithmetic. In addition, the library was developed taking into account the IEEE 1788-2015 standard for interval arithmetic on digital computers, which guarantees high accuracy of the results and compatibility with the other existing software

© Андросов А. С., Шарый С. П., 2022

products. The top-level functionality of the IntvalPy library implements state-of-the-art methods for recognizing and estimating solution sets for interval linear systems of equations, computing their formal solutions, and visualizing solution sets for interval equations and systems of equations. As an example of the library application, we solve the practically important problem of estimating the parameters of the electrochemical process of the formation of loose metal powder precipitates. Additionally, numerical computation was carried out, as well as qualitative comparisons with other interval libraries, in order to demonstrate the functionality and optimality of implemented interval classes.

Keywords

interval analysis, interval computing library, Python, inaccurate data, interval systems of equations, visualization

For citation

Androsova A. S., Shary S. P. IntvalPy — a Python Interval Computation Library. *Vestnik NSU. Series: Information Technologies*, 2022, vol. 20, no. 4, pp. 5–23. (in Russ.) DOI 10.25205/1818-7900-2022-20-4-5-23

Введение

Настоящая работа посвящена созданию библиотеки программ на языке Python, реализующей интервальные вычисления и численные методы интервального анализа для решения ряда стандартных математических задач. Это современные методы для определения существования решений как квадратных, так и переопределенных интервальных систем линейных алгебраических уравнений, вычисление внешних и внутренних оценок множеств решений, а также визуализации множеств решений интервальных уравнений и систем уравнений.

Перед нами стояла задача реализовать на популярном алгоритмическом языке развитый инструментарий современного интервального анализа, главным образом касающийся интервальных систем линейных алгебраических уравнений и родственных к ним постановок.

Мотивацией для данной работы послужил также тот факт, что в языке Python отсутствуют хорошо развитые и известные интервальные библиотеки, в которых реализована подобная функциональность. Последняя созданная библиотека, в которой реализованы интервальные методы, уже не поддерживается ее разработчиками.

Кроме того, среди большого количества программных модулей, реализованных к настоящему моменту для Python, отсутствуют высокоуровневые интервальные средства, необходимые для решения таких практически важных задач, как распознавание разрешимости интервальных систем уравнений, оценивание и визуализация их множеств решений и т. п. Кроме того, нет хороших продуктов для визуализации множеств решений линейных неравенств $Ax \leq b$ для случаев двух и трех переменных. В силу известных результатов такая функциональность позволила бы реализовать визуализацию множества решений для интервальных линейных уравнений. Другой ключевой мотивацией нашей работы было то, что в большинстве созданных интервальных библиотек для Python отсутствует возможность работать в интервальных арифметиках, отличных от классической. Но классическая интервальная арифметика не имеет обратных элементов относительно основных арифметических операций и у нее плохие порядковые свойства. Таким образом, порой необходимо работать в полной интервальной арифметике или арифметике Каухера. Все вышеперечисленные соображения послужили причиной для создания указанной библиотеки IntvalPy (см. [1]).

1. Теория

Интервалом $[a, b]$ вещественной оси $\mathbb{R}$ будем называть множество всех чисел, расположенных между числами a и b , включая их самих, т. е.

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}.$$

Система обозначений в работе соответствует неформальному международному стандарту в интервальном анализе [2]. В частности, интервалы и интервальные величины обозначаются

жирным шрифтом, к примеру, $\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots$, тогда как точечные величины специальным образом не выделяются. Кроме того, будем обозначать концы интервала (т. е. его правую и левую границы), как $\underline{\mathbf{a}}$ и $\overline{\mathbf{a}}$, т. е. $\mathbf{a} = [\underline{\mathbf{a}}, \overline{\mathbf{a}}]$. Интервал называется *вырожденным* при $\underline{\mathbf{a}} = \overline{\mathbf{a}}$ и *невырожденным* иначе.

Прежде всего определим наиболее важные и часто встречающиеся арифметические операции между интервалами. Основным руководящим принципом будет служить соотношение

$$\mathbf{a} * \mathbf{b} = \{a * b \mid a \in \mathbf{a}, b \in \mathbf{b}\}, \quad * \in \{+, -, \cdot, /\},$$

согласно которому интервальные арифметические операции выполняются «по представителям». Конструктивные определения интервальных арифметических операций имеют вид:

$$\text{сложение: } \mathbf{a} + \mathbf{b} = [\underline{\mathbf{a}} + \underline{\mathbf{b}}, \overline{\mathbf{a}} + \overline{\mathbf{b}}],$$

$$\text{вычитание: } \mathbf{a} - \mathbf{b} = [\underline{\mathbf{a}} - \overline{\mathbf{b}}, \overline{\mathbf{a}} - \underline{\mathbf{b}}],$$

$$\text{умножение: } \mathbf{a} \cdot \mathbf{b} = [\min S, \max S], \text{ где } S = \{\underline{\mathbf{a}}\underline{\mathbf{b}}, \underline{\mathbf{a}}\overline{\mathbf{b}}, \overline{\mathbf{a}}\underline{\mathbf{b}}, \overline{\mathbf{a}}\overline{\mathbf{b}}\},$$

$$\text{деление: } \mathbf{a}/\mathbf{b} = \mathbf{a} \cdot [1/\overline{\mathbf{b}}, 1/\underline{\mathbf{b}}], \text{ если } 0 \notin \mathbf{b}.$$

Алгебраическая система, носителем которой является множество всех интервалов вещественной оси $\mathbb{R}$ с определенными для них интервальными арифметическими операциями сложения, вычитания, умножения и деления, называется *классической интервальной арифметикой* и обозначается $\mathbb{IR}$.

К сожалению, алгебраические свойства этой арифметики довольно скудны по сравнению с привычными числовыми системами, такими как поля рациональных чисел $\mathbb{Q}$ или вещественных чисел $\mathbb{R}$, поскольку:

1. Всякий невырожденный интервал не имеет обратного элемента относительно арифметических операций.
2. Отсутствует полноценная дистрибутивность умножения относительно сложения и вычитания.
3. Неудовлетворительны порядковые свойства относительно упорядочения по включению.

Из-за отсутствия дистрибутивности невозможно приводить подобные члены, что является серьезным недостатком интервальной арифметики. Для преодоления этой проблемы необходимо в значительной степени изменить данную арифметику как алгебраическую систему, что ставит под вопрос саму целесообразность этого шага. Однако проблему необратимости арифметических операций и плохих порядковых свойств можно частично решить с помощью построения $\mathbb{IR}$ до некоторой более широкой и полной интервальной арифметики $\mathbb{KR}$, которая была предложена Э. Каухером.

Элементами полной арифметики Каухера $\mathbb{KR}$ являются пары вещественных чисел $[a, b]$ которые не обязательно связаны соотношением $a \leq b$. В результате $\mathbb{KR}$ получается при соединении *неправильных интервалов* к множеству обычных (правильных) интервалов $\mathbb{IR} = \{[a, b] \mid a, b \in \mathbb{R}, a \leq b\}$ и вещественных чисел. Для перехода от неправильных к правильным интервалам применяется операция *дуализации* $\text{dual}: \mathbb{KR} \rightarrow \mathbb{KR}$.

Распространение операций сложения и вычитания на полную интервальную арифметику выполняется достаточно просто, а явные формулы для умножения удобно выделить в следующие подмножества, которые можно представить в виде табл. 1:

$$\mathcal{P} := \{\mathbf{a} \in \mathbb{KR} \mid \underline{\mathbf{a}} \geq 0, \overline{\mathbf{a}} \geq 0\} \text{ – неотрицательные интервалы,}$$

$$\mathcal{Z} := \{\mathbf{a} \in \mathbb{KR} \mid \underline{\mathbf{a}} \leq 0 \leq \overline{\mathbf{a}}\} \text{ – нульсодержащие интервалы,}$$

$$-\mathcal{P} := \{\mathbf{a} \in \mathbb{KR} \mid -\mathbf{a} \in \mathcal{P}\} \text{ – неположительные интервалы,}$$

$$\text{dual } \mathcal{Z} := \{\mathbf{a} \in \mathbb{KR} \mid \text{dual } \mathbf{a} \in \mathcal{Z}\} \text{ – интервалы, содержащиеся в нуле.}$$

Таблица 1

Умножение в арифметике Каухера

Table 1

Multiplication in Kaucher Arithmetic

$a \cdot b$	$b \in \mathcal{P}$	$b \in \mathcal{Z}$	$b \in -\mathcal{P}$	$b \in \text{dual } \mathcal{Z}$
$a \in \mathcal{P}$	$[\underline{a} \underline{b}, \overline{a} \overline{b}]$	$[\overline{a} \underline{b}, \underline{a} \overline{b}]$	$[\overline{a} \underline{b}, \underline{a} \overline{b}]$	$[\underline{a} \underline{b}, \underline{a} \overline{b}]$
$a \in \mathcal{Z}$	$[\underline{a} \underline{b}, \overline{a} \overline{b}]$	$[\min\{\underline{a} \underline{b}, \overline{a} \underline{b}\}, \max\{\underline{a} \underline{b}, \overline{a} \underline{b}\}]$	$[\overline{a} \underline{b}, \underline{a} \overline{b}]$	0
$a \in -\mathcal{P}$	$[\underline{a} \underline{b}, \overline{a} \overline{b}]$	$[\underline{a} \underline{b}, \underline{a} \overline{b}]$	$[\overline{a} \underline{b}, \underline{a} \overline{b}]$	$[\overline{a} \underline{b}, \underline{a} \overline{b}]$
$a \in \text{dual } \mathcal{Z}$	$[\underline{a} \underline{b}, \overline{a} \overline{b}]$	0	$[\overline{a} \underline{b}, \underline{a} \overline{b}]$	$[\max\{\underline{a} \underline{b}, \overline{a} \underline{b}\}, \min\{\underline{a} \underline{b}, \overline{a} \underline{b}\}]$

2. Постулаты

При создании любого программного модуля разработчики руководствуются некоторыми основными требованиями. В данном разделе в порядке убывания значимости указаны наиболее приоритетные условия, которыми руководствовались мы.

1. Библиотека должна быть проста в использовании, а написанный код – читаемым и понятным. Хорошо спроектированный модуль, построенный в соответствии с ясной и продуманной структурой, будет проще для изучения и использования. Вряд ли пользователи будут разбираться с беспорядочной структурой, даже если программный инструмент наиболее современен.

2. Библиотека должна удовлетворять общепринятым стандартам – IEEE 754-2008 для арифметики с плавающей точкой на ЭВМ и IEEE 1788-2015 для интервальных вычислений на ЭВМ. По умолчанию пользователь работает в режиме повышенной точности для концов интервалов, в котором количество знаков после запятой – варьируемая величина. Но при желании, если позволяет постановка задачи, пользователь может перейти в режим округления к ближайшему четному числу, что существенно сократит скорость выполнения алгоритмов.

3. Библиотека должна обеспечивать высокую производительность. Если решение задач полиномиальной сложности будет требовать изрядного количества времени, то вряд ли подобную функциональность можно назвать полезной.

4. Библиотека должна быть кроссплатформенной.¹

5. Библиотека должна обеспечивать возможность работы в разных интервальных арифметиках (классическая, Каухера, Кэхэна и т. д.). Также в случае, если арифметики совместимы друг с другом, то система должна автоматически переходить от одной к другой.

6. Библиотека должна быть гибкой к изменениям и обеспечивать возможность расширения функциональности. Архитектура библиотеки должна позволять вносить изменения как на низкоуровневой функциональности, так и на верхнеуровневых интервальных методах (решатели уравнений, визуализация множеств решений и т. д.).

3. Архитектура библиотеки

Библиотека IntvalPy, написанная на языке программирования Python, реализует алгебраически замкнутую систему для работы с интервалами, для решения интервальных систем

¹ Отметим, что в отличие от других интервальных библиотек, написанных на языке Python, IntvalPy кроссплатформенна, и это является одним из ее главных достоинств.

как линейных, так и нелинейных уравнений, а также для визуализации множества решений интервальных линейных систем уравнений.

В этом разделе представлены основные идеи, которые определяют архитектуру библиотеки и ее реализацию. Предназначенный в большей степени для практических задач модуль по умолчанию хранит концы интервалов в формате `np.float64` (стандартная «двойная точность») и не использует для каждой операции направленные округления. Это позволяет значительно ускорить вычисления и сократить время работы алгоритмов. Однако в некоторых случаях современных научных вычислений требуется более высокая точность, и поэтому пользователю предоставлена возможность работать с интервалами, концы которых представляются в формате `mpf` (`multiprecision format`, где хранится задаваемое пользователем количество знаков мантиссы). При этом режим округления при арифметических операциях с интервалами на данный момент остается стандартным – округление к ближайшему числу с четной младшей значащей цифрой.

Язык Python для реализации нашей библиотеки был выбран потому, что это универсальный и свободно распространяемый язык с открытым исходным кодом, а также кроссплатформенной переносимостью. Кроме того, у языка сформировалось огромное сообщество пользователей и разработчиков, качественная документация и большое количество доступных библиотек, в которых реализованы самые разнообразные инструменты и методы. В частности, в пакете `IntvalPy` для хранения левых и правых концов интервалов (которые могут являться многомерными) активно используется широко распространенный модуль `NumPy`. Его применение особенно полезно потому, что позволяет создавать не только векторы, но и матрицы с соответствующими поэлементными операциями. Такой подход в значительной степени ускоряет вычисления, так как сам модуль `NumPy` интегрирован с языком `C++` и именно в нем реализуются циклы.

Реализация библиотеки `IntvalPy` соответствует схеме, показанной на рис. 1, где стрелки обозначают включение (подчинение) отдельных структурных элементов в другие элементы. Наиболее подробно разработаны низкоуровневые операции, а также численные методы для интервальных линейных систем.

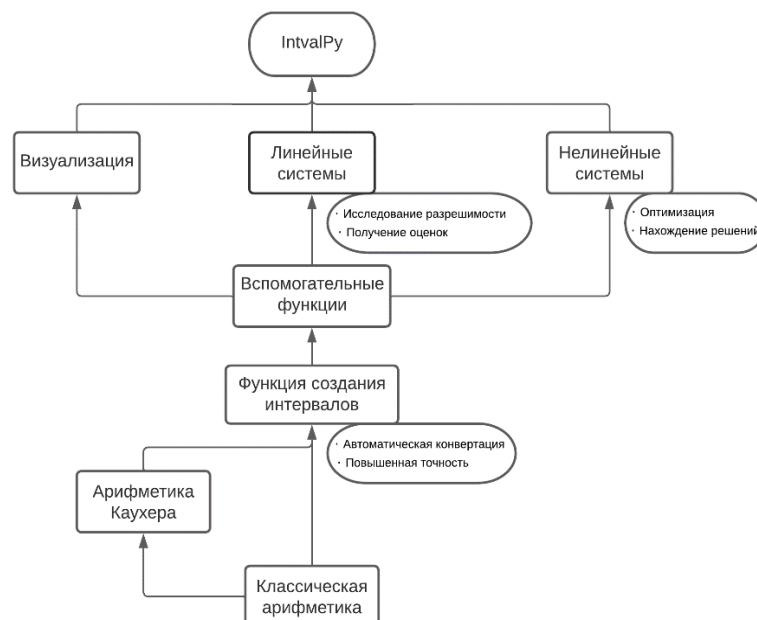


Рис. 1. Архитектура `IntvalPy`

Fig. 1. `IntvalPy` architecture

Базовыми классами библиотеки IntvalPy являются Classical и Kaucher, в которых определяются все арифметические операции, абсолютные характеристики интервалов, операции для работы с многомерными интервальными массивами. Также предусмотрены некоторые методы, которые могут менять принадлежность интервала к классу (например, дуализация или получение алгебраически обратного интервала). Для удобства создания интервалов была написана отдельная функция Interval, которая определяет, к какой арифметике относится интервал или массив интервалов произвольной размерности. Стоит отметить, что по умолчанию стоит автоматическая конвертация концов интервалов, т. е. интервалы всегда «правильные». Поэтому, если пользователь хочет работать в полной арифметике Каухера, то ему необходимо отключить эту функциональность. Кроме того, в функции Interval определяется, будут ли данные храниться в формате pr.float64 или mpf.

Введение подобной функции соответствует основным принципам, которые были описаны ранее. В целом библиотека спроектирована как расширяемая, и все ее будущие реализации можно будет без особых усилий сделать совместимыми с предшествующими версиями.

Другие библиотеки, использованные при создании IntvalPy, также широко известны и могут применяться для других операционных систем. Это соответствует сформулированному выше принципу кроссплатформенности библиотеки.

4. Функциональность

Практическая пригодность интервального пакета IntvalPy обусловлена имеющимися в нем функциональными возможностями. При этом программный модуль продолжает развиваться, в него вводится новый инструментарий.

4.1. Низкоуровневая функциональность

Операции нижнего функционального слоя модуля IntvalPy, доступные для использования, – это операции интервальных арифметик, классической интервальной арифметики и полной арифметики Каухера. Однако в некоторых высокоуровневых функциях встречаются выражения с применением деления из арифметики Кэхэна, допускающие нуль в интервале делителя. Таким образом, реализованы основные арифметические операции, элементарные функции, а также различные характеристики интервалов, которые соответствуют стандарту IEEE 1788-2015. В табл. 2 резюмированы операции доступные для применения.

Таблица 2

Доступные операции в IntvalPy

Table 2

Available Operations in IntvalPy	
Операции	Описание
$+$, $-$, $\cdot$, $/$	Арифметические операции
pow, abs, exp, log, sin, cos	Элементарные функции
matmul, dot, transpose	Матричные операции
$<$, $\leq$, $=$, $\geq$, $>$, $\neq$, $\subset$, $\subseteq$, $\cap$	Бинарные отношения и операции
rad, wid, mid, mig, mag	Характеристики интервалов
dual, pro, opp, inv	Преобразования интервалов

С помощью созданных классов пользователь может конструировать интервалы, работать с интервальными арифметическими операциями, интервальными расширениями элементарных функций, а также применять различные служебные функции – получать характеристики

интервалов и т. п. Нахождение числовых характеристик интервалов и различные преобразования интервалов вызываются как соответствующие методы классов.

Отметим следующее. Поскольку программный модуль предоставляет возможность работы с интервальными векторами и матрицами, то не следует забывать о том, что аргумент передается в функцию «по ссылке» (поверхностная копия), а не «по значению». Поэтому при расширении функциональности библиотеки необходимо создавать глубокие копии аргументов внутри функций (т. е. отдельные объекты), чтобы не столкнуться с неприятными сюрпризами в дальнейшем.

```

1  import intvalpy as ip
2  import numpy as np
3
4  # Barth-Nuding system
5  A = ip.Interval([
6      [[2, 4], [-2, 1]],
7      [[-1, 2], [2, 4]]
8  ])
9  b = ip.Interval([-2, 2], [-2, 2])
10 x = np.array([-0.2, 0.4])
11
12 ip.subset(A @ x, b)      # True
13 # A @ x = Interval(['[-1.6, 0]', '[0.4, 1.8]'])
14
15 a = A[0, 1].opp          # Interval(['[2, -1]'])
16 type(A[0, 1].opp)
17 # intvalpy.RealInterval.KaucherArithmetic

```

Листинг 1. Основные операции
Listing 1. Basic operations

Рассмотрим программный код в листинге 2, который реализует вычисление выражения из известного примера Румпа [3; 4]:

$$f(a, b) = 333.75 \cdot b^6 + a^2 \cdot (11 \cdot a^2 \cdot b^2 - b^6 - 121 \cdot b^4 - 2) + 5.5 \cdot b^8 + a/(2 \cdot b). \quad (1)$$

Этот пример иллюстрирует тот факт, что результаты численных расчетов на компьютере не обязательно должны приближаться к истинному ответу, несмотря на применение возрастающей точности вычислений (float, double, long double и т. д.).

```

1  import intvalpy as ip
2
3  ip.precision.extendedPrecisionQ = True
4  ip.precision.dps(40)
5
6  f = lambda a, b: 333.75 * b**6 + \
7      a**2 * (11 * a**2 * b**2 - b**6 - \
8      121 * b**4 - 2) + 5.5 * b**8 + a/(2*b)
9
10 a = ip.Interval(77617, 77617)
11 b = ip.Interval(33096, 33096)
12 f(a, b)
13 # Interval(['[-0.827396, -0.827396]'])

```

Листинг 2. Пример Румпа
Listing 2. Rump example

Интервальные библиотеки, реализующие интервальные вычисления с внешним направленным округлением, для примера Румпа дают довольно широкие интервалы (шириной порядка 10^{21}), которые содержат правильный ответ. К сожалению, эти ответы малопригодны для практического использования из-за своей огромной ширины. Если же увеличивать точность представления концов, то результирующий интервал будет сужаться. С помощью библиотеки `IntvalPy`, используя повышенную точность представления концов интервалов с 40 знаками после запятой, можно получить довольно узкие интервальные оценки результата.

Кроме явно указанных функций, в библиотеке `IntvalPy` есть и другие, например, `zeros`, `eye`, `Shary`, `Neumeier` и т. п. Использование этих функций помогает писать более короткие и выразительные программы, а также создавать тестовые задачи с хорошо изученными свойствами. Читатель может найти описания этих функций в документации, однако библиотека регулярно расширяется, и документация может несколько отставать от фактического состояния библиотеки. Впрочем, самую актуальную информацию можно найти непосредственно в исходном коде библиотеки, который расположен в программном репозитории `GitHub`.

4.2. Верхнеуровневая функциональность

Одна из отличительных особенностей интервальной библиотеки `IntvalPy` – реализация в ней высокоуровневых решателей для различных задач интервального анализа и его приложений. В частности, в библиотеке имеются развитые методы решения как квадратных, так и переопределенных систем интервальных линейных уравнений. Для двух- и трехмерных интервальных линейных систем на основе метода граничных интервалов, предложенного И. А. Шарой [5, 6], реализованы функции для визуализации различных множеств решений. Кроме того, в библиотеке представлены функции вычисления и максимизации распознающих функционалов множеств решений, которые являются основой метода максимума согласования [7, 8] для восстановления зависимостей по интервальным данным.

Далее мы продемонстрируем, как именно применять `IntvalPy` на нескольких конкретных примерах.

В качестве первого примера разберем исследование разрешимости интервальной системы линейных алгебраических уравнений (ИСЛАУ), понимаемое, как непустота соответствующего множества решений (ее также называют «сильной разрешимостью», см. [9]). Ниже мы рассмотрим допустимое множество решений ИСЛАУ (см., к примеру, [8, 10]), и для его исследования применим *метод распознающего функционала*. Он сводит исходную задачу к безусловной максимизации некоторой специальной функции, называемой распознающим функционалом множества решений, и по итогам нахождения ее максимума делается заключение о пустоте/непустоте множества решений, а также получается точка, обладающая наибольшей совместимостью с данной системой уравнений.

Во фрагменте кода, приведенном на листинге 3, рассмотрена тестовая система, предложенная А. Ноймайером в работе [11].

```
1 | import intvalpy as ip
2 |
3 | A, b = ip.Neumeier(3, 3.33, infb=-1, supb=2)
4 | x = np.array([-3, 7, 2])
5 | print('Tol: ', ip.linear.Tol(A, b, x=x))
6 | print('max Tol: ', ip.linear.Tol(A, b, maxQ=True))
```

Листинг 3. Метод распознающих функционалов
Listing 3. The method of recognizing functionals

Представленный код проверяет, принадлежит ли точка x допустовому множеству решений (в первом операторе `print`), а далее, с помощью максимизации распознающего функционала `Tol`, находит точку, которая в наибольшей степени совместна с рассматриваемой ИСЛАУ (во втором операторе `print`). Напомним, что для интервальных линейных систем уравнений распознающий функционал `Tol` обладает хорошими свойствами: он является непрерывным, вогнутым и достигает конечного максимума, что упрощает задачу нахождения его безусловного максимума. В частности, в функции `ip.linear.Tol` используется алгоритм негладкой оптимизации `ralgb5`, разработанный П.И. Стецюком [12].

Отметим, что помимо сильной разрешимости (относительно допустового множества решений) часто рассматривается также слабая разрешимость ИСЛАУ, понимаемая как непустота ее объединенного множества решений. Это множество решений является, по-видимому, наиболее популярным из различных множеств решений для интервальных систем уравнений. Его пустота/непустота тоже может быть исследована с помощью распознающего функционала `Uni`, но в общем случае эта задача является труднорешаемой (NP-трудной), а распознающий функционал `Uni` в общем случае является многоэкстремальным. При нахождении его глобального максимума большую роль играет выбор начального приближения, и это также характерно для максимизации распознающих функционалов нелинейных интервальных систем уравнений.

Полное исчерпывающее описание множеств решений интервальных систем уравнений невозможно и практически не нужно, так что в случае непустоты множеств решений обычно находят для них приближенные оценки в том или ином смысле, требуемом смыслом рассматриваемой задачи [10]. Очень популярны, в частности, внешние оценки множеств решений, с помощью объемлющих интервальных брусков. Для их получения предложены интервальные методы Гаусса, Гаусса–Зейделя, процедура Хансена–Блика–Рона, метод Рона, методы дробления решений (PSS-методы), методы дробления параметров (PPS-методы) и др. Среди всех этих численных методов наиболее интересны для нас PSS-методы, которые позволяют находить оптимальную внешнюю интервальную оценку множеств решений и, кроме того, обладают свойством последовательной гарантии.

Напомним, что задача вычисления оптимальных или гарантированно близких к оптимальным оценок объединенного множества решений ИСЛАУ является NP-трудной, поэтому в библиотеке `IntvalPy` реализована гибридная версия PSS-методов, удобная для решения практических задач значительной размерности. Кроме того, эти методы позволяют решать переопределенные линейные интервальные системы, возникающие при обработке данных. В листинге 4 представлено численное решение модельной тестовой системы С.П. Шарого [13] с размерностью 20×20 .

```

1 | import intvalpy as ip
2 |
3 | A, b = ip.Shary(20)
4 | pss = ip.linear.PSS(A, b)
```

Листинг 4. Метод дробления решений
Listing 4. Solution splitting methods

В качестве последнего примера рассмотрим задачу визуализации объединенного множества решений ИСЛАУ (в библиотеке `IntvalPy` реализована также процедура визуализации допустового множества решений). Типичные подходы к визуализации полиэдральных множеств, какими являются множества решений ИСЛАУ, основаны на перечислении их вершин, но этот путь и основанные на нем процедуры имеют ряд недостатков. Они работают только с квадратными системами, плохо обрабатывают неограниченные и тощие множества решений (с пустой

внутренностью). В библиотеке IntvalPy визуализация множеств решений использует метод граничных интервалов, предложенный И.А. Шарой для исследования и отрисовки полиэдральных множеств [5, 6, 14]. Главным преимуществом этого подхода является возможность работать с неограниченными и тощими множествами решений, а также с линейными системами, в которых количество уравнений не равно количеству неизвестных.

```

1  import intvalpy as ip
2  %matplotlib notebook
3
4  A = ip.Interval([
5      [-1, 1], [-2, 2], [-2, 2]],
6      [[-2, 2], [-1, 1], [-2, 2]],
7      [[-2, 2], [-2, 2], [-1, 1]]
8  ])
9  b = ip.Interval([[2, 2], [2, 2], [2, 2]])
10
11 bounds = [[-2, -2, -2], [2, 2, 2]]
12 ip.IntLinIncR3(A, b, alpha=0.5, s=0, \
13               bounds=bounds, size=(11,11))

```

Листинг 5. Метод граничных интервалов
Listing 5. Boundary intervals method

Результат работы программы листинга 5 представлен слева на рис. 2, а справа, на этом же рисунке, – пример визуализации тощего множества решений переопределенной интервальной системы (2) линейных алгебраических уравнений.

$$\begin{pmatrix} [-1,1] & [0,0] & [0,0] \\ [0,0] & [-1,1] & [0,0] \\ [0,0] & [0,0] & [-1,1] \\ [1,1] & [0,0] & [0,0] \\ [0,0] & [1,1] & [0,0] \\ [0,0] & [0,0] & [1,1] \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} [1,1] \\ [1,1] \\ [1,1] \\ [-1,2] \\ [-1,2] \\ [-1,2] \end{pmatrix}. \quad (2)$$

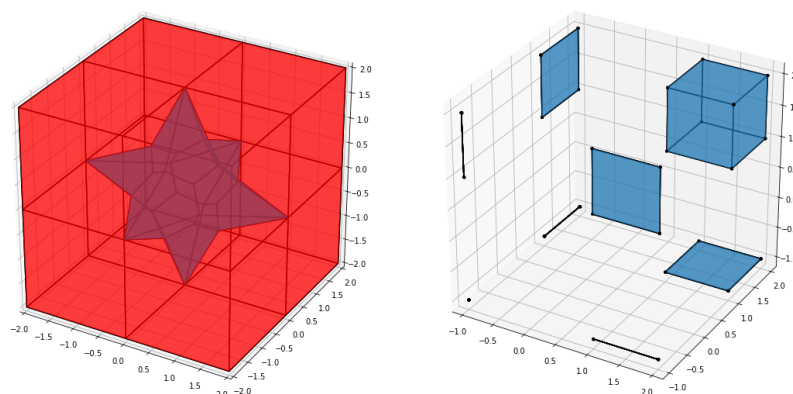


Рис. 2. Объединенные множества решений интервальных линейных систем
Fig. 2. United sets of solutions of interval linear systems

Поясним, что множество решений слева на рис. 2. является неограниченным. По этой причине оно было обрезано, и для понимания, где именно это произошло, плоскости обрезки окрашены в красный цвет.

5. Приложение

Рассмотрим практически важную задачу оценивания параметров электрохимического процесса формирования рыхлых осадков порошков металла. Постановка задачи вкратце такова: экспериментаторы имеют набор замеров, полученных в результате электрохимического процесса, и нуждаются в оценках его параметров, которые необходимы для решения вопроса об остановке процесса с высокой точностью по времени, основываясь на заранее полученных значениях.

Но получение оценок параметров процесса осложняется неточностью информации, полученной при измерении его характеристик, а также невозможностью применения традиционных вероятностных методов обработки погрешностей измерений (см. [16]). При небольшом количестве наблюдений стандартные методы теоретико-вероятностной математической статистики дают очень широкие доверительные интервалы оценок, так что применять их для отслеживания момента остановки процесса невозможно. Нам необходимо находить более точные двусторонние оценки состояния.

Для достижения поставленной цели можно использовать подходы интервального анализа, оперирующие с заданной оценкой максимального значения (по модулю) погрешностей. Они в самом деле позволяют получить искомые интервалы оценки состояний с более узкой шириной.

Для описания процесса формирования рыхлых осадков порошков металла была предложена следующая функция:

$$W(t) = \frac{D + t}{R + t}, \quad (3)$$

которая была получена эмпирическим способом [15, 16]. D, R – параметры, измеряемые в секундах, на которые налагаются условия $R > D$. Время (t , сек) – независимый аргумент.

Исходными данными для обработки являются шесть выборок измерений, по 19 измерений каждая:

$$\{t_{n,m}, W_{n,m}\}, \quad t_{n,m} \in [0, 2400], \quad n = 1, \dots, 19, \quad m = 1, \dots, 6.$$

На этапе предобработки выполняем очистку измерений от выбросов и затем формируем одну большую выборку $\{t_n, W_n\}_{n=1}^{100}$, пользуясь тем, что все моменты времени, в которые сделаны измерения, различны. В дальнейшем будем обозначать полученную общую выборку базовых измеренных значений [17] как $\{(t_n, W_n)\}_{n=1}^{100}$. Далее осуществляем интервализацию этих базовых значений для зависимой переменной W , добавляя уравновешенный интервал, который соответствует абсолютной погрешности ε :

$$W_n = W_n + [-\varepsilon, \varepsilon].$$

Моменты времени t_n , в которые производятся измерения, считаются заданными точно. Итак, получаем выборку интервальных данных $\{t_n, W_n\}_{n=1}^{100}$.

Для решения задачи восстановления зависимости по интервальным данным мы подставляем данные из нашей выборки в равенство (3), получая в результате интервальную систему уравнений вида

$$\begin{cases} \frac{D + t_n}{R + t_n} = W_n, \\ n = 1, 2, \dots, 100. \end{cases} \quad (4)$$

Далее нужно найти ее множество решений и выбрать из него точку, в случае пустоты этого множества решений, найти точку, на которой достигается максимальная совместность этой системы. Отметим, что так как интервальность присутствует лишь в правой части этой системы, то ее объединенное множество решений совпадает с допусковым множеством решений, и поэтому можно не различать их и говорить просто о «множестве решений».

Величина ε , вообще говоря, неизвестна точно, и в процессе обработки данных мы даже можем варьировать ее. Ее начальное значение было задано как $\varepsilon_{\text{init}} = 0.0175$. Но при этом множество решений системы (4) оказывается пустым. С помощью максимизации распознающего функционала множества решений системы (4) можно найти предельное значение для ε , при котором множество решений непусто, и оно равно $\varepsilon_{\text{crit}} = 0.0232$. Основываясь на этом результате, было принято решение задать для практического расчета ограничение с 10%-м запасом, т. е. $\varepsilon = 0.0255$.

В случае непустоты множества решений системы (4) часто бывают необходимы его оценки, в частности, внешняя оценка. Для решения этого вопроса можно воспользоваться тем обстоятельством, что система (4) может быть приведена к интервальной системе линейных алгебраических уравнений, имеющей следующий вид:

$$\begin{cases} (1, -W_n) \begin{pmatrix} D \\ R \end{pmatrix} = t_n \cdot (W_n - 1), \\ n = 1, 2, \dots, 100. \end{cases} \quad (5)$$

Найдя известными методами внешнюю оценку для ее множества решений, ответим на поставленный выше вопрос. Следует лишь иметь в виду, что в системе (5) интервальные параметры являются связанными, так как одновременно входят в матрицу и в правую часть. Как следствие, известные интервальные численные методы, ориентированные на решение ИСЛАУ без связей, с независимыми параметрами, могут давать в этой задаче довольно грубые оценки. По этой причине для получения более точных оценок лучше рассматривать задачу в ее изначальной дробно-линейной форме (4).

В общем случае для отыскания оценки параметров можно использовать метод максимума совместности, в котором оценкой берется точка максимума распознающего функционала множества решений системы (4) (см. [7, 8]). Так как независимые аргументы восстанавливаемой функции задаются точно, то неважно какой именно функционал будет использован, объединенного или допускового множеств решений: в данном случае они совпадают.

С помощью реализованных в библиотеке IntvalPy функций была выполнена глобальная оптимизация распознающего функционала Uni и найден его максимум, который оказался примерно равен 0.0018. Аргумент максимума функционала примерно равен $\arg \max Uni \approx \approx (170.69, 223.72) = (c_1, c_2)$. Интересно сравнить этот результат с оценкой параметров (180.7, 235.5), которая была получена для той же задачи в работе [18] с помощью метода наименьших квадратов.

Отметим, что независимый аргумент t (время) дважды встречается в функции (3) для описания физического процесса. Выше мы рассмотрели решение задачи восстановления зависимости в случае, когда t является точной величиной. Если же t известно неточно и задается интервалами t_n , то нужно озаботиться вычислением области значений отображения по $t \in t_n$. В частности, естественное интервальное расширение здесь неприменимо из-за неединственности вхождения переменной. Для преодоления этой проблемы можно видоизменить выражение и работать с его эквивалентной формой:

$$W(t) = \frac{D + t + R - R}{R + t} = 1 + \frac{D - R}{R + t}. \quad (6)$$

Опираясь на специфические свойства множества решений, изложим способ вычисления внешней оценки множества решений для частного случая. Заменим исходную задачу нахож-

дения внешнего бруса для объединенного множества решений $\Xi_{uni}(F, t, \mathbf{W})$ на эквивалентную совокупность подзадач. Каждая из них требует нахождения отдельных покомпонентных оценок множества решений $\min \{x_v \mid x \in \Xi_{uni}(F, t, \mathbf{W})\}$ и $\max \{x_v \mid x \in \Xi_{uni}(F, t, \mathbf{W})\}$, где $v = 1, 2$. Далее для определенности рассмотрим нахождение максимума первой компоненты, остальные оценки находятся аналогично. Для нахождения интервала изменения параметра D мы используем идею, которая аналогична основной идее методов дробления графика (см. [10, 19]): оценки находятся на основе информации о сечениях множества решений прямыми, параллельными координатным осям.

Для нахождения интервала D изменения параметра D выразим его из равенства (3), а затем подставим в полученную формулу интервальную оценку $\mathbf{W}$ и значение параметра R , равное c_2 и вычисленное ранее:

$$D = W \cdot (c_2 + t) - t.$$

Далее, подставляя в это соотношение моменты времени t_n и интервалы $\mathbf{W}_n$, где $n = 1, \dots, 100$, мы получим систему равенств

$$\begin{cases} D = W_n \cdot (c_2 + t_n) - t_n, \\ n = 1, 2, \dots, 100. \end{cases} \quad (7)$$

Из нее пересечением отдельных интервалов D находится начальная оценка $d = [d^{(0)}, d^{(1)}]$ среза множества решений, т. е. интервал, в котором варьируется величина D при фиксированном значении параметра R .

Учитывая вид множества решений, которое было получено с помощью перебора на сетке в работе [18], нетрудно понять, что если интервал d вырождается в точку, то мы достигли наибольшего или наименьшего значения соответствующей компоненты точек из множества решений, т. е. получили точную верхнюю или точную нижнюю оценку этой компоненты (рис. 3).

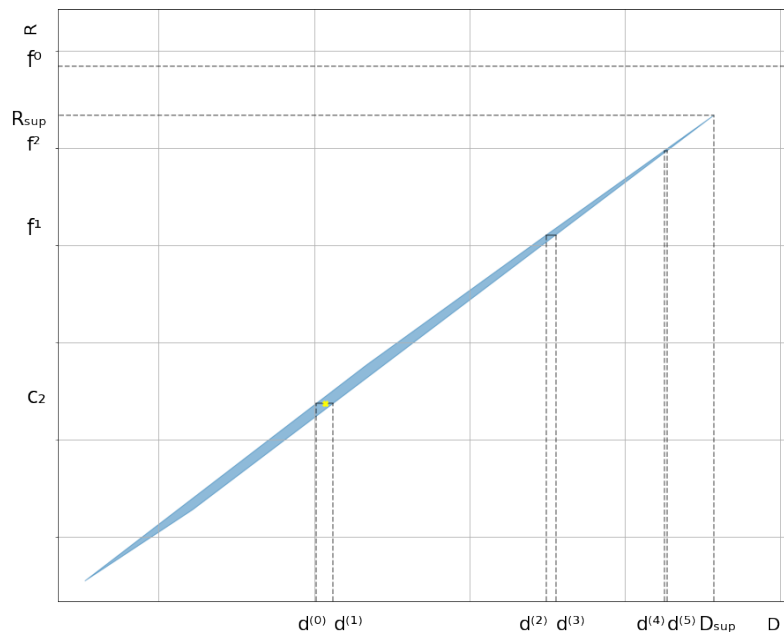


Рис. 3. Иллюстрация получения внешней оценки для значений первого измерения
Fig. 3. Illustration of obtaining an external estimate for the first measurements

Опишем более подробно алгоритм получения D_{sup} , т. е. наименьшего значения первой компоненты точек из множества решений. Сначала найдем какое-нибудь значение f^0 , при котором уравнение (7) не имеет решений. Проверить существование решений можно с помощью под-

становки выбранного f^0 в систему (7) и пересечения отдельных D . Если получаем непустой интервал, то система разрешима, иначе – неразрешима. После нахождения f^0 , доставляющего неразрешимость системе, мы, фактически, получили интервал $r = [c_2, f^0]$, где локализован максимум значений второй компоненты точек из множества решений. Организуя его дробление и проверку существования решений системы (7), мы сможем уточнить этот максимум. Более точно, станем последовательно дробить интервал r и его потомки пополам, подставляя их середины в уравнение (7) и проверяя пустоту/непустоту пересечений отдельных D .

Если на каком-то шаге последовательного дробления мы получаем непустой интервал (система (7) неразрешима), то отбрасываем ту часть интервала, где значения меньше его центральной точки («нижнюю» часть), иначе – «верхнюю». После этого переходим на следующий шаг с новым интервалом r . Описанные действия выполняются до тех пор, пока ширина интервала r больше заданной величины, или же пока не исчерпано выделенное на алгоритм количество шагов.

В целом, опираясь на высокоуровневую функциональность интервальной библиотеки IntvalPy, удалось распознать разрешимость системы (4) и реализовать алгоритм для поиска внешней оценки множества решений практически значимой задачи. Отметим, что построенный алгоритм является последовательно гарантирующим, что обеспечивает получение интервальной оценки параметров, даже если алгоритм будет прерван при исчерпании выделенного ему количества шагов (см. [10]).

6. Сравнение с другими библиотеками

Существует немало интервальных библиотек на самых разных языках программирования, среди которых C/C++, Matlab/Octave, Julia, Java и др. Для демонстрации возможностей созданной библиотеки IntvalPy был проведен качественный анализ ее функциональности, а также реализованы вычислительные тесты для сравнения с другими модулями на языке Python.

Целью вычислительных тестов является сравнение быстродействия кода по отношению к другим реализациям. Для проведения подобного рода экспериментов был реализован простейший алгоритм интервальной глобальной оптимизации, и его работа рассматривалась только на библиотеках, доступных на языке Python. Из немалого количества интервальных библиотек, написанных к настоящему моменту для языка Python, практический интерес имеют, как нам кажется, только две – PyInterval [20], PyIbex. Возможность использовать другую реализацию интервального типа на одном и том же языке позволяет корректно сравнить вычислительную эффективность.

В качестве тестовых задач глобальной оптимизации были рассмотрены целевые функции различной сложности, имеющие много локальных экстремумов, с различной трудоемкостью решения. В первую очередь решались задачи безусловной оптимизации, в частности, использовались тесты из работы [21]. Табл. 3 демонстрирует полученные результаты.

Сравнивая результаты, полученные различными интервальными пакетами, видим, что модуль IntvalPy, в целом, работает медленнее на интервальных операциях, чем другие библиотеки. Это связано с наличием дополнительных проверок типов при определении операций, при создании интервалов, а также использование типа данных повышенной точности вычислений. Например, в интервальном классе модуля IntvalPy всякая арифметическая операция проверяет, к какой интервальной арифметике относится объект, что влияет на дальнейшее определение операций с ним. В будущем, возможно, стоит ожидать интеграции модуля IntvalPy с языком C/C++ для дальнейшей модификации интервального класса, что должно в существенной степени увеличить скорость выполнения кода.

Помимо скорости решения задач оптимизации стоит обратить внимание на тот факт, что некоторые задачи так и не были до конца решены из-за исчерпания выделенного количества итераций. Однако, если такая ситуация возникала при выполнении в одной из библиотек,

Таблица 3

Численные результаты безусловной глобальной оптимизации

Table 3

Numerical Results of Unconditional Global Optimization

IntervalPy	ext.	0,517	0,278	1,637	0,316	0,157	0,217	57,32	1,137	68,18	0,468
	t_{cpu}	0,107	0,061	0,373	0,079	0,029	0,061	15,94	0,311	11,57	0,097
	$\in$	$8,90^{-15}$	$6,55^{-15}$	$5,66^{-15}$	$7,90^{-15}$	$8,73^{-15}$	$4,44^{-15}$	$1,06^{-15}$	$7,11^{-15}$	$2,83^{-15}$	$8,57^{-15}$
	N_{it}	142	66	342	146	80	70	10000	129	10000	68
PyIbex	t_{cpu}	0,047	0,016	0,087	0,035	0,011	0,016	5,450	78,03	3,897	4,085
	$\in$	$8,88^{-15}$	$6,66^{-15}$	$5,66^{-15}$	$7,90^{-15}$	$8,73^{-15}$	$5,55^{-15}$	$1,06^{-15}$	$4,26^{-14}$	$2,83^{-7}$	$3,14^{-14}$
	N_{it}	164	66	342	146	80	63	10000	10000	10000	10000
PyInterval	t_{cpu}	0,301	0,149	0,833	0,214	0,070	0,116	246,7	42,73	27,30	30,57
	$\in$	$7,10^{-15}$	$6,66^{-15}$	$5,66^{-15}$	$7,90^{-15}$	$8,73^{-15}$	$8,10^{-15}$	$1,06^{-15}$	$4,26^{-14}$	$2,83^{-7}$	$1,70^{-14}$
	N_{it}	160	66	342	146	80	63	10000	10000	10000	10000
	dim	2	2	2	2	3	2	2	4	2	2
	F	Branin (RC)	Bohachevsky (B_m)	Beale (BL)	Booth (BO)	De Jong (DJ)	Easom (ES)	Eggholder	Rastrigin (RT_n)	Schaffer № 4	Levy № 13

Примечание. F – целевая функция, dim – размерность задачи, N_{it} – ограниченное количество итераций для остановки оптимизации, $\in$ – минимальная допустимая погрешность, t_{cpu} – усредненное 50-ю запусками время выполнения задачи, double – двойная точность, extended – повышенная точность.

то не очевидно, что подобная ситуация возникнет и в другой реализации. По-видимому, большое различие в трудоемкости решения одних и тех же задач различными библиотеками связано с тем, что некоторые из них используют направленные округления, результатами которых являются более широкие интервалы по сравнению с интервалами, полученными с повышенной точностью. В нелинейных задачах оптимизации, где рассматриваются сложные целевые функции с большим количеством локальных экстремумов, это обстоятельство может потребовать большего количества итераций для завершения алгоритма.

Целью качественного анализа, представленного в табл. 4, является сравнение возможностей, которые предоставляют другие интервальные библиотеки. Для этого рассмотрим подробнее их функциональность, точнее, такие ее аспекты, как возможность работы с различными интервальными арифметиками, возможность выполнения матричных операций, наличие гото-

вых инструментов для визуализации решений, наличие встроенных процедур для вычисления внешних и/или внутренних оценок параметров линейных систем, нахождение решений уравнений, а также возможность решения задач оптимизации.

Таблица 4

Качественное сравнение интервальных библиотек

Table 3

Qualitative Comparison of Interval Libraries

Библиотека	Арифм.	Мат. операции	Визуал.	Оценка парам.	Решение уравнений	Оптим.
IntvalPy	+	+	+	+	+	+
PyInterval	+	—	—	—	+	—
PyIbex	—	+	+	—	+	+
Octave ²	—	+	—	+	+	—
WM ³	+	+	—	+	—	—
JInterval	+	+	+	+	+	+
JI ⁴	—	+	+	—	+	+
IntLab	+	+	—	+	+	+

Стоит отметить, что инструменты для нахождения решений нелинейных систем уравнений в библиотеке IntvalPy хотя и присутствуют, но не очень развиты (как и в некоторых других библиотеках). Реализованные алгоритмы являются простейшими, в отличие от методов для линейных систем, среди которых присутствуют наиболее современные и развитые на данный момент численные методы.

Заключение

В работе были изложены особенности разработанной библиотеки IntvalPy, главные идеи, положенные в ее основу, а также ее архитектура и реализация. Объяснены преимущества и достоинства библиотеки, в число которых входят востребованные сегодня кроссплатформенность, скорость вычислений и наиболее современные алгоритмы.

Библиотека позволяет учитывать специфические запросы пользователей, например возможность отключения определенных опций в случае необходимости. Другой важной характеристикой созданного модуля является его практическая ориентированность – наличие в его составе готовых численных методов для решения различных прикладных задач интервального анализа. Кроме того, библиотека IntvalPy продолжает развиваться и в нее добавляется новый инструментарий.

Список литературы

1. Андросов А. С., Шарый С. П. Библиотека IntvalPy. Программный репозиторий // URL: <https://github.com/AndrosovAS/intvalpy>
2. Kearfott R. B., Nakao M., Neumaier A., Rump S. M. et al. Standardized notation in interval analysis // *Reliable Computing*. 2010, vol. 15, no. 1, pp. 7–13.
3. Rump S. M. Verification methods: Rigorous results using floating-point arithmetic // *Acta Numerica*. 2010, vol. 19, pp. 287–449.

² Подразумевается интервальная библиотека на свободно распространяемой программной системе GNU Octave.

³ Интервальные возможности проприетарной системы Wolfram Mathematica.

⁴ JuliaIntervals.

4. Хансен Э., Уолстер Дж. У. Глобальная оптимизация с помощью методов интервального анализа // Институт компьютерных исследований. М., 2012, 516 с.
5. Sharaya I. A. Boundary intervals and visualization of AE-solution sets for interval systems of linear equations // *Reliable Computing*. 2012, pp. 435–467.
6. Шарая И. А. Метод граничных интервалов для визуализации полиэдральных множеств решений // *Вычислительные технологии*. 2015, Т. 20, С. 75–103.
7. Шарый С. П. Разрешимость интервальных линейных уравнений и анализ данных с неопределенностями // *Автоматика и телемеханика*. 2012, С. 111–125.
8. Шарый С. П. Сильная согласованность в задаче восстановления зависимостей при интервальной неопределенности данных // *Вычислительные технологии*. 2017, Т. 22, С. 150–172.
9. Fiedler M., Nedoma J., Ramik J., Rohn J., Zimmermann K. Linear optimization problems with inexact data // Springer Science+Business Media. New York, 2006, 223 p.
10. Шарый С. П. Конечномерный интервальный анализ // Новосибирск: XYZ, 2022, 654 с.
11. Neumaier A. Interval methods for systems of equations // Cambridge. New York, 1990, 270 p.
12. Стецюк П. И. Субградиентные методы `ralgb5` и `ralgb4` для минимизации овражных выпуклых функций // *Вычислительные технологии*. 2017, Т. 22, С. 127–149.
13. Shary S. P. On optimal solution of interval linear equations // *SIAM Journal on Numerical Analysis*. 1995, pp. 610–630.
14. Шарая И. А. Пакет `IntLinIncR3`. Руководство пользователя // 2014, 26 с, URL: <http://www.nsc.ru/interval/Programing/MCodes/IntLinIncR3.pdf>
15. Ostanina T. N., Rudoi V. M., Patrushev A. V., Darintseva A. B., Farlenkov A. S. Modelling the dynamic growth of copper and zinc dendritic deposits under the galvanostatic electrolysis conditions // *Journal of Electroanalytical Chemistry*. 2015, vol. 750, pp. 9–18.
16. Ostanina T. N., Rudoy V. M., Nikitin V. S., Darintseva A. B., Zalesova O. L., Porotnikova N. M. Determination of the surface of dendritic electrolytic zinc powders and evaluation of its fractal dimension // *Russian Journal of Non-Ferrous Metals*. 2017, vol. 57, pp. 47–51.
17. Баженов А. Н., Жилин С. И., Кумков С. И., Шарый С. П. Обработка и анализ данных с интервальной неопределенностью // 2022, 242 с, URL: <http://www.nsc.ru/interval/Library/AplBooks/InteDataProcessing.pdf>
18. Kumkov S. I., Nikitin V. S., Ostanina T. N., Rudoy V. M. Interval processing of electrochemical data // *Computational and Applied Mathematics*. 2020, vol. 380, 7 p.
19. Shary S. P. Graph subdivision methods in interval global optimization // *Studies in Computational Intelligence*. 2014, vol. 539, pp. 153–170.
20. Taschini S. Interval Arithmetic: Python Implementation and Applications // *Proceedings of the 7th Python in Science Conference (SciPy 2008)*. 2008, pp. 16–22.
21. Hedar A. R., Ahmed A. Studies on metaheuristics for continuous global optimization problems // Kyoto University, Kyoto, Japan. 2004, 148 p.
22. Nadezhin D. Yu., Zhilin S. I. Interval Library: Principles, Development, and Perspectives // *Reliable Computing*. 2014, vol. 19, no. 3, pp. 229–247.
23. Shary S. P. Interval regularization for imprecise linear algebraic equations // 2018, 21 p., URL: https://www.researchgate.net/publication/328063512_Interval_regularization_for_imprecise_linear_algebraic_equations
24. Oettli W., Prager W. Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides // *Numerische Mathematik*. 1964, vol. 6, pp. 405–409.
25. Jr D. M., Tortelli L. M., Finger A. F., Loret A. B. KaucherPy: interval package for Kaucher arithmetic // 2019, 12 p.

References

1. **Androsov A. S., Shary S. P.** The IntvalPy library [Online]. URL: <https://github.com/AndrosovAS/intvalpy>.
2. **Kearfott R. B., Nakao M., Neumaier A., Rump S. M. et al.** Standardized notation in interval analysis. *Reliable Computing*, 2010, vol. 15, no. 1, pp. 70–13.
3. **Rump S. M.** Verification methods: Rigorous results using floating-point arithmetic. *Acta Numerica*, 2010, vol. 19, pp. 287–449.
4. **Hansen E., Walster G. W.** Global Optimization Using Interval Analysis: Revised And Expanded CRC Press; 2nd edition (December 19, 2003). 2003, 728 p.
5. **Sharaya I. A.** Boundary intervals and visualization of AE-solution sets for interval systems of linear equations. *Reliable Computing*, 2012, pp. 435–467.
6. **Sharaya I. A.** Boundary interval method for visualization of polyhedral sets of solutions. *Reliable Computing*, 2015, vol. 20, pp. 75–103. (in Russ.)
7. **Shary S. P.** Solvability of interval linear equations and data analysis with uncertainties. *Automatics and Telemechanics*, 2012, pp. 111–125. (in Russ.).
8. **Shary S. P.** Strong compatibility in the problem of dependency recovery under interval uncertainty. *Reliable Computing*, 2017, vol. 22, no. 2, pp. 150–172. (in Russ.).
9. **Fiedler M., Nedoma J., Ramik J., Rohn J., Zimmermann K.** Linear optimization problems with inexact data Springer Science+Business Media. New York, 2006. 223 p.
10. **Shary S. P.** Finite-dimensional interval analysis Novosibirsk: Publishing «XYZ», 2022. 654 p. (in Russ.).
11. **Neumaier A.** Interval methods for systems of equations Cambridge. New York, 1990. 270 p.
12. **Stetsyuk P. I.** Subgradient methods ralgb5 and ralgb4 for minimizing convex functions. *Reliable Computing*, 2017, vol. 22, no. 2, pp. 127–149. (in Russ.).
13. **Shary S. P.** On optimal solution of interval linear equations. *SIAM Journal on Numerical Analysis*, 1995, pp. 610–630.
14. **Sharaya I. A.** The IntLinIncR3 package. User's Guide [Online]. 2014. 26 p. URL: <http://www.nsc.ru/interval/Programing/MCodes/IntLinIncR3.pdf>. (in Russ.)
15. **Ostanina T. N., Rudoi V. M., Patrushev A. V., Darintseva A. B., Farlenkov A. S.** Modelling the dynamic growth of copper and zinc dendritic deposits under the galvanostatic electrolysis conditions. *Journal of Electroanalytical Chemistry*, 2015, vol. 750, pp. 9–18.
16. **Ostanina T. N., Rudoy V. M., Nikitin V. S., Darintseva A. B., Zalesova O. L., Porotnikova N. M.** Determination of the surface of dendritic electrolytic zinc powders and evaluation of its fractal dimension Russian. *Journal of Non-Ferrous Metals*, 2017, vol. 57, pp. 47–51.
17. **Bazenov A. N., Zhilin S. I., Kumkov S. I., Shary S. P.** Processing and analysis of data with interval uncertainty [Online]. 2022, 242 p. URL: <http://www.nsc.ru/interval/Library/AplBooks/InteDataProcessing.pdf>. (in Russ.)
18. **Kumkov S. I., Nikitin V. S., Ostanina T. N., Rudoy V. M.** Interval processing of electrochemical data. *Computational and Applied Mathematics*, 2020, vol. 380, 7 p.
19. **Shary S. P.** Graph subdivision methods in interval global optimization. *Studies in Computational Intelligence*, 2014, vol. 539, pp. 153–170.
20. **Taschini S.** Interval Arithmetic: Python Implementation and Applications. Proceedings of the 7th Python in Science Conference (SciPy 2008), 2008, pp. 16–22.
21. **Hedar A. R., Ahmed A.** Studies on metaheuristics for continuous global optimization problems. Kyoto: Kyoto University, 2004. 148 p.
22. **Nadezhin D. Yu., Zhilin S. I.** JInterval Library: Principles, Development, and Perspectives. *Reliable Computing*, 2014, vol. 19, no. 3, pp. 229–247.

23. **Shary S. P.** Interval regularization for imprecise linear algebraic equations [Online]. 2018. 21 p. URL: https://www.researchgate.net/publication/328063512_Interval_regularization_for_imprecise_linear_algebraic_equations.
24. **Oettli W., Prager W.** Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides. *Numerische Mathematik*, 1964, vol. 6, pp. 405–409.
25. **Jr D. M., Tortelli L. M., Finger A. F., Loret A. B.** KaucherPy: interval package for Kaucher arithmetic. 2019. 12 p.

Информация об авторах

Артём Станиславович Андросов, аспирант

Сергей Петрович Шарый, доктор физико-математических наук

Information about the Authors

Androsov Artem Stanislavovich, postgraduate student

Shary Sergey Petrovich, Doctor of Sciences (Physics and Mathematics)

*Статья поступила в редакцию 03.12.2022;
одобрена после рецензирования 16.01.2023; принята к публикации 16.01.2023*

*The article was submitted 03.12.2022 ;
approved after reviewing 16.01.2023; accepted for publication 16.01.2023*

Научная статья

УДК 004.822:004.89

DOI 10.25205/1818-7900-2022-20-4-24-38

Модель представления онтологии предметных областей на основе графовых баз данных

Владислав Александрович Лисин¹, Алексей Сергеевич Серый²
Елена Анатольевна Сидорова³

¹Новосибирский государственный университет
Новосибирск, Россия

^{2,3}Институт систем информатики им. А.П. Ершова СО РАН
Новосибирск, Россия

¹vladlisin2@gmail.com

²alexey.seryj@iis.nsk.su

³lsidorova@iis.nsk.su

Аннотация

В статье представлен подход к моделированию онтологий предметных областей на основе графовых баз данных. Онтология традиционно рассматривается как средство изучения и формализации предметной области. На основе онтологий формируются базы знаний информационных систем, которые в дальнейшем можно пополнять и использовать для исследования определенных прикладных аспектов предметной области. В то же время с развитием технологий NoSQL и графовых баз данных, направленных на оптимизацию работы со связанными данными, появляется возможность проектирования хранилища данных без строгой предустановленной модели предметной области. В силу очевидной графовой природы онтологий графовые базы данных выступают перспективным решением для хранения и работы с онтологиями. Тем не менее при всем сходстве онтологии и модели данных применяемые в графовых СУБД не так просто объединить. В работе предложен подход к интеграции онтологической и графовой модели представления знаний и созданию веб-редактора онтологий. Описываются визуальные средства редактирования онтологий, приведен алгоритм послойной отрисовки ориентированного графа, описан механизм генерации динамических форм для редактирования классов и экземпляров онтологии, а также реализована базовая поддержка паттернов онтологического проектирования.

Ключевые слова

онтология, редактор онтологии, графовая база данных, визуализация графа, алгоритм послойной отрисовки направленного графа, редактор паттернов содержания

Для цитирования

Лисин В. А., Серый А. С., Сидорова Е. А. Модель представления онтологии предметных областей на основе графовых баз данных // Вестник НГУ. Серия: Информационные технологии. 2022. Т. 20, № 4. С. 24–38. DOI 10.25205/1818-7900-2022-20-4-24-38

© Лисин В. А., Серый А. С., Сидорова Е. А. 2022

Domain Ontology Representation Model Based on Graph Databases

Vladislav A. Lisin¹, Aleksey S. Sery²,
Elena A. Sidorova³

¹Novosibirsk State University
Novosibirsk, Russian Federation

^{2,3}Ershov Institute of Informatics Systems SB RAS
Novosibirsk, Russian Federation

¹vladlisin2@gmail.com

²alexey.seryj@iis.nsk.su

³lsidorova@iis.nsk.su

Abstract

The article presents an approach to modeling domain ontologies based on graph databases. Ontology is traditionally considered as a means of studying and formalizing the subject area. Based on ontologies, knowledge bases of information systems are formed, which can later be replenished and used to study certain applied aspects of the subject area. At the same time, with the development of NoSQL technologies and graph databases aimed at optimizing work with related data, it becomes possible to design a data warehouse without a strict pre-established domain model. Due to the obvious graph nature of ontologies, graph databases are a promising solution for storing and working with ontologies. However, with all the similarities, the ontology and data models used in graph DBMSs are not so easy to combine. The paper proposes an approach to the integration of ontological and graph models of knowledge representation and considers its application in creating a database for a prototype of a web ontology editor. Visual tools for editing ontologies are described, an algorithm for layer-by-layer rendering of a directed graph is given, a mechanism for generating dynamic forms for editing classes and ontology instances is described, and basic support for ontological design patterns is implemented.

Keywords

ontology, ontology editor, graph database, graph visualization, directed graph layer-by-layer drawing algorithm, content pattern editor

For citation

Lisin V. A., Sery A. S., Sidorova E. A. Domain Ontology Representation Model Based on Graph Databases. *Vestnik NSU. Series: Information Technologies*, 2022, vol. 20, no. 4, pp. 24–38. (in Russ.) DOI 10.25205/1818-7900-2022-20-4-24-38

Введение

Онтология традиционно рассматривается как средство изучения и формализации предметной области. На основе онтологий формируются базы знаний информационных систем, которые в дальнейшем можно пополнять и использовать для исследования определенных прикладных аспектов предметной области. Описание декларативных знаний в виде классов, объектов и отношений между ними формирует семантический базис для коммуникации между человеком и компьютерными агентами. Для формального представления онтологий традиционно применяются модель RDF и язык OWL [1], позволяющие хранить данные в виде множества RDF-триплетов, используя для этого одно из известных хранилищ, например Jena TDB [2] или Virtuoso [3], и язык запросов SPARQL. При всех своих достоинствах хранилища триплетов часто проигрывают в скорости обработки запросов популярным СУБД, вынуждая разработчиков информационных систем, основанных на онтологиях, искать альтернативные пути представления своих данных, используя, в том числе, реляционные и графовые СУБД [4].

Технологии традиционных реляционных баз данных сфокусированы на оптимальной организации данных для повышения эффективности их хранения, управления и поиска. С другой стороны, набирающая популярность технология графовых баз данных представляет собой решение, ориентированное на поэлементный просмотр компонентов графа, понимание их взаимосвязей, а не на пакетную обработку данных. Несмотря на взаимодополняемость и взаимосвязь между этими технологиями их полная интеграция и взаимодействие еще не стандар-

тизованы, в том числе по причине отсутствия формализма [5], описывающего схемы графовых баз данных и обеспечивающего совместимость данных, а также задания соответствия между графовой моделью БД, выраженной этим формализмом, и стандартной онтологической моделью. В силу очевидной графовой природы онтологий графовые базы данных выступают перспективным решением для хранения и работы с онтологиями. В то же время при всем сходстве структуры NoSQL баз данных RDF и NoSQL графовых баз данных относятся к разным графовым моделям, которые не так просто объединить.

Графовые базы данных предоставляют более обобщенную структуру, основанную на модели LPG – Labeled Property Graphs (графы меток и атрибутов) [6]. В данной структуре узлы и дуги могут обладать персональными именами (метки и типы соответственно), а также способны хранить свойства, представленные в виде полей «ключ – значение». Также LPG предоставляет возможность создания множества дуг между узлами графа, наделяя каждую дугу уникальным идентификатором. Основными отличиями между LPG и RDF являются:

- атомарность узлов RDF: возможность хранения информации непосредственно в узлах LPG обеспечивает более компактную структуру, уменьшая общее число необходимых узлов в десятки раз;
- RDF не способен находить различие между встречающимися одинаковыми связями между одинаковыми элементами, в то время как LPG предоставляет уникальные идентификаторы дуг и узлов;
- RDF не поддерживает прикрепление свойств к экземплярам связей (данное ограничение можно преодолеть при помощи трансформации экземпляра связи в объект, но это влечет ухудшение читабельности);
- RDF поддерживает атрибуты, принимающие множество значений (данное ограничение преодолевается в LPG путем хранения данных атрибута в соответствующем векторе);
- в LPG есть только один вид узлов, а в RDF-графах – два (URI или литеральные значения для объектов триплетов).

В табл. 1 представлено сопоставление терминов, используемых для обозначения схожих концептов в базах данных, онтологиях и LPG.

Таблица 1

Сопоставление терминов баз данных, онтологий и LPG

Table 1

Matching Database, Ontology, and LPG Terms

Базы данных	Онтологии	LPG
Экземпляр	Экземпляр	Узел / Дуга
Значение	Значение	Значение
Атрибут	Свойство – значение (Datatype Property)	Свойство
Отношение	Свойство – объект (Object Property)	Дуга (тип)
Таблица	Класс	Узел (метка)

В данной работе предлагается модель представления онтологий предметных областей на основе графовых баз данных, а также приведено описание подхода к разработке инструмента для их визуализации и редактирования.

Оставшаяся часть статьи организована следующим образом. В разделе 1 представлен метод организации хранения онтологий и графов знаний в графовых базах данных. Раздел 2 посвящен описанию алгоритма послойной отрисовки графа на плоскости для визуального представления онтологий. В разделе 3 приведено описание редактора онтологий, включающего конструктор паттернов онтологического проектирования и систему полнотекстового поиска и фильтрации.

1. Представление онтологии в формате графовой базы данных

В областях научной деятельности, связанных с компьютерными технологиями, часто используется общее определение онтологии Т. Грубера [7]. Грубер описывает онтологию как полную спецификацию концептуализации. Она представляет собой описание предметной области (концептуализации), состоящее из системы понятий, определений и описаний предметов, а также методов исследования. Основные составляющие онтологии – это сущности (классы, концепты) предметной области, отношения между классами и экземпляры классов, также называемые объектами.

На 12-й Европейской конференции по искусственному интеллекту в 1996 г. А. Бернарас, И. Ларесгоити и Х. Корера в своей работе [8] предложили метод построения онтологии предметной области электрических сетей. Работа была проделана в рамках проекта Esprit KACTUS (KACTUS is a European ESPRIT project). Одной из ключевых задач проекта являлась проверка возможности повторного использования знаний в сложных технических системах, а также использование онтологий для поддержки данного процесса. Метод заключался в гибком построении онтологий – архитектура и требования, накладываемые на онтологию, были тесно связаны с разработкой программного обеспечения. При разработке новой программы параллельно шел процесс построения онтологии для представления знаний, применяемых при разработке. Онтологии такого вида могут быть построены путем повторного использования фрагментов других онтологий, а затем, в свою очередь, могут применяться как составные части онтологий, которые будут разрабатываться позднее. Таким образом, была предложена методология разработки онтологий, использующая уже созданные онтологии и/или их фрагменты.

В данной работе была разработана модель представления онтологии в виде ориентированного помеченного графа (Labeled Property Graph, LPG), совместимого с моделью представления данных графовой БД Neo4j, обеспечивающая поддержку рассмотренной выше методологии. Предлагаемая модель обеспечивает возможность одновременного хранения нескольких онтологий в одном графе, а также заимствования элементов существующих онтологий при создании новых онтологий без необходимости дублировать данные.

1.1. Метки вершин в графовом представлении онтологии

Элементы онтологий, такие как классы, свойства, экземпляры и т. д. обладают уникальными идентификаторами (URI), позволяющими однозначно идентифицировать их принадлежность. В графовом представлении (LPG) эту функцию берут на себя метки. Множество меток вершины графа также включает пространство имен соответствующей онтологии, что позволяет хранить множество онтологий в одном графе. При этом одни и те же вершины могут одновременно быть частью более чем одной онтологии.

Уникальной меткой вершины является ее идентификатор URI. Идентификаторы новых элементов создаются простой конкатенацией пространства имен онтологии и уникального токена, полученного функцией генерации случайной строки. Далее будут приведены примеры основных меток и типов, используемых для описания вершин в графовом представлении онтологий.

На рис. 1 изображен пример проекции owl-описания онтологий и их классов в графовое представление.

На рис. 1 показаны графовые представления двух онтологий Ex1 и Ex2. Для их обозначения на графе созданы вершины с метками **Ontology**. Каждая онтология состоит из одного класса с именем C1. Как можно видеть из рис. 1, вершина C1, представляющая класс онтологии Ex1, имеет метку **Class**, метку <http://www.ex1.com>, указывающую на принадлежность класса онтологии Ex1 и метку <http://www.ex1.com/c1> – URI класса. Аналогичным образом устроена вершина, соответствующая классу онтологии Ex2. Несмотря на идентичность имен двух

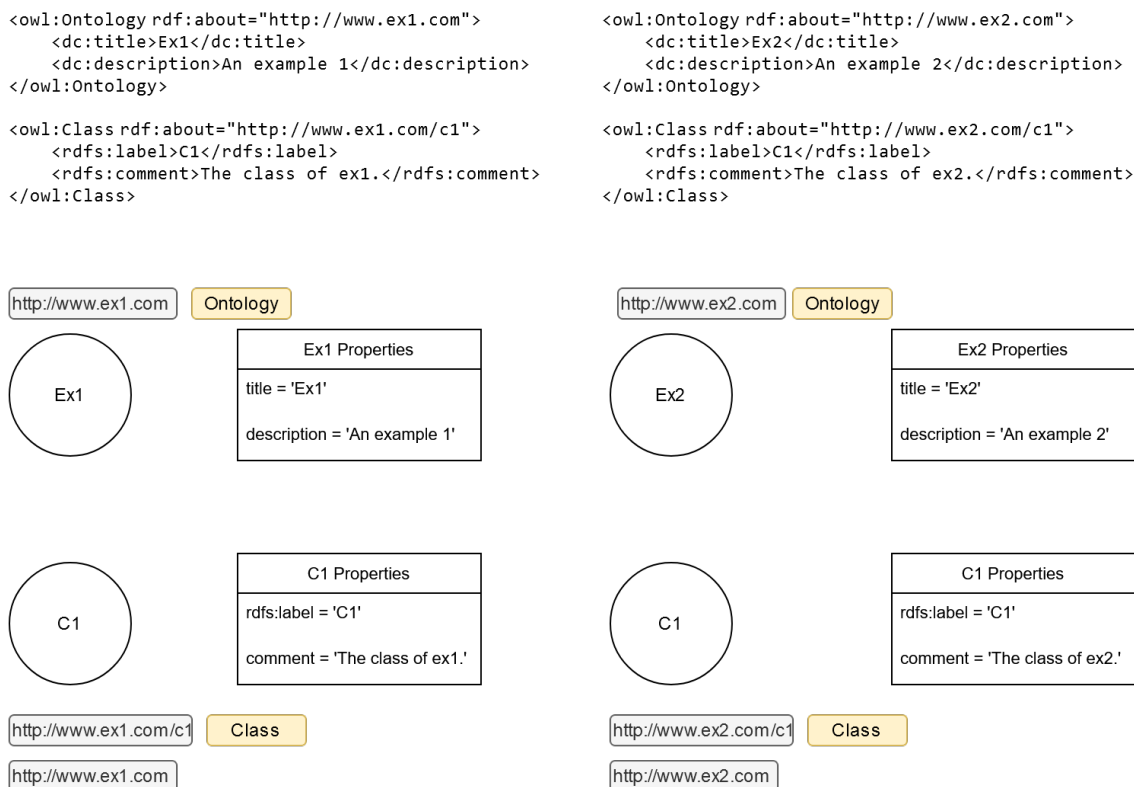


Рис. 1. Пример описания онтологий и их классов в графовой БД
 Fig. 1. An example of describing ontologies and their classes in a graph database

классов на рис. 1, им соответствуют две разные вершины графа. Их различие заключается в принадлежности двум разным онтологиям, о чем свидетельствуют метки пространства имен и URI. На рисунке приведены примеры основных меток, применяемых для типизации узлов – **Ontology** и **Class**. Данные метки используются для определения основных типов элементов графа: главный узел онтологии, классы, объекты и узлы атрибутов классов. Использование данных меток упрощает последующие запросы к графовой базе данных.

Для корректного отображения в пользовательском интерфейсе каждый элемент онтологии обладает атрибутом `rdfs:label`. Данный атрибут заимствован из OWL, и его значениями, как правило, являются корректные названия элемента на всех необходимых языках.

1.2. Иерархия онтологий

Разработанная модель представления позволяет выстраивать онтологии в иерархические структуры. Данное решение избавляет пользователя от необходимости создания каждой онтологии с нуля, предоставляя возможность использовать созданную ранее уже готовую онтологию, расширив ее новыми элементами или, наоборот, взяв ее подмножество. Модель различает базовые и наследуемые онтологии. Онтология, которая не наследуется от других, называется корневой. Метки **Secondary** и **Exclude** присваиваются расширенным и ограниченными онтологиям соответственно. Каждый граф, представляющий онтологию, имеет главную вершину, в которой содержится информация об онтологии в целом, в том числе об имеющихся иерархических связях с другими онтологиями. Подобный подход позволяет расширять и ограничивать уже существующие онтологии без необходимости полного дублирования их содержания. Помимо главной вершины, метками **Secondary** или **Exclude** помечаются и все классы соответствующей онтологии.

При расширении онтологии пользовательскими классами каждому новому элементу присваивается метка **Secondary**. При ограничении некоторой онтологии O , т. е. порождении из ее подмножества онтологии $O1$, каждому элементу O , который мы хотим исключить из $O1$, присваивается метка, составленная как конкатенация URI онтологии $O1$ и ключевого слова **Exclude**. Таким образом, порождается онтология, являющаяся подмножеством существующей. На рис. 2 приведен пример расширения корневой онтологии.

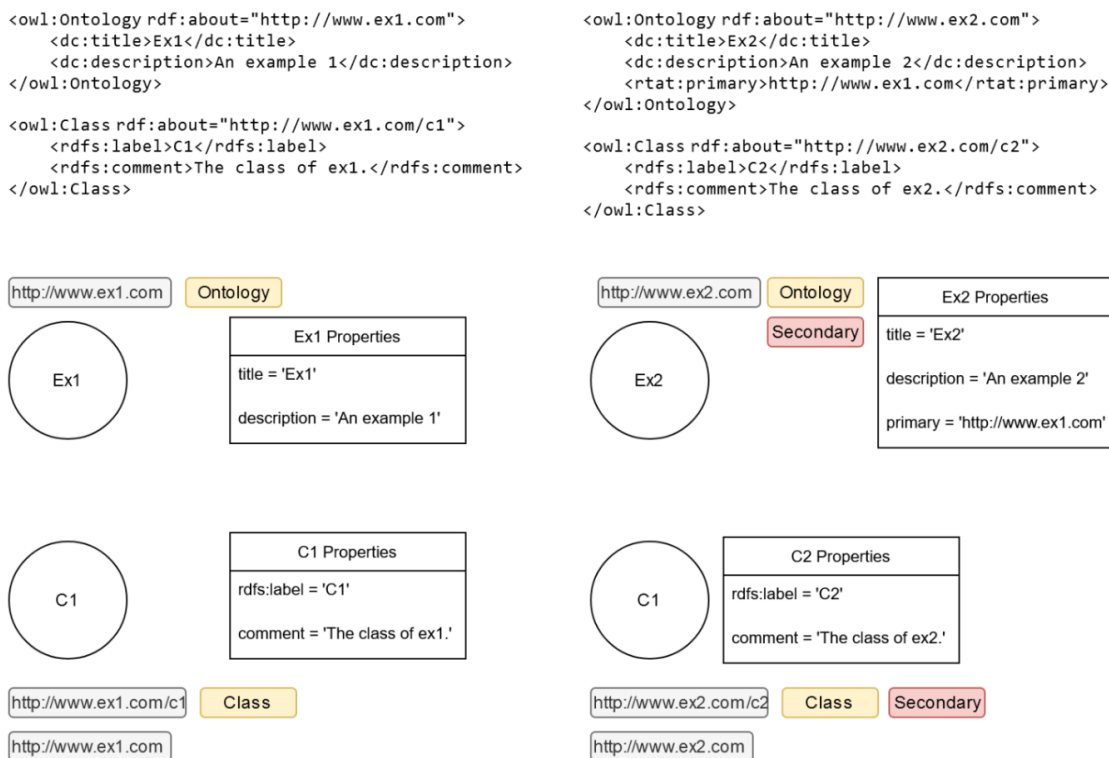


Рис. 2. Пример наследования онтологией элементов корневой онтологии с использованием меток типа Secondary на графе

Fig. 2. An example of ontology inheritance of elements of the root ontology using labels of the “Secondary” type on the graph

Здесь узлу Ex2 присвоен атрибут primary, в котором записан URI онтологии, от которой наследуется онтология Ex2, т. е. онтологии Ex1. Таким образом, при поиске элементов онтологии Ex2 к запросу будут прикреплены все элементы онтологии Ex1 – в частности класс C1. Метки Secondary используются для упрощения фильтрации запросов к графовой БД.

1.3. Свойства классов

Класс в разработанной модели представляет собой узел в графе, помеченный основной меткой **Class**. Для задания и отображения иерархии классов используются дуги с основным типом **subClassOf**.

Представлением атрибута класса является узел, наделенный основной меткой **DatatypeProperty**. URI атрибута задается типом дуги, связывающей узел **DatatypeProperty** и узел класса. Связи класса представляются вершинами с меткой **ObjectProperty**. URI отношения задается типом дуги, связывающей узел **ObjectProperty** и узел класса.

На рис. 3 приведен пример иерархии классов C1 и C3 и описание **DatatypeProperty** и **ObjectProperty** класса C3.

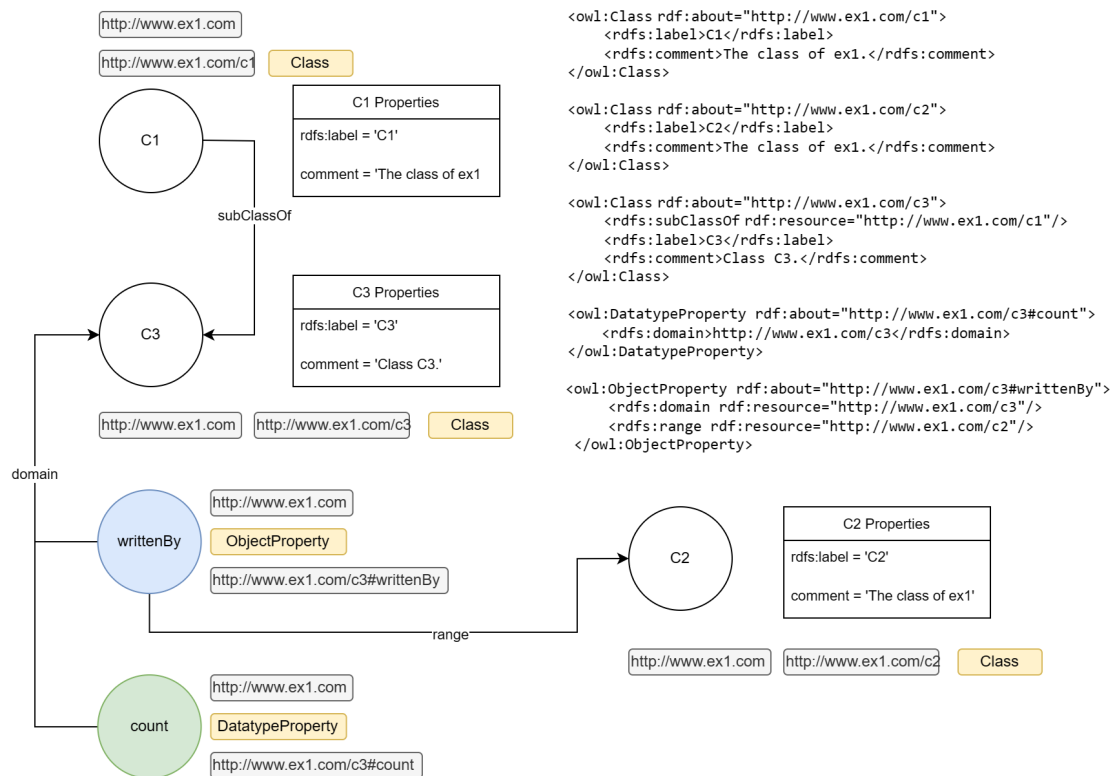


Рис. 3. Пример реализации иерархии классов онтологии и их атрибутов на графе
Fig. 3. An example of the implementation of the hierarchy of ontology classes and their attributes on a graph

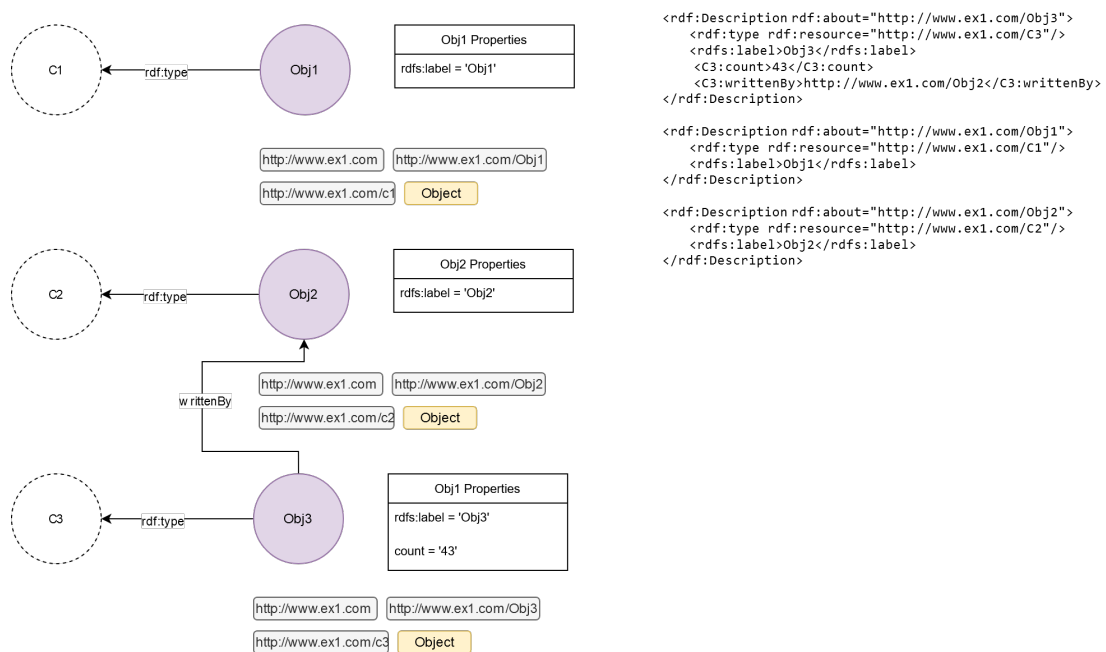


Рис. 4. Пример реализации экземпляров классов C1, C2 и C3 онтологии на графе
Fig. 4. An example of the implementation of instances of classes C1, C2 and C3 of an ontology on a graph

1.4. Свойства экземпляров классов

Каждый экземпляр класса онтологии на графе представляется узлом с меткой **Object**. Каждый такой объект связан со своим классом дугой типа **rdf:type**. Атрибуты объекта хранятся в структуре `dict` узла объекта. Связи объекта строятся через дуги соответствующих типов, определенные в классе объекта. На основе примера, изображенного на рис. 3, на рис. 4 приведен пример экземпляров классов C1, C2 и C3.

2. Визуализация онтологии

Визуализация является неотъемлемой частью процесса исследования и анализа онтологий предметных областей. Корректная и наглядная визуализация предоставляет инженеру знаний средства для эффективного управления, просмотра и навигации по онтологиям, а также – получения необходимой информации и выводов на основе данных онтологий. Кроме того, непрерывный рост количества и размеров онтологий требует современных и качественных методов визуализации элементов, способных справляться с возникающими алгоритмическими проблемами масштабируемости.

Из этого следует, что разработчики средств визуализации и редактирования онтологий должны принять меры по улучшению дизайна посредством анализа существующих и создания новых инструментов, обладающих более продвинутой функциональностью для конечных пользователей.

В данной работе для визуализации онтологии был выбран алгоритм послойной отрисовки графа, основанный на работе [9], посвященной техникам представления ориентированных графов. Реализация выполнена средствами библиотеки Javascript Dagre. Этот алгоритм прост в реализации, дает наглядное послойное отображение графа с возможностью интерактивной настройки расстояний между слоями или объектами слоя, а также достаточно эффективен для интерактивного использования. Последнее является ключевым фактором в его применении на ReactFlow.

По результатам анализа разработанных вручную визуальных представлений конечных автоматов, авторы в [9] сформулировали следующие принципы эстетического восприятия, которым должно удовлетворять визуальное представление графа.

1. Дуги графа следует по возможности располагать таким образом, чтобы они были ориентированы примерно в одном направлении, например сверху вниз. Такой подход придает графу иерархическую структуру и помогает пользователю в навигации по онтологии предметной области, в частности наследования атрибутов классов.

2. Следует избегать пересечения дуг и их изгиба под острыми углами. По мнению авторов [8], подобные визуальные аномалии искажают восприятие графа.

3. Дуги должны быть как можно более короткими, что делает более наглядными связи между вершинами.

4. По возможности отдавать предпочтение симметрии и балансу при расположении вершин. Данный принцип задает общий подход, которому необходимо следовать для создания корректных и наглядных визуальных представлений и обеспечивает минимизацию рабочего пространства, в котором расположен граф.

Для рендера графа (т. е. получения изображения по модели) был использован ReactFlow – инструмент на основе языка Javascript и фреймворка React с широкими возможностями настройки для создания редакторов на основе узлов и интерактивных диаграмм.

Для графового представления онтологии была предложена следующая цветовая схема:

- синий цвет узла – класс онтологии;
- красный цвет узла – экземпляр класса;
- оранжевый цвет узла – атрибут класса.

Согласно алгоритму визуализации, на первых слоях каждого подграфа располагаются вершины, соответствующие классам онтологии с последующим спуском к их экземплярам. Каждая дуга помечена названием соответствующего ей отношения в онтологии. Каждый подграф или группа вершин могут быть произвольно перемещены для организации рабочего пространства пользователя.

На рис. 5 показан пример визуализации графа, соответствующего фрагменту онтологии из области фольклора «Культурные универсалии народов Сибири» [10].

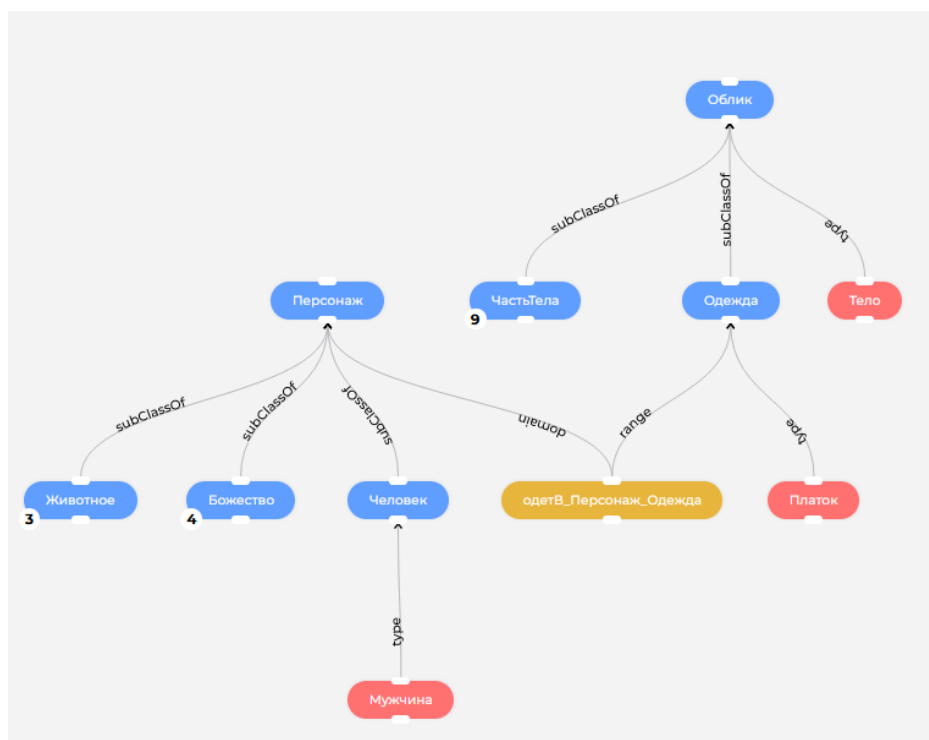


Рис. 5. Визуализация фрагмента онтологии из области фольклора «Культурные универсалии народов Сибири» на основе алгоритма послойной отрисовки графа

Fig. 5. Visualization of a fragment of the ontology from the field of folklore “Cultural universals of the peoples of Siberia” based on the algorithm of layered graph rendering

Объекты классов «Животное», «Божество» и «Часть тела» были скрыты разработанной функцией интерактивной фильтрации элементов графа. Число скрытых объектов определяется специальным маркером, расположенным в левой нижней части маркера узла (у «Животного» скрыто 3 объекта, у класса «Божество» – 4 объекта, у класса «Часть тела» – 9 объектов). Несмотря на то что класс «Персонаж» является корневым, его расположение на графе относительно других элементов было смещено на один слой ниже, чтобы избежать пересечения дуг.

3. Редактор онтологий

Для реализации модели представления и средств визуализации, приведенных в разделах 1 и 2, разрабатывается веб-редактор онтологий. Данный редактор предоставляет пользователю возможность работать непосредственно с графовыми представлениями онтологий и включает следующие базовые функции.

- А. Создание и редактирование элементов: корневых вершин онтологий, вершин-классов, вершин-экземпляров и т. д. Таким образом, пользователь редактора имеет возможность добавить новую или отредактировать существующую онтологию.

- В. Полнотекстовый поиск вершин. При положительном результате поиска граф в рабочем пространстве ориентируется таким образом, чтобы найденная вершина занимала центральную позицию.
- С. Фильтрация вершин путем выбора определенного набора меток.
- Д. Создание онтологических паттернов содержания в виде «фрагмента» онтологии, снабженного описанием.
- Е. Применение паттернов при разработке онтологий.

3.1. Редактор элементов онтологии

Каждый элемент графа является представлением определенного элемента онтологии. Для совершения действий над элементом в интерфейсе приложения-редактора предусмотрено персонализированное контекстное меню, пример которого показан на рис. 6.

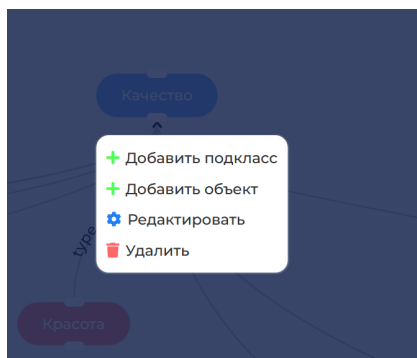


Рис. 6. Контекстное меню узла графа

Fig. 6. Context menu of the graph node

Содержимое данного меню зависит от типа узла графа. Для объектов онтологии существуют опции редактирования и удаления, в то время как для классов предусмотрен список отфильтрованных структурных паттернов, а также кнопки для создания их подклассов и объектов.

Таблица 2

Запросы для извлечения информации о структуре объектов

Table 2

Queries to Retrieve Information about the Structure of Objects

Запрос	Описание
match(n) - [:{subClass}*] -> (s) <- [:{domain}] - (d: {type}) where n.uri = '{uri}' return d	Сбор всех Datatype атрибутов (d) всех родителей класса n (s), где URI класса (n) равен переменной (uri)
match (d: {dataType})-[:{domain}]->(n) where n.uri = '{uri}' return d	Сбор Datatype атрибутов (d) конкретного класса (n), где URI класса (n) равен переменной (uri)
match(n)-[:{subClass}*]->(s)<-[:{domain}]- (d: {type})-[:{range}]->(r) where n.uri = '{uri}' return d,r	Сбор Object атрибутов (d) и их типов (r) всех родителей класса n (s), где URI класса (n) равен переменной (uri)
match(n)<-[:{domain}]- (d: {type})-[:{range}]->(r) where n.uri = '{uri}' return d,r	Сбор Object атрибутов (d) и их типов (r) конкретного класса (n), где URI класса (n) равен переменной (uri)

Разрабатываемый редактор предполагает его применимость для любой предметной области, поэтому поля редактора генерируются автоматически с помощью шаблонных запросов

на языке запросов Cypher графовой СУБД Neo4j. В табл. 2 представлены примеры шаблонных запросов для генерации полей.

На рис. 7 приведен пример формы для создания/редактирования объектов класса «Персонаж», сгенерированной на основе приведенных запросов. Фрагмент структуры класса «Персонаж» представлен на рис. 5.

Объект: Мужчина

Рис. 7. Форма редактирования узла

Fig. 7. Node editing form

В соответствии с онтологическим представлением, объекты класса «Персонаж» имеют атрибуты (DatatypeProperty) «Название» и «Описание», которые пользователь может заполнить в верхней части формы. Связи объекта с другими объектами (ObjectProperty) заполняются в нижней части формы, в данном примере связь «Персонаж_одеВ_Одежда».

3.2. Редактор паттернов содержания

Чтобы упростить разработку онтологий, была поставлена задача обеспечения поддержки паттернов онтологического проектирования (Ontology Design Patterns, ODP) [11]. На практике чаще всего используются онтологические паттерны содержания (Ontology Content Design Patterns) – небольшие целостные фрагменты онтологии, формализующие обобщенные ситуации предметной области (например, участие в событии, исполнение роли, наличие частей у объекта и др.). Паттерны содержания применяются в качестве строительных блоков при разработке и редактировании онтологий.

Для поддержки онтологических паттернов в редакторе онтологии был разработан инструмент для создания макетов паттернов на основе классов текущей предметной области. В основу редактора макетов входят 5 основных строительных блоков: «якорь», «создание», «существование», «выбор», «условие».

Узел «якорь» представляет собой начальную точку, от которой будет построен паттерн. Узел выражает тип элемента онтологии – класс или объект некоторого класса. Узел «создание» описывает элемент онтологии, который будет создан при применении паттерна. Узел «существование» является уже существующим элементов онтологии, который будет прикреплен к некому узлу паттерна, а узел «выбор» – существующий элемент онтологии, который будет выбран непосредственно при применении паттерна.

Для установления более детализированных условий на вышеописанные узлы используется узел «условие». Данный тип узлов не используется при непосредственном построении, а диктует тип используемых узлов путем установления ограничений на их связи.

Рассмотрим данный инструмент на элементарном примере: создание паттерна «Добавить сына». На рис. 8 изображен финальный макет паттерна «Добавить сына».



Рис. 8. Паттерн «Добавить сына»

Fig. 8. Pattern «Add son»

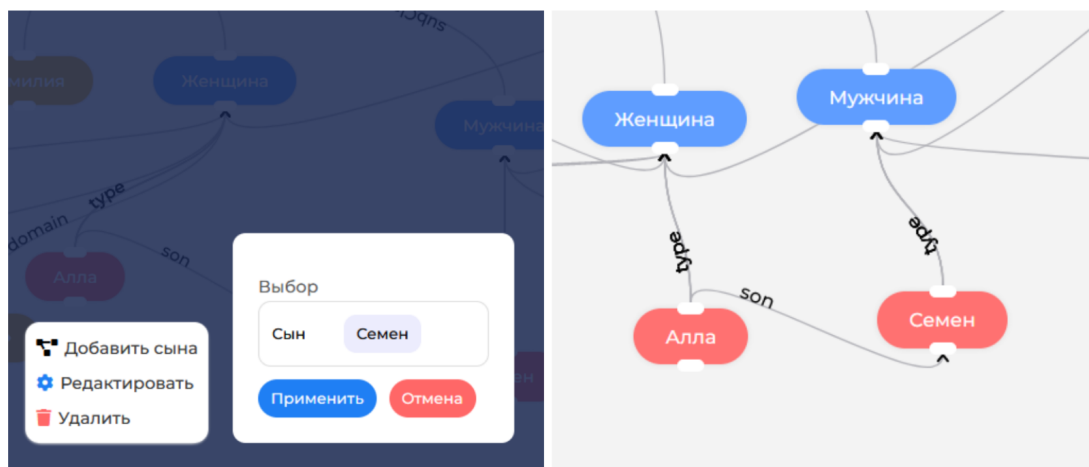


Рис. 9. Результат применения паттерна

Fig. 9. The result of applying the pattern

Применение паттернов осуществляется через контекстное меню определенного узла. Благодаря использованию узлов «условия», инструмент автоматически привязывает определенные паттерны к нужным узлам типа «якорь». На рис. 9 представлен результат применения паттерна «Добавить сына» к объекту «Алла» класса «Женщина».

3.3. Поиск и фильтрация

Приложение содержит функцию простого полнотекстового поиска по узлам. При обнаружении элемента у пользователя имеется возможность сфокусировать рабочее пространство графа на найденном элементе с его последующей подсветкой.

Фильтрация предполагает наличие фильтров по типам узлов графа. Такими фильтрами являются:

- фильтр по классам;
- фильтр по классам и их атрибутам;
- фильтр по объектам;
- фильтр по объектам и их атрибутам;
- фильтр по классам и объектам.

Заключение

В данной работе предлагается метод представления онтологий в графовых базах данных на примере графовой БД Neo4j. Данный метод основан на принципе наследования и расширения онтологий и предполагает использование специальных меток для обозначения ключевых элементов, таких как онтология, класс, экземпляр, атрибут, отношение, а также иерархических связей между онтологиями.

Создаваемый веб-редактор использует графовое представление онтологий, а для визуализации вершин и ребер графа — алгоритм послойной отрисовки помеченных графов, что позволяет повысить наглядность и визуальную привлекательность пользовательских онтологий. Особенностью предлагаемого решения является поддержка применения паттернов онтологического проектирования. Данный функционал позволит в дальнейшем создать полноценное хранилище и среду разработки паттернов онтологического проектирования, а также обеспечит их автоматическую интеграцию с разрабатываемыми пользовательскими онтологиями.

Список литературы

1. OWL 2 Web Ontology Language Document Overview (Second Edition) [Электронный ресурс]. URL: <https://www.w3.org/TR/owl2-overview> (дата обращения: 13.01.2023).
2. **Curé O, Blin G.** RDF database systems: triples storage and SPARQL query processing // Burlington: Morgan Kaufmann, 2014.
3. Virtuoso [Электронный ресурс]. URL: <https://virtuoso.openlinksw.com> (дата обращения: 13.01.2023)
4. **Ben Mahria B., Chaker I., Zahi A.** An empirical study on the evaluation of the RDF storage systems // Journal of Big Data. 2021. Pp. 1–20. DOI 10.1186/s40537-021-00486-y
5. **Stefano Ferilli.** Integration Strategy and Tool between Formal Ontology and Graph Database Technology // Electronics. 2021. V. 10(21). P. 2616. DOI 10.3390/electronics10212616
6. **Jesús Barrasa.** RDF Triple Stores vs. Labeled Property Graphs: What's the Difference? [Электронный ресурс] // URL: <https://neo4j.com/blog/rdf-triple-store-vs-labeled-property-graph-difference> (дата обращения: 13.01.2023).
7. **Gruber T. R.** Toward Principles for the Design of Ontologies Used for Knowledge Sharing // International Journal Human-Computer Studies. 1995. V. 43. Pp. 907–928.

8. **Bernaras, A.; Laresgoiti, I.; Corera, J.** Building and Reusing Ontologies for Electrical Network Applications // *Proceedings of the European Conference on Artificial Intelligence. ECAI 1996*. V. 96. Pp. 298–302.
9. **Gansner, Emden & Koutsofios, Eleftherios & North, Stephen & Vo, Khoi.** A Technique for Drawing Directed Graphs // *IEEE Transactions on Software Engineering*. 1993. V. 19. Pp. 214–230.
10. **Гриневи́ч А., Серый А.** Анализ культурных универсалий фольклора народов Сибири и Дальнего Востока // *Материалы XXIII Международной конференции по аналитике и управлению данными в информационно емких областях (DAMDID/RCDL 2021)*. Москва, Россия, 26–29 октября 2021 г. Материалы семинара CEUR, 2021. Vol. 3036. С. 387–401.
11. **Gangemi A., Presutti V.** Ontology Design Patterns // *Handbook on Ontologies*, Eds., Staab, S. and R. Studer. Berlin: Springer Verlag, 2009. Pp. 221–243.

References

1. OWL 2 Web Ontology Language Document Overview (Second Edition) [Online]. URL: <https://www.w3.org/TR/owl2-overview/> (accessed on: 13.01.2023).
2. **Curé O, Blin G.** RDF database systems: triples storage and SPARQL query processing. Burlington: Morgan Kaufmann, 2014.
3. Virtuoso [Online]. URL: <https://virtuoso.openlinksw.com> (accessed on: 13.01.2023).
4. **Ben Mahria B., Chaker I., Zahi A.** An empirical study on the evaluation of the RDF storage systems. *Journal of Big Data*, 2021, pp. 1–20. DOI 10.1186/s40537-021-00486-y
5. **Stefano Ferilli.** Integration Strategy and Tool between Formal Ontology and Graph Database Technology. *Electronics*, 2021, vol. 10(21), p. 2616. DOI 10.3390/electronics10212616
6. **Jesús Barrasa.** RDF Triple Stores vs. Labeled Property Graphs: What's the Difference? [Online]. URL: <https://neo4j.com/blog/rdf-triple-store-vs-labeled-property-graph-difference> (accessed on: 13.01.2023).
7. **Gruber T. R.** Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal Human-Computer Studies*, 1995, no. 43, pp. 907–928.
8. **Bernaras A., Laresgoiti I., Corera J.** Building and Reusing Ontologies for Electrical Network Applications. *Proceedings of the European Conference on Artificial Intelligence. ECAI 96*, 1996. Pp. 298–302.
9. **Gansner E., Koutsofios E., North S., Vo K.** A Technique for Drawing Directed Graphs. *Software Engineering, IEEE Transactions on*. 19. 1993. pp. 214–230.
10. **Grinevich A., Sery A.** Analyzing the Cultural Universals of the Folklore of Peoples of Siberia and the Far East. *Proceedings of the XXIII International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2021)*. Moscow, Russia, October 26–29, 2021. CEUR Workshop Proceedings, 2021, Vol-3036. Pp. 387–401.
11. **Gangemi A., Presutti V.** Ontology Design Patterns. *Handbook on Ontologies*, Eds., Staab, S. and R. Studer. Berlin: Springer Verlag, 2009. Pp. 221–243.

Информация об авторах

Владислав Александрович Лисин, студент аспирантуры

Елена Анатольевна Сидорова, старший научный сотрудник, кандидат физико-математических наук

Алексей Сергеевич Серый, младший научный сотрудник

Information about the Authors

Vladislav A. Lisin, PhD student, Novosibirsk State University (Novosibirsk, Russia)

Elena A. Sidorova, senior researcher, Ph.D., Ershov Institute of Informatic Systems SB RAS, AI laboratory (Novosibirsk, Russia)

Alexey S. Sery, junior researcher, Ershov Institute of Informatic Systems SB RAS, AI laboratory (Novosibirsk, Russia)

*Статья поступила в редакцию 26.12.2022;
одобрена после рецензирования 12.01.2023; принята к публикации 12.01.2023*

*The article was submitted 26.12.2022;
approved after reviewing 12.01.2023; accepted for publication 12.01.2023*

Научная статья

УДК 519.81

DOI 10.25205/1818-7900-2022-20-4-39-60

Исследование взаимосвязи между конфликтными свойствами и показателями эффективности организационно-технических систем на примере сети связи специального назначения

Алексей Вячеславович Попов¹

Олег Викторович Пьянков²

Воронежский институт МВД России

Воронеж, Россия

¹Alex_std_ex@mail.ru

²ovpyankov@mail.ru

Аннотация

Используется системный подход к исследованию организационно-технических систем, включающих организационные и технические элементы и взаимодействия между ними. Проводится анализ литературы по исследуемой проблематике, делается выбор в пользу теории конфликтов для исследования организационно-технических систем и определяется необходимость решения научной задачи по уточнению применимости существующих оценок конфликтности системы. Рассматривается гипотеза о наличии связи между оценками конфликтных отношений элементов и показателями эффективности функционирования организационно-технических систем. На примере локальной сети связи специального назначения разрабатывается математическая модель конфликтных взаимодействий, вводятся функции полезности ее элементов, рассматриваются существующие и предлагаются новые методы оценки конфликтных отношений элементов и системы в целом.

На макете рассматриваемой локальной сети проводится натурный эксперимент по определению значений показателей эффективности работы сети (отношение скорости передачи данных к пропускной способности канала связи, загрузка центрального процессора, отношение времени передачи фиксированного объема данных к минимально достижимому при использовании всей пропускной способности канала связи) в различных условиях ее функционирования, в том числе при воздействии негативных факторов. Для каждого исследуемого состояния локальной сети осуществляется модификация математической модели с вычислением значений оценок конфликтных отношений элементов и системы в целом.

Осуществляется выбор оценок конфликтных отношений, потенциально имеющих связь с показателями эффективности. Между выбранным набором оценок конфликтных отношений и показателями эффективности методами корреляционного анализа определяется наличие взаимосвязи. В результате анализа выделены оценки, которые могут применяться для прогнозирования степени изменения конфликтности системы и ее элементов при модификации или изменении условий функционирования.

Ключевые слова

организационно-техническая система, теория конфликтов, оценки конфликтности, показатели эффективности, натурный эксперимент

Для цитирования

Попов А. В., Пьянков О. В. Исследование взаимосвязи между конфликтными свойствами и показателями эффективности организационно-технических систем на примере сети связи специального назначения // Вестник НГУ. Серия: Информационные технологии. 2022. Т. 20, № 4. С. 39–60. DOI 10.25205/1818-7900-2022-20-4-39-60

Study of the Relationship between Conflict Properties and Performance Indicators of Organizational and Technical Systems on the Example of a Special Purpose Communication Network

Alexey V. Popov¹, Oleg V. Pyankov²

Voronezh Institute of the Ministry of Internal Affairs of Russia
Voronezh, Russia

¹Alex_std_ex@mail.ru

²ovpyankov@mail.ru

Abstract

The system approach to the study of organizational and technical systems, including organizational and technical elements and the interaction between them is used. The analysis of the literature on the investigated problem is carried out, the choice in favor of the conflict theory for the study of organizational-technical systems is made and the necessity of solving the scientific problem of clarifying the applicability of existing assessments of the conflicts of the system is determined. The hypothesis of the existence of a relationship between estimates of the conflict relations of elements and performance indicators of organizational-technical systems is considered.

By the example of a special purpose local communication network a mathematical model of conflict interactions is developed, utility functions of its elements are introduced, existing and new methods of evaluation of conflict relations of elements and the system as a whole are considered.

On the model of the considered local network we conduct a full-scale experiment to define the network performance (the ratio of the data transmission rate to the channel capacity, the central processor load, the ratio of the transmission time of a fixed amount of data to the minimum achievable with the use of the entire channel capacity) in various conditions of its operation, including the impact of negative factors. For each investigated state of the local network, we modify the mathematical model to calculate the values of the conflict relations of the elements and the system as a whole.

The choice of estimates of the conflict relations, potentially having connection with efficiency indicators is carried out. The correlation between the selected set of assessments of conflict relations and performance indicators is determined by methods of correlation analysis. As a result of the analysis, we singled out the estimates, which can be used to predict the degree of change in the conflicts of the system and its elements when modifying or changing the conditions of functioning.

Keywords

organizational and technical system, conflict theory, conflict assessment, performance indicators, field experiment

For citation

Popov A. V., Pyankov A. V. Study of the Relationship between Conflict Properties and Performance Indicators of Organizational and Technical Systems on the Example of a Special Purpose Communication Network. *Vestnik NSU. Series: Information Technologies*, 2022, vol. 20, no. 4, pp. 39–60. (in Russ.) DOI 10.25205/1818-7900-2022-20-4-39-60

Введение

Органы внутренних дел МВД России в пределах своих компетенций реализуют широкий спектр полномочий, связанных с охраной общественного порядка и обеспечением общественной безопасности. Для централизованного управления силами и средствами органов внутренних дел, осуществления коммуникаций между ними, а также взаимодействия с другими органами государственной власти, местного самоуправления, организациями и гражданами в системе Министерства внутренних дел Российской Федерации используются сети связи специального назначения (СС СН), администрирование которых осуществляют сотрудники центров/отделов информационных технологий, связи и защиты информации (Ц(О)ИТСиЗИ) территориальных органов МВД России.

Использование СС СН невозможно без осуществления исследования и анализа их текущего состояния, оптимизации и прогнозирования эффективности их функционирования, исходя из сложившейся оперативной обстановки. Для исследования свойств таких сетей связи целесообразно применение системного подхода, рассматривающего СС СН как организаци-

онно-техническую систему (ОТС), включающую организационные и технические элементы и воздействия их друг на друга [1]. Особенностью ОТС является значимость роли человека в функционировании системы и реализации процессов. К организационным элементам относятся пользователи (в том числе технический персонал), осуществляющие эксплуатацию и обслуживание СС СН, к техническим – программные и аппаратные средства СС СН, включая технические средства связи, системы отображения информации индикации, органы управления и программные интерфейсы [2].

В ряде исследований [3–11] рассмотрены различные подходы к моделированию ОТС, а также исследованию их свойств и описанию реализуемых в них процессов. Так, в работе [3] предложена методология, реализующая математическое и программное обеспечение, направленное на повышение уровня профессиональной подготовки специалистов по эксплуатации технических систем. В исследовании [4] основной задачей является анализ информационных процессов, реализуемых в ОТС, и автоматизация управления, направленного на повышение защищенности таких систем. В [5] предложен комплексный показатель, характеризующий надежность функционирования ОТС, основанный на вероятностях нахождения техники в состоянии готовности и безошибочной работы пользователей, а также рассмотрен перечень состояний, в которых может находиться такая система и причины перехода из одного состояния в другое. В [6] использован подход, основанный на разработке структурно-параметрических моделей социально-технических систем в виде взвешенных гиперграфов, а эффективность функционирования систем оценивается вкладом каждого элемента, направленным на достижение общей цели системы.

В исследованиях [8; 11] авторы используют методы системного анализа для синтеза, оценки эффективности функционирования, регулирования и контроля ОТС управления и выработки соответствующих управленческих решений, направленных на оптимизацию ОТС.

Работы [9; 10] посвящены разработке моделей и исследованию свойств адаптивных тренировочных комплексов, рассматриваемых как эргатические системы, что для предметной области является эквивалентом ОТС. В [10] такая эргатическая система включает в себя аппаратное и программное обеспечение, а также математические модели технической системы, деятельности персонала и воздействий на комплекс. Каждый из структурных элементов дополнительно состоит из подсистем, позволяющих описать необходимые свойства исследуемого комплекса.

Основываясь на структурных особенностях ОТС [1; 2; 12; 13] в процессе функционирования, согласно назначению, она может быть представлена как [14]:

$$\mathfrak{A} = \langle S, R, Q \rangle,$$

где $S = \{s_i\}$ – множество элементов системы, $R = \{p_{i,j}^k, z_{i,j}^k\}$ – множество воздействий элементов системы друг на друга (p – класс процессов, реализуемых организационными элементами, z – класс процессов, реализуемых техническими элементами, k – индекс процесса), $Q = \{q_{s_i}\}$ – множество функций полезности элементов, которые характеризуют степень достижения цели каждого i -го элемента.

Индекс процесса k позволяет идентифицировать процесс конкретной предметной области, реализуемый в исследуемой системе. Так, в работе [15] были рассмотрены и проанализированы процессы p -класса, которые реализуются между организационным и техническим элементами: p^1 – эксплуатация СС СН; p^2 – модернизация; p^3 – управление; p^4 – построение (конфигурация) СС СН или ее сегмента; p^5 – мониторинг качества оказываемых услуг связи; p^6 – техническое обслуживание; p^7 – ремонт; p^8 – хранение; p^9 – транспортирование. К процессам z -класса относятся: z^1 – передача информации; z^2 – формирование; z^3 – прием; z^4 – обработка; z^5 – хранение информации.

Реализуемые в системе процессы позволяют выделить отношения между элементами системы, участвующими в этих процессах. Тип отношений определяется исходя из изменений функций полезности элементов в соответствии с положениями теории конфликтов.

Основная идея этой теории заключается в разработке системной модели конфликта, которая связывает объекты и факторы, участвующие в конфликте [16]. Для теории конфликтов свойственна интеграция качественных и количественных законов, и, наряду с этим, одной из ее особенностей является применение технических измерений и наличие констант. Этот факт обуславливает важность построения функционального пространства со своей метрикой при разработке системной модели конфликта. Метрика должна охватывать как технические, так и эргономические компоненты исследуемых систем. Временной масштаб системных моделей может варьироваться в зависимости от условий решаемой задачи, однако в любой модели можно проследить разнообразные воздействия элементов друг на друга, их поведение. Исследование ОТС с точки зрения теории конфликтов позволяет решать аналитические задачи и строить различного вида прогнозы их развития.

Ряд зарубежных источников [17–25], содержащих исследования по проблематике теории конфликтов, используют аппарат математического моделирования для анализа объектов и систем в различных предметных областях. В работах [17; 22; 23] исследуются причины, динамика конфликтов, возникающих при реализации социальных процессов в сфере жизнедеятельности общества. Определяется множество участников конфликта и моделируются пути его разрешения с применением дифференциальных уравнений. Также категориально-понятийный аппарат теории конфликтов может быть применен при возникновении противоречий в политике предоставления доступа к данным в географических информационных системах [18], для прогнозирования и анализа траекторий движения воздушных судов и недопущения столкновений [19], для разработки математических моделей систем, оценивающих эпидемиологическую обстановку и динамику распространения инфекционных заболеваний [20], а также в моделях, разрабатываемых для оценки бизнес-процессов [21].

Основываясь на системном подходе при моделировании ОТС целесообразно использовать определения теории конфликтов, что позволит охарактеризовать отношения между элементами системы. Как правило, принято выделять три основных типа отношений: сотрудничество, противоречие и безразличие [26; 27]. Определение типа отношений между элементами связано с выполнением следующих условий:

1. Если существует воздействие элемента i на элемент j , такое что при возрастании (убывании) функции полезности i -го элемента возрастает (убывает) функция полезности j -го элемента, то отношение i -го элемента к j -му называется сотрудничеством.

Если $\exists p_{i,j}^k \cup z_{i,j}^k : q_{s_i} \uparrow (\downarrow) \Rightarrow q_{s_j} \uparrow (\downarrow)$, то $s_i > |_c s_j$.

Аналогично:

2. Если $\exists p_{i,j}^k \cup z_{i,j}^k : q_{s_i} \uparrow (\downarrow) \Rightarrow q_{s_j} \downarrow (\uparrow)$, то $s_i > |_n s_j$ (противоречие).

3. Если $\exists p_{i,j}^k \cup z_{i,j}^k : q_{s_i} \uparrow (\downarrow) \Rightarrow q_{s_i} = const$, то $s_i > |_o s_j$ (безразличие).

Если $\nexists (p_{i,j}^k \cup z_{i,j}^k) \wedge \nexists (p_{j,i}^k \cup z_{j,i}^k)$, то элементы i и j независимы друг от друга.

Совокупность отношений сотрудничества и противоречий между элементами ОТС позволяет оценить ее конфликтность и сбалансированность, оказывающие влияние на реализуемые процессы. В связи с этим целью исследования является проверка гипотезы о наличии взаимосвязи между оценками конфликтности и показателями эффективности ОТС.

1. Моделирование организационно-технической системы

Разработаем модель ОТС на примере локальной СС СН, состоящей из трех последовательно соединенных друг с другом коммутаторов и двух пользователей, один из которых подключен к порту первого коммутатора, а второй – к порту третьего (рис. 1).

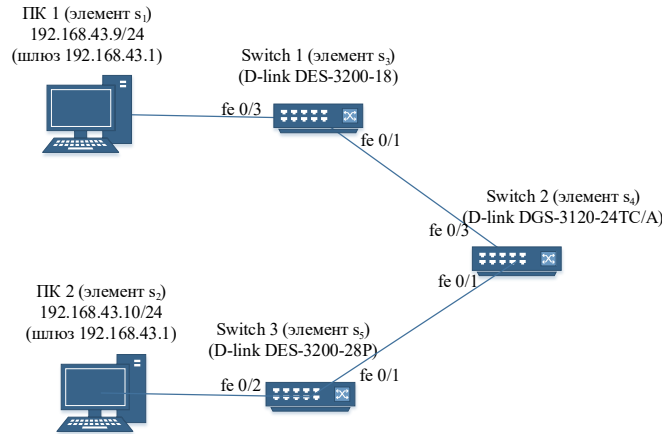


Рис. 1. Схема локальной CC CH
Fig. 1. Diagram of the local special-purpose communication network

К организационным элементам ОТС отнесем двух пользователей, осуществляющих одностороннюю и двухстороннюю передачу данных друг другу. С учетом требуемого уровня декомпозиции [13], отражающего перечень конфликтных отношений между элементами, персональные компьютеры пользователей не будут выделяться как отдельные элементы, а будут находиться в составе организационных элементов, образуя тем самым эргатические (социально-технические) элементы s_1, s_2 . Коммутаторы отнесем к техническим элементам s_3, s_4, s_5 . Таким образом, множество элементов исследуемой ОТС в ее исходном состоянии зададим следующим образом: $S_0 = \{s_1, s_2, s_3, s_4, s_5\}$.

Чтобы задать множество R , определим, какие процессы будут реализовываться в системе. Рассмотрим ситуацию, когда пользователю s_1 необходимо передать некоторый объем данных пользователю s_2 , а пользователю s_2 передать подтверждение о получении сообщения. В таком случае эргатический элемент s_1 будет оказывать прямое воздействие на первый коммутатор

(элемент s_3), т. е. $s_1 \xrightarrow{p_{1,3}^1} s_3$ (см. рис. 1), где $p_{1,3}^1$ означает эксплуатацию устройства. Аналогичным образом будут взаимодействовать элементы s_2 и s_5 при обратной передаче данных ($s_2 \xrightarrow{p_{2,5}^1} s_5$). Между элементами s_3, s_4, s_5 будут реализовываться процессы класса z : $z_{3,4}^1, z_{4,3}^1, z_{4,5}^1, z_{5,4}^1$, связанные с передачей информации между узлами связи, т. е. $s_i \xrightarrow{z_{i,j}^1} s_j$. Помимо этого, реализуются также два процесса передачи данных эргатическим элементам от элементов s_3 и s_5 : $s_5 \xrightarrow{z_{5,2}^1} s_2$ и $s_3 \xrightarrow{z_{3,1}^1} s_1$. Таким образом, множество воздействий для исходного состояния исследуемой системы $R_0 = \{p_{1,3}^1, p_{2,5}^1, z_{3,1}^1, z_{5,2}^1, z_{3,4}^1, z_{4,3}^1, z_{4,5}^1, z_{5,4}^1\}$.

Далее перейдем к определению множества функций полезности элементов Q_0 . Отметим, что функции полезности $q_{s1}(X_1, c_1)$ и $q_{s2}(X_2, c_2)$ являются аналитически эквивалентными в силу идентичности природы элементов s_1 и s_2 . Поскольку каждый пользователь заинтересован в скорейшей отправке/приеме сообщения, то его функция полезности будет определяться выражением:

$$q_{s1}(X_1, c_1) = X_1 \cdot c_1 = q_{s2}(X_2, c_2) = X_2 \cdot c_2, \quad (1)$$

где c_1 и c_2 – состояния элементов, принимающие значения 0 (если элемент не участвует в функционировании системы) и 1 (элемент участвует в функционировании системы); X_i – отношение суммы скоростей передачи и приема информации i -м пользователем к пропускной способности канала связи, определяемое по формуле:

$$X_i = \frac{v_{ni} + v_{npi}}{2 \cdot C}, \quad i = 1, 2, \quad (2)$$

где v_{ni} – скорость передачи информации i -м пользователем, v_{npi} – скорость приема информации i -м пользователем, C – пропускная способность канала связи. Значения функций полезности q_{s1} , q_{s2} могут изменяться в диапазоне от 0 до 1. Для исследуемой предметной области справедливо предположение, что чем больше значение функций полезности пользователей, тем эффективнее функционирует ОТС, целью которой является удовлетворение потребностей пользователей.

Функции полезности сетевых устройств сегмента СС СН q_{s3} , q_{s4} , q_{s5} зададим следующим образом:

$$q_{si}(X_i, c_i) = X_i \cdot c_i = \left(\frac{1}{N_i} \sum_{k=1}^{N_i} \frac{v_{nk}}{C_k} \right) \cdot c_i, \quad i = 3, 4, 5, \quad (3)$$

где X_i – среднее значение отношений скоростей передачи информации пользователями к пропускной способности задействованных каналов связи, N – количество пользователей, подключенных к портам коммутаторов и осуществляющих одновременную передачу данных.

Таким образом, $Q_0 = \{q_{s1}, q_{s2}, q_{s3}, q_{s4}, q_{s5}\}$.

В работе [13] дано описание причинно-следственных связей возникновения определенного типа конфликтного отношения между элементами. Эта связь обусловлена характером процессов, реализующих функции СС СН. Поскольку в моделируемой системе рассматриваются процессы эксплуатации и передачи данных, то при отправке сообщений каждый из элементов множества S_0 будет повышать функцию полезности смежного с ним элемента, таким образом, $s_i > |_c s_j$ для всех i и j .

Широкое применение при моделировании ОТС нашли ориентированные графы, в которых вершинам v_i соответствуют элементы системы s_i , а дугам $e(z, k)_{ij}$ и $e(p, k)_{ij}$ – реализуемые между элементами процессы класса p или z [13; 28], причем если при реализации процесса $p_{i,j}^k$ возникает отношение сотрудничества между элементами i и j , то дуга $e(p, k)_{ij}$ будет сплошной, а если отношение противоречия – пунктирной. Таким образом, модель рассматриваемой системы может быть представлена в виде графа $G_0 = G(V, E)$, где V – множество вершин, E – множество дуг (см. рис. 2).

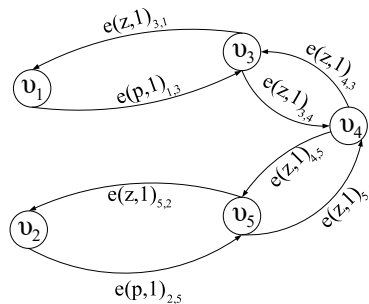


Рис. 2. Граф G_0

Fig. 2. Graph G_0

Пусть в начальный момент времени t_0 $q_{si}(X_i, c_i) = 0$ для всех i . Когда s_i начинает передавать сообщение, то будем считать, что данные передаются с максимально возможной скоростью, и, соответственно, $q_{s1} \rightarrow 0,5$, так как $v_{n1} \approx C$, $v_{npi} = 0$. Поскольку все отношения между элементами в системе являются отношениями сотрудничества, то функции полезности остальных

элементов начинают аналогичным образом пропорционально возрастать, тогда с учетом (3) для $i = 3, 4, 5$ функции полезности $q_{s1} \rightarrow 0,5$, а $q_{s2} \rightarrow 0,5$.

1.1. Модель системы с нарушителем

Добавим в систему нарушителя s_6 , осуществившего несанкционированный доступ к порту коммутатора (s_3) и загружающего канал связи паразитным трафиком.

Элемент s_6 является организационным элементом, реализующим процесс $p_{6,3}^1$ (атаку паразитным трафиком, в силу идентичности природы процессов эксплуатации и атаки, отнесем к p^1). Его собственная функция полезности определяется как

$$q_{s6}(X_6, c_6) = X_6 \cdot c_6 = \frac{v_{n6}}{C}. \quad (4)$$

В исследуемой сети такая DoS-атака⁵ вызывает снижение скорости передачи данных пользователем s_1 , что приводит к соответствующему убыванию функций полезности q_{si} , для всех $i < 6$ и возрастанию q_{s6} . Такое изменение функций полезности соответствует конфликтному отношению $s_6 > |s_3$ процесса $s_6 \xrightarrow{p_{6,3}^1} s_3$. Таким образом, ОТС для рассматриваемого случая задается следующим образом:

$$\mathcal{S}_1 = S_1, R_1, Q_1 >$$

где $S_1 = S_0 \cup \{s_6\}$, $R_1 = R_0 \cup \{p_{6,3}^1\}$, $Q_1 = Q_0 \cup \{q_{s6}\}$.

Ее модель в виде графа G_1 представлена на рис. 3.

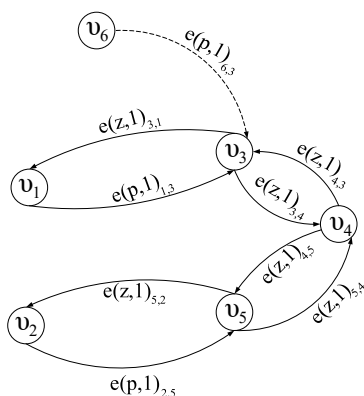


Рис. 3. Граф G_1
Fig. 3. Graph G_1

1.2. Разработка модели системы с противодействием нарушителю

При возникновении различных атак на сети связи возникает необходимость реализации комплекса мероприятий, направленных на противодействие угрозам, нарушающим бесперебойную работу средств связи, а также снижающим эффективность функционирования СС СН. Такие мероприятия проводятся сотрудниками Ц(О)ИТСиЗИ МВД России, в частности инженерами. При возникновении угрозы, рассмотренной в разделе 1.1, в целях ее нейтрализации либо минимизации негативных последствий для деятельности органов внутренних дел в структуру исследуемой ОТС $\mathcal{S}_1$ может быть добавлен элемент s_7 (инженер/системный администратор),

⁵ Англ. Denial of Service – «отказ в обслуживании».

идентифицирующий возникающую угрозу и принимающий меры по противодействию ей. В случае если DoS-атака осуществлялась на коммутатор s_3 , инженер реализует процесс управления $p_{7,3}^3$ по отношению к коммутатору, подвергающемуся негативному воздействию. Таким образом, в системе будет реализован новый процесс $s_7 \xrightarrow{p_{7,3}^3} s_3$. Функцию полезности q_{s_7} зададим выражением:

$$q_{s_7}(X_7, c_7) = X_7 \cdot c_7 = \left(\frac{1}{N} \sum_{k=1}^N q_{s_k} \right) \cdot c_7, \quad (5)$$

где q_{s_k} – функция полезности пользователя, N – количество пользователей.

Управляющее воздействие на коммутатор со стороны инженера будет способствовать возрастанию функции полезности q_{s_3} , что соответствует отношению сотрудничества $s_7 > |_c s_3$.

Структура и состав такой системы аналогично, как и в разделе 1.1, задается как:

$$\mathfrak{A}_2 = \langle S_2, R_2, Q_2 \rangle,$$

где $S_2 = S_1 \cup \{s_7\}$, $R_2 = R_1 \cup \{p_{7,3}^3\}$, $Q_2 = Q_1 \cup \{q_{s_7}\}$.

Модель системы представлена в виде графа G_2 на рис. 4.

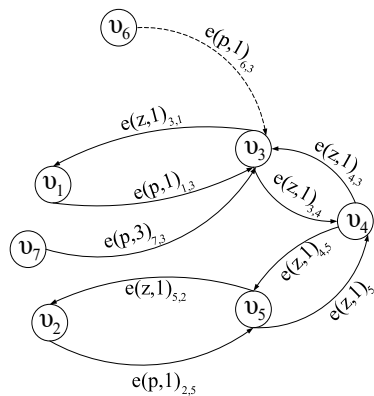


Рис. 4. Граф G_2
Fig. 4. Graph G_2

В рамках исследования были разработаны 63-графовые модели СС СН, обладающие различным составом и структурой. Рассматривалась возможность подключения к СС СН до 9 нарушителей, реализующих масштабную DDoS-атаку сразу на все коммутаторы. Например, самая крупная система задавалась множествами следующей мощности:

$$|S_n|_{\max} = |Q_n|_{\max} = 15; |R_n|_{\max} = 20.$$

1.3. Разработка оценок конфликтности организационно-технических систем

Для определения конфликтности систем используются оценки конфликтности, рассчитываемые по знаковым графам, предложенные в работах [13; 29; 30].

Основными из них являются:

1. M – баланс по Харари, определяемый как

$$\begin{aligned} M^- &= C^- / C = (C - C^+) / C; \\ M^+ &= C^+ / C = (C - C^-) / C, \end{aligned} \quad (6)$$

где C – общее число циклов в системе, C^+ – число сбалансированных циклов (в которых все существующие отношения между элементами синергетические, либо количество антагонистических отношений чётно); C^- – число несбалансированных циклов (циклы, в которых число антагонистических отношений нечётно).

2. M_c – степень сбалансированности графа по вершинам, определяемая формулой:

$$M_c = |v^c| / |V|, \quad (7)$$

где $|V|$ – мощность множества вершин графа, $|v^c|$ – мощность множества вершин, сбалансированных по циклам.

3. M_s – баланс по системе:

$$M_s = |v^s| / |V|, \quad (8)$$

где $|v^s|$ – мощность множества вершин, сбалансированных по системе.

4. M_n – степень несбалансированности:

$$M_n = |v^n| / |V|, \quad (9)$$

где $|v^n|$ – мощность множества несбалансированных вершин.

5. M_m – средняя сбалансированность графа:

$$M_m = \frac{1}{|V|} \sum_{i=1}^{|V|} M_i; \quad (10)$$

$$M_i = C^+(v_i) / C(v_i), \quad (11)$$

где $C(v_i) = C^+(v_i) + C^-(v_i)$ – общее число циклов, которым принадлежит вершина v_i ; $C^+(v_i)$ и $C^-(v_i)$ – число сбалансированных и несбалансированных циклов, которым принадлежит вершина v_i .

6. $\vec{P}_i$ – воздействие i -го фактора на систему:

$$\vec{P}_i = \frac{1}{|V|} \sum_{j=1}^{|V|} d_{i,j}, \quad (12)$$

где $d_{i,j}$ – показатель воздействия i -го фактора на j -й фактор. Этот показатель определяется из матрицы транзитивного замыкания $W' = \|w'_{i,j}\|_{2V \times 2V}$, где $w'_{i,j}$ – веса взаимных воздействий факторов друг на друга, определяемые экспертами. Элементы $d_{i,j}$ получаются из матрицы W' по правилу: $d_{i,j} = w'_{2i,2j}$, если $w'_{2i,2j} > w'_{(2i-1),(2j-1)}$; в противном случае $d_{i,j} = 0$.

7. $\vec{P}_j$ – воздействие системы на j -й фактор:

$$\vec{P}_j = \frac{1}{|V|} \sum_{i=1}^{|V|} d_{i,j}. \quad (13)$$

8. X_{pz} – баланс по взаимным воздействиям элементов системы:

$$X_{pz} = \sum_{i=1}^{|R|} r_{i,j}, \quad (14)$$

где $r_{i,j}$ – вес отношения между i -м и j -м элементами в пределах одной компоненты связности графовой модели. Модуль $r_{i,j}$ определяется как

$$|r_{i,j}| = \frac{1}{|R|},$$

$$\begin{aligned} r_{i,j} &> 0, \text{ если } s_i > |S_j|; \\ r_{i,j} &< 0, \text{ если } s_i < |S_j|. \end{aligned} \quad (15)$$

Разработаем дополнительные оценки конфликтности, позволяющие исследовать ОТС с различными топологическими свойствами, не привязанными к наличию обратных связей (циклов), а также получать интегральные (усредненные) значения по всем элементам системы или только для отдельного множества элементов (например, только для элементов, входящих в базовую модель G_0).

9. $I\bar{P}_\phi$ – интегральная оценка воздействий элементов, входящих в состав базовой модели G_0 :

$$I\bar{P}_\phi = \sum_{i=1}^{|V_0|} \bar{P}_i, \quad (16)$$

где V_0 – множество вершин в модели G_0 .

10. $I\bar{P}_s$ – интегральная оценка влияний элементов на систему:

$$I\bar{P}_s = \sum_{i=1}^{|V|} \bar{P}_i. \quad (17)$$

11. $C\bar{P}_s$ – концентрация влияний элементов на систему:

$$C\bar{P}_s = \frac{1}{|V|} I\bar{P}_s. \quad (18)$$

12. $I\bar{P}_\phi$ – интегральная оценка влияния системы на элементы, входящие в состав базовой модели G_0 :

$$I\bar{P}_\phi = \sum_{j=1}^{|V_0|} \bar{P}_j, \quad (19)$$

13. $I\bar{P}_s$ – интегральная оценка влияния системы на элементы:

$$I\bar{P}_s = \sum_{j=1}^{|V|} \bar{P}_j. \quad (20)$$

14. $C\bar{P}_s$ – концентрация влияний системы на элементы:

$$C\bar{P}_s = \frac{1}{|V|} I\bar{P}_s. \quad (21)$$

15. Y_{in} – баланс по сторонним входным воздействиям. Данная оценка рассчитывается для G_0 по следующей формуле:

$$Y_{in} = \sum_{i=1}^{|V_0|} \left[\sum_{k=1}^{E1} (\varpi_k(t) \cdot \alpha_k \cdot y_k^-(v_i)) + \sum_{k=1}^{E2} (\varpi_k(t) \cdot \beta_k \cdot y_k^+(v_i)) \right], \quad (22)$$

где $y_k^-(v_i)$, $y_k^+(v_i)$ – входные воздействия на вершину $v_i \in V_0$, принимающие значение -1 (в случае противоречия) или 1 (в случае сотрудничества); α_k , β_k – коэффициенты множественности воздействия принимают значение 1 , если входные воздействия скомпенсированы, в противном случае $\alpha_k, \beta_k > 1$; $E1$, $E2$ – количество сторонних воздействий противоречия и сотрудничества; $\varpi_k(t)$ – коэффициент силы воздействия k -го элемента на базовую модель.

2. Проведение натурального эксперимента

Показатели эффективности ОТС, в частности, СС СН задаются множеством $\Phi = \{\phi_i\}$ и могут включать различные качественные и количественные характеристики, учитываемые при организации информационного обмена и любых коммуникаций между пользователями сетей связи.

В литературных источниках в качестве показателей эффективности функционирования сетей связи используются различные подходы к моделированию и оценке эффективности функционирования сетей связи. Например, в работе [31] эффективность функционирования сетей передачи данных оценивается показателем вероятности выхода из строя или отключения узлов связи, входящих в структуру сети. С применением аппарата комбинаторики из значения данного показателя рассчитываются уровень стабильности сети, величина детерминированного ущерба, а также производится оценка суммарного риска. Исследование [32] сопряжено с разработкой математических моделей и методов повышения эффективности функционирования сетей сотовой связи, основным показателем для которых является среднее время отклика системы. А в [33] оценивается плотность и скорость вызовов в беспроводных сетях связи для последующего ввода эволюционных дифференциальных уравнений, позволяющих охарактеризовать дальнейшую динамику развития таких систем и, соответственно, разработать их стохастические или детерминированные модели.

Исходя из заданных в разделе 1 функций полезности элементов исследуемой ОТС, в качестве основных показателей эффективности функционирования СС СН были выбраны:

- отношение скорости передачи данных к пропускной способности канала связи (ϕ_1);
- загрузка центрального процессора компьютера источника (ϕ_2);
- отношение времени передачи фиксированного объема данных к минимально достижимому при использовании всей пропускной способности канала связи (ϕ_3).

Таким образом, множество показателей эффективности имеет следующий вид:

$$\Phi = \{\phi_1, \phi_2, \phi_3\}.$$

Для подтверждения выдвинутой гипотезы о зависимости между конфликтными свойствами СС СН и эффективностью ее функционирования был проведен натурный эксперимент, основанный на локальной сети, конфигурирование которой осуществлялось согласно схеме (см. рис. 1). Экспериментальная установка для проведения натурного эксперимента представлена на рис. 5.



Рис. 5. Макет локальной СС СН
Fig. 5. Layout of a special-purpose local communications network

Для натурного эксперимента использовалось программное обеспечение PRTG Network Monitor [34], позволяющее оценить выбранные показатели эффективности.

В ходе натурного эксперимента была произведена тестовая передача данных от пользователя s_1 пользователю s_2 . Объем данных в ходе каждой передачи был фиксирован и составлял 2,2 Гб. Значение функций полезности рассчитывалось на основе данных о скорости исходящего/входящего трафика на коммутаторах или пользовательских ПК.

Изначально оценивалась скорость исходящего трафика пользователя s_1 в сети, модель которой представлена графом G_0 . График скорости передачи данных пользователем s_1 изображен на рис. 6.

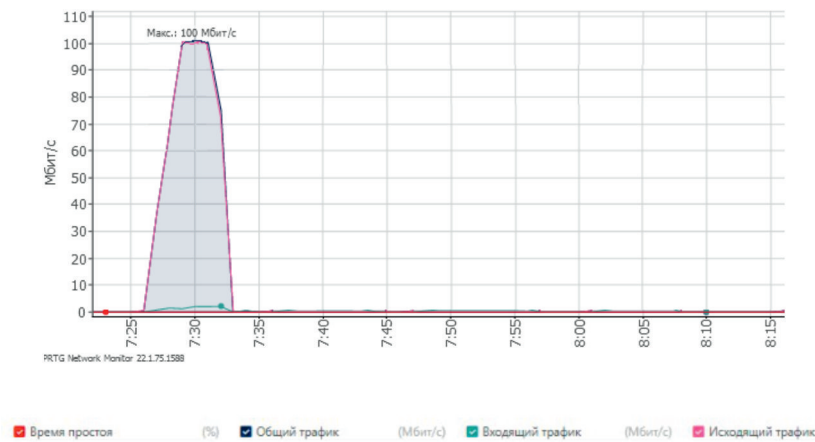


Рис. 6. Скорость исходящего трафика пользователя s_1

Fig. 6. User outbound traffic speed s_1

Из графика видно, что $v_{n1} \approx 100$ Мбит/с $\Rightarrow q_{s1} = q_{s2} = 0,05$; $q_{s3} = q_{s4} = q_{s5} = 1$. Поскольку все функции полезности элементов системы возросли, утверждение о полном синергизме в системе подтвердилось, т. е. $s_i > |_c s_j$ для всех i, j .

При осуществлении DoS-атаки был получен график скорости исходящего трафика пользователя s_1 (рис. 7), из которого видно, что средняя скорость v_{n1} составляет 51 Мбит/с. Такая ситуация соответствует модели G_1 (рис. 3).

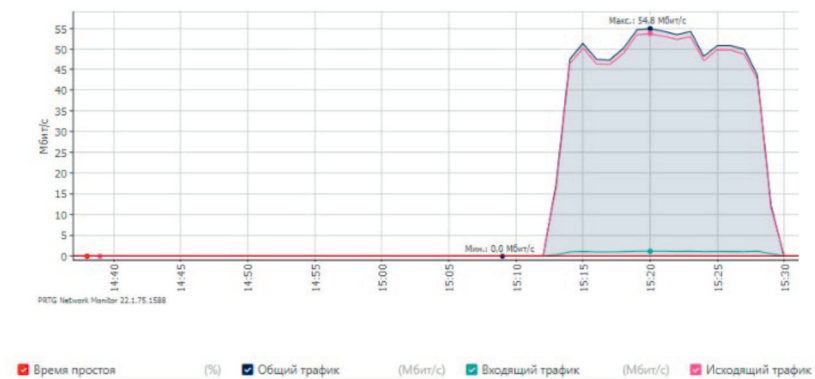


Рис. 7. Скорость исходящего трафика от пользователя s_1

Fig. 7. User outbound traffic speed s_1

Рассчитаем значения функций q_{si} по формулам (2) и (3):

$$q_{s1} = q_{s2} = \frac{51}{2 \cdot 100} = 0,255;$$

$$q_{s3} = q_{s4} = q_{s5} = \frac{1}{100} \cdot 51 = 0,51;$$

$$q_{s6} = \frac{49}{100} = 0,49.$$

Таким образом, элементы модели G_1 имеют значения функций полезности меньше, чем в модели G_0 .

Если рассматривать ситуацию, соответствующую модели G_2 (рис. 4), то при появлении в структуре системы инженера им выполняется комплекс мероприятий, направленных на противодействие возникшей угрозе, в частности DoS-атаке. Воздействие инженера на коммутатор будет осуществлено в ходе процесса управления и заключаться в ограничении скорости трафика на порту, к которому осуществил подключение нарушитель. В ходе эксперимента скорость входящего и исходящего трафика на порту была ограничена до 6 Мбит/с. В результате график скорости исходящего трафика от пользователя s_1 будет выглядеть следующим образом (рис. 8).

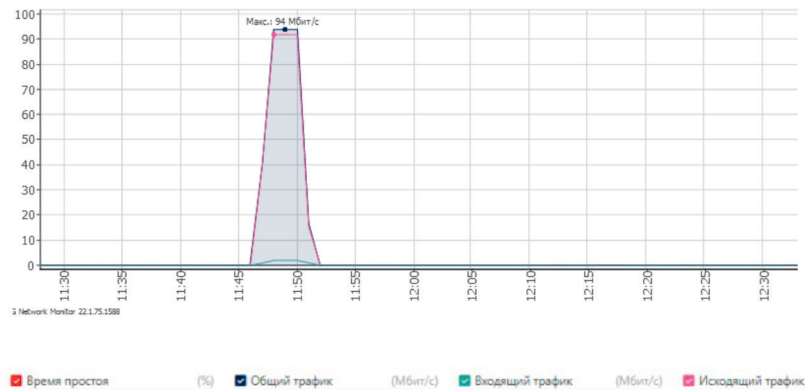


Рис. 8. Скорость исходящего трафика от пользователя s_1
Fig. 8. User outbound traffic speed s_1

Рассчитаем значения функций полезности элементов такой системы по формулам (2) – (5):

$$q_{s_1} = q_{s_2} = \frac{94}{2 \cdot 100} = 0,47; \quad q_{s_3} = q_{s_4} = q_{s_5} = \frac{1}{100} \cdot 94 = 0,94; \quad q_{s_6} = \frac{7}{100} = 0,07; \quad q_{s_7} = 0,47 \cdot 2 = 0,94.$$

Если сравнивать результаты вычислений для моделей G_1 и G_2 , то видно, что:

$$q_{s_i} \uparrow, \text{ для всех } i \neq 6; \quad q_{s_6} \downarrow.$$

Отсюда следует справедливость заключения о том, что $s_7 > |_c s_3$. Таким образом, при появлении элемента s_7 в системе, деятельность которого направлена на противодействие нарушителю, повышается эффективность функционирования системы.

В ходе эксперимента аналогично трем рассмотренным случаям производилась оценка скорости исходящего трафика пользователя s_1 , а также рассчитывались такие показатели, как t_d и CPU. Соответственно для формирования набора статистических данных были построены 63 графика скорости передачи данных пользователем s_1 для различных моделей СС СН, 8 из которых представлены на рис. 9.

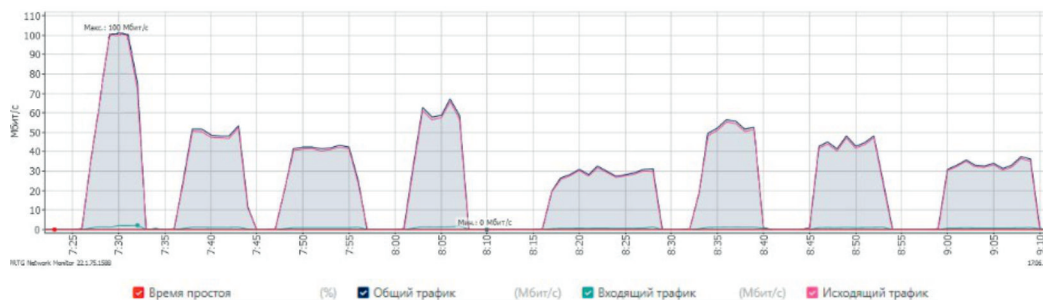


Рис. 9. Исходящий трафик пользователя s_1
Fig. 9. User outbound traffic speed s_1

Первый график соответствует модели G_0 ; второй график – модели G_1 ; третий – DOS-атака на Switch 2; четвертый – DOS-атака на Switch 3; пятый – DDOS-атака на все три коммутатора; шестой – DDOS-атаки на Switch 2 и Switch 3; седьмой – DDOS-атаки на Switch 1 и Switch 3; восьмой – DDOS-атака на Switch 1 и Switch 2.

В ходе эксперимента были получены наборы данных, определяемые множествами:

$$v_{n1} = \{v_{n1}^1, v_{n1}^2, \dots, v_{n1}^{63}\};$$

$$t_{\partial} = \{t_{\partial}^1, t_{\partial}^2, \dots, t_{\partial}^{63}\};$$

$$CPU = \{CPU^1, CPU^2, \dots, CPU^{63}\}.$$

Для сокращения размерности решаемой задачи на основе данных, полученных в ходе натурного эксперимента, было получено подтверждение существования зависимости между элементами множества Φ . Для анализа данных использовались такие методы, как регрессионный анализ и канонический корреляционный анализ, реализованные как модули программы STATISTICA 10 [35]. При сравнении значений показателя v_{n1} с t_{∂} и CPU в процессе регрессионного анализа были получены следующие графики (рис. 10 и 11):

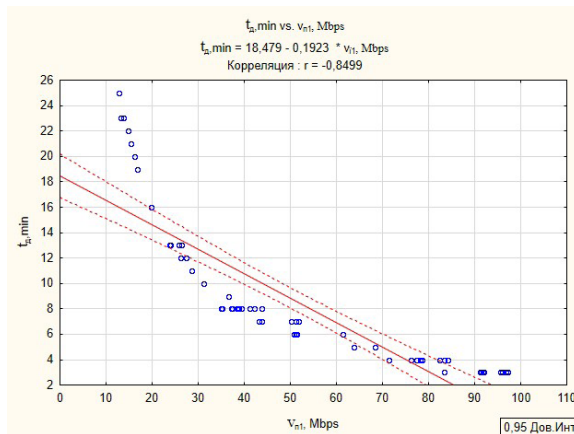


Рис. 10. Диаграмма рассеяния значений v_{n1} и t_{∂}
Fig. 10. Scatter diagram of values v_{n1} and t_{∂}

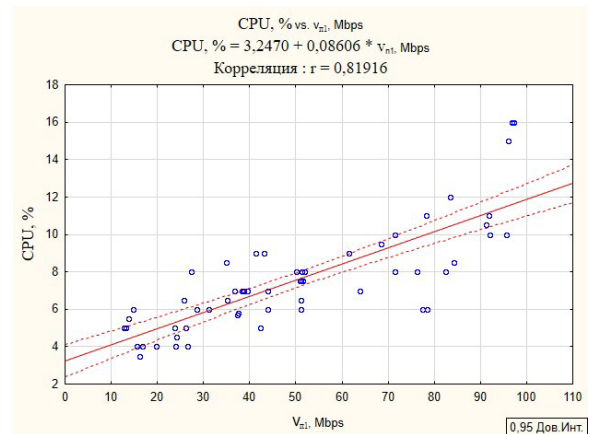


Рис. 11. Диаграмма рассеяния значений v_{n1} и CPU
Fig. 11. Scatter diagram of values v_{n1} and CPU

Как видно из рис. 10 и 11, значение модуля коэффициента корреляции $|r| > 0,8$, что свидетельствует о наличии высокой корреляционной связи показателей.

Таким образом, можно прийти к выводу о том, что для установления взаимосвязи между конфликтностью ОТС и эффективностью их функционирования достаточным будет использование одного из показателей множества Φ , например, скорости передачи данных v_{n1} .

3. Исследование взаимосвязи между показателями эффективности и оценками конфликтности организационно-технических систем

Для проверки гипотезы о наличии связи между оценками конфликтности и показателями эффективности найдем значения следующих оценок конфликтности:

$$\Pi = \{M, M_c, M_s, M_n, M_m, \bar{P}_b, \bar{P}_j, X_{pz}, \bar{I}P_0, \bar{I}P_s, \bar{C}P_s, \bar{I}P_0, \bar{I}P_s, \bar{C}P_s, Y_{in}\},$$

где $\bar{P}_i = \{\bar{P}_1, \bar{P}_2, \dots, \bar{P}_{|V_0|}\}$, $\bar{P}_j = \{\bar{P}_1, \bar{P}_2, \dots, \bar{P}_{|V_0|}\}$.

С учетом 63 рассмотренных моделей каждая оценка будет иметь множество значений с индексами в диапазоне от 1 до 63, например, $M = \{M^1, M^2, \dots, M^{63}\}$.

В ходе расчетов значений оценок конфликтности были получены следующие результаты.

1. При выявлении корреляционной зависимости между множеством значений v_{n1} и множествами значений оценок M_c и M_s были получены графики, представленные на рис. 12 и 13 соответственно.

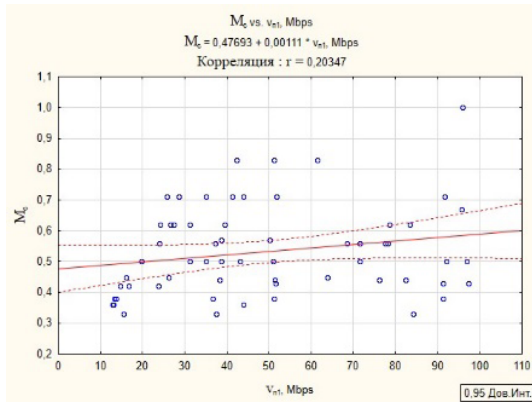


Рис. 12. Диаграмма рассеяния значений v_{n1} и M_c
Fig. 12. Scatter diagram of values v_{n1} and M_c

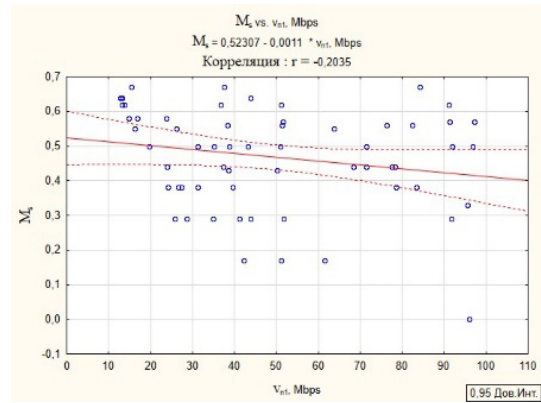


Рис. 13. Диаграмма рассеяния значений v_{n1} и M_s
Fig. 13. Scatter diagram of values v_{n1} and M_s

$$\begin{cases} M^i = 1 \\ M_n^i = 0, \forall i. \\ M_m^i = 1 \end{cases} \quad (23)$$

Из рис. 12, 13 и выражения (23) следует вывод о том, что зависимость между показателем v_{n1} и оценками M , M_c , M_s , M_n , M_m отсутствует. Это обусловлено тем, что при построении и исследовании графовых моделей G_i их циклическая структура не менялась, т. е. при добавлении элементов s_i , $i > 5$ появлялись только односторонние отношения на основные пять элементов без формирования отношений обратной связи.

2. При анализе взаимосвязи между v_{n1} и множествами элементов $\bar{P}_i$ и $\bar{\bar{P}}_i$ для всех $i \leq 5$ (т. е. для элементов, входящих в состав базовой модели) были получены следующие значения коэффициентов корреляции $r(v_{n1}, \bar{P}_i \vee \bar{\bar{P}}_i)$ (табл.):

Значения коэффициента корреляции

Correlation coefficient values

i	$r(v_{n1}, \bar{P}_i)$	$r(v_{n1}, \bar{\bar{P}}_i)$
1	0,286	0,588
2	0,286	0,588
3	0,308	0,692
4	-0,177	0,220
5	-0,06	0,311

Из таблицы видно, что корреляция между показателем v_{n1} и оценками влияний элементов на систему отсутствует, однако имеется средняя корреляционная связь между v_{n1} и влиянием

системы на первый, второй и третий элементы, т. е. на пользователей и коммутатор, к которому подключен пользователь, передающий сообщение.

3. При выявлении зависимости между v_{n1} и значениями оценок $\bar{IP}_6, \bar{IP}_s, C\bar{P}_s, \bar{IP}_6, \bar{IP}_s, C\bar{P}_s$ были получены следующие значения коэффициентов корреляции: $r(v_{n1}, \bar{IP}_6) = 0,057$; $r(v_{n1}, \bar{IP}_s) = 0,518$; $r(v_{n1}, C\bar{P}_s) = 0,474$; $r(v_{n1}, \bar{IP}_6) = 0,519$; $r(v_{n1}, \bar{IP}_s) = 0,523$; $r(v_{n1}, C\bar{P}_s) = 0,475$.

Диаграммы рассеяния значений трех оценок, обладающих наибольшим коэффициентом корреляции ($r > 0,5$), представлены на рис. 14, 15 и 16:

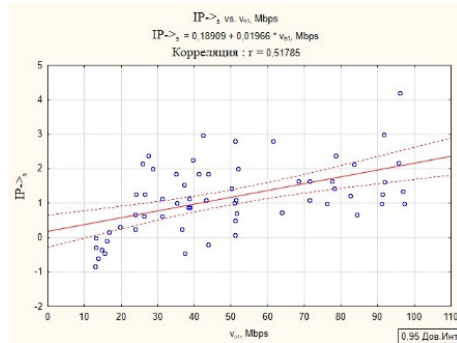


Рис. 14. Диаграмма рассеяния значений v_{n1} и $\bar{IP}_s$
Fig. 14. Scatter diagram of values v_{n1} and $\bar{IP}_s$

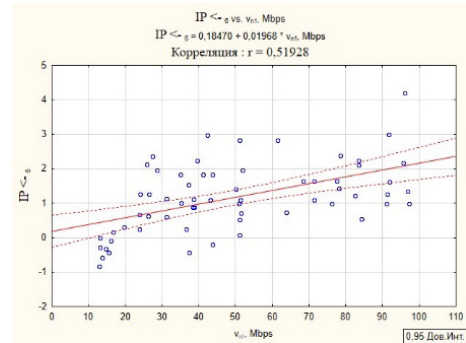


Рис. 15. Диаграмма рассеяния значений v_{n1} и $\bar{IP}_6$
Fig. 15. Scatter diagram of values v_{n1} and $\bar{IP}_6$

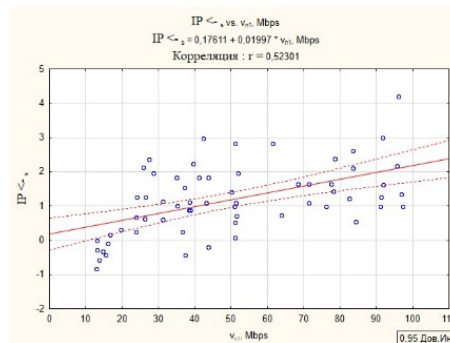


Рис. 16. Диаграмма рассеяния значений v_{n1} и $\bar{IP}_s$
Fig. 16. Scatter diagram of values v_{n1} and $\bar{IP}_s$

Значение коэффициента корреляции $r \approx 0,5$ не позволяет сделать вывод о наличии зависимости между v_{n1} и данной группой оценок.

4. Более сильная корреляционная связь выявляется при нахождении $r(v_{n1}, X_{pz}) = 0,737$, о чем свидетельствует диаграмма рассеяния, представленная на рис. 17.

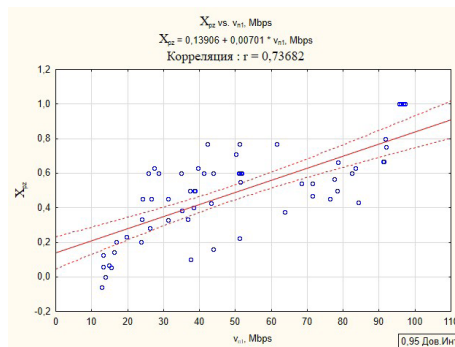


Рис. 17. Диаграмма рассеяния значений v_{n1} и X_{pz}
Fig. 17. Scatter diagram of values v_{n1} and X_{pz}

5. Более подробно рассмотрим расчет оценки Y_{in} в соответствии с формулой (22). Поскольку в ходе натурного эксперимента в целях противодействия конфликтному элементу осуществлялось снижение скорости на порту коммутатора, к которому он был подключен, то значение $\pi k(t)$, где s_k – инженер, может быть рассчитано следующим образом: если сравнить модели G_1 и G_2 , то q_{s6} изменилась с 0,49 до 0,07, т. е. в процессе воздействия со стороны элемента s_7 функция полезности конфликтного элемента уменьшилась на 85 %. Если бы инженер полностью блокировал сетевой трафик на порту, то полезность s_6 обратилась в 0. Из этого можно сделать вывод, что для рассматриваемой ситуации определим коэффициент силы синергетического влияния со стороны инженера статическим (для упрощения) и равным 0,85.

При определении коэффициентов α_k, β_k учтем, что при реализации процесса управления инженер оказывает противодействие сразу всем нарушителям, подключившимся к коммутатору, поэтому $\beta_k = E1, \alpha = 1$. После расчета значений $Y_{in}^i, i = [1, 2, \dots, 63]$ была получена диаграмма рассеяния значений v_{n1} и Y_{in} (рис. 18) и определено значение коэффициента корреляции $r(v_{n1}, Y_{in}) = 0,83$.

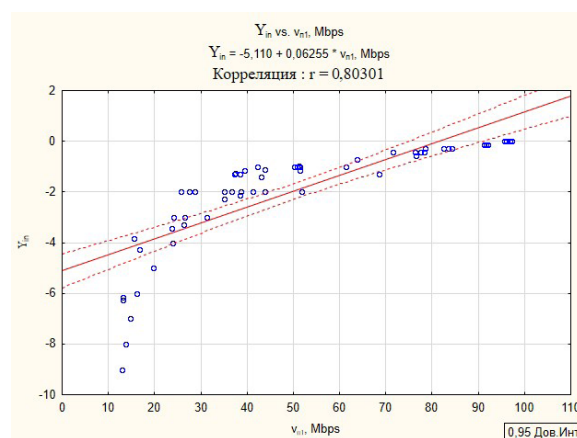


Рис. 18. Диаграмма рассеяния значений v_{n1} и Y_{in}

Fig. 18. Scatter diagram of values v_{n1} and Y_{in}

Таким образом, можно сделать вывод, что наиболее сильная взаимосвязь между показателями эффективности ОТС на примере СС СН и оценками конфликтности прослеживается в тех случаях, когда расчет оценки производится с учетом силы взаимных воздействий элементов друг на друга, а также при ограничении области расчетов основной компонентой связности графовой модели, внутри которой реализуются процессы, направленные на достижение цели функционирования исследуемой системы. Сделанный вывод подтверждают значения $r(v_{n1}, X_{pz})$ и $r(v_{n1}, Y_{in})$. Благодаря выявленной взаимосвязи становится возможным с некоторой вероятностью оценить эффективность функционирования реальной ОТС по ее графовой модели, и наоборот, по значениям показателей эффективности определить преобладающие типы отношений между элементами системы.

Заключение

В ходе исследования была осуществлена разработка моделей ОТС на примере СС СН с учетом типов отношений между их элементами. Теоретическая значимость полученных результатов заключается в подтверждении гипотезы о существовании взаимосвязи между показателями эффективности функционирования ОТС и рядом оценок конфликтности. Результат исследования корреляционных зависимостей позволил выбрать из множества оценок кон-

фликтности те, которые могут быть использованы для предсказания и оценки эффективности функционирования системы исходя из ее структурных особенностей. Наличие взаимосвязи позволит на практике при проектировании ОТС осуществлять оценку эффективности их функционирования на основе оценок конфликтности их графовых моделей. Кроме того, появляется возможность при необходимости изменения состояния ОТС осуществлять поиск наиболее предпочтительных структур и путей перехода к ним.

Список литературы

1. **Алексеев В. В., Корыстин С. И., Малышев В. А., Сысоев В. В.** Моделирование информационного воздействия на эргатический элемент в эрготехнических системах. М.: Стенсвил. 2003. 200 с.
2. **Меньших В. В., Сысоев В. В.** Структурная адаптация систем управления. М.: ИПРЖР. 2002. 150 с.
3. **Алексеев В. В., Зайцев А. В., Лысункин П. С.** Методология повышения качества эргатического элемента в эрготехнических системах на основе искусственного интеллекта // Надежность и качество сложных систем. 2018. № 3. С. 17–22. <https://doi.org/10.21685/2307-4205-2018-3-3>.
4. **Талах А. Н., Алексеев В. В., Жуков А. О., Кулаков А. В., Тютюнник В. М.** Математическая модель управления качеством информационного процесса в эрготехнических системах // Инженерная физика. 2018. № 2. С. 25–30. <https://www.elibrary.ru/item.asp?id=32612824>.
5. **Емелин Н. М.** Подходы к оценке влияния безошибочности действий персонала на надежность эргатических систем // Известия Института инженерной физики. 2022. № 1. С. 112–114. <https://www.elibrary.ru/item.asp?id=47988760>.
6. **Меньших В. В., Зверев Г. И.** Выбор функций элементов эргатических систем для обеспечения их функционирования в условиях деструктивных воздействий // Моделирование, оптимизация и информационные технологии. 2020. Т. 8. № 3. <https://doi.org/10.26102/2310-6018/2020.30.3.004>.
7. **Асяев И. Ю., Данилов А. М.** Моделирование эргатических систем // Образование и наука в современном мире. Инновации. 2021. № 5. С. 68–76. <https://www.elibrary.ru/item.asp?id=46566654>.
8. **Железнов Э. Г., Комиссаров П. В., Цымай Ю. В.** Исследование эргатических систем управления // Современные наукоемкие технологии. 2021. № 4. С. 37–41. <https://doi.org/10.17513/snt.38612>.
9. **Krasnyanskiy M. N., Obukhov A. D., Dedov D. L.** Problem formulation for determining the effectiveness of development and application of adaptive training systems for ergatic systems, {it International Review on Modelling and Simulations}, 2020, vol. 13, no. 3, pp. 159–169. <https://doi.org/10.15866/iremos.v13i3.17275>.
10. **Obukhov A. D., Dedov D. L., Arkhipov A. E.** Development of structural model of adaptive training complex in ergatic systems for professional use, {it IOP Conference Series: Materials Science and Engineering, Tomsk, 04–06 December 2017}, Tomsk: Institute of Physics Publishing, 2018. P. 022075. <https://doi.org/10.1088/1757-899X/327/2/022075>.
11. **Kulakov S. M., Kojnov R. S., Taraborina E. N.** On the functional structure of the ergatic system of precedent management of a complex production facility, {it Izvestiya Saratovskogo universiteta. Novaya seriya. Seriya: Matematika. Mekhanika. Informatika}, 2022, vol. 22, pp. 241–249. <https://doi.org/10.18500/1816-9791-2022-22-2-241-249>.
12. **Попов А. В.** Модель взаимодействия между элементами эрготехнической системы на примере сети связи специального назначения // Математические методы в технологиях и технике, 2022. № 3. С. 48–51. https://doi.org/10.52348/2712-8873_MMTT_2022_3_42.

13. **Попов А. В.** Декомпозиционный подход к построению модели эрготехнической системы на примере сети связи специального назначения // Инфокоммуникационные технологии, 2022. № 1. С. 8–17. <https://doi.org/10.18469/ikt.2022.20.1.01>.
14. **Микони С. В., Соколов Б. В., Юсупов Р. М.** Квалиметрия моделей и полимодельных комплексов // М.: РАН. 2018. 314 с.
15. **Попов А. В.** Сравнение методов оценки процессов эрготехнических систем органов внутренних дел // Автоматизация процессов управления, 2022. № 3 (69). С. 67–76. https://doi.org/10.35752/1991-2927_2022_3_69_67.
16. **Дружинин В. В., Конторов Д. С., Конторов М. Д.** Введение в теорию конфликта // М.: Радио и связь. 1989. 288 с.
17. **Syed A. Raza, Standing C.** A systemic model for managing and evaluating conflicts in organizational change. *Systemic Practice and Action Research*, 2011, vol. 24, no. 3, p. 187–210. <https://ro.ecu.edu.au/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1409&context=ecuworks2011>.
18. **Aijuan Zhang, Cheng Ji., Yu Bao, Xin L.** Conflict Analysis and Detection Based on Model Checking for Spatial Access Control Policy. *Tsinghua SCIENCE and technology*, 2017, vol. 22, no. 5, p. 478–488. <https://dc.tsinghuajournals.com/cgi/viewcontent.cgi?article=1166&context=tsinghua-science-and-technology>.
19. **Mitici M., Blom A. P. H.** Mathematical models for air traffic conflict and collision probability estimation. *IEEE Transactions on Intelligent Transportation Systems*, 2018, vol. 20, no. 3, p. 1052–1068. <http://dx.doi.org/10.1109/TITS.2018.2839344>.
20. **Bekesiene S., Meidute-Kavaliauskiene I., Samoilenko I., Nikitin A.** The Complex Systems for Conflict Interaction Modelling to Describe a Non-Trivial Epidemiological Situation. *Mathematics*, 2022, vol. 10, no. 4. <https://doi.org/10.3390/math10040537>.
21. **Kubiv S., Balanyuk Y.** Development of a mathematical model of conflict between the parties in the implementation of the offset transaction. *Technology audit and production reserves*, 2020, vol. 2, no. 52, p. 28–31. <http://doi.org/10.15587/2706-5448.2020.201260>.
22. **Kenetova R. O.** On the issue of mathematical modeling based on the theory of conflict. *Vestnik KRAUNC. Fiz.-Mat. Nauki*, 2021, vol. 36, no. 3, p. 72–79. <https://doi.org/10.26117/2079-6641-2021-36-3-72-79>.
23. **Petukhov A. Y., Malkhanov A. O., Sandalov V. M., Petukhov Y. V.** Modeling conflict in a social system using diffusion equations. *Simulation: Transactions of the Society for Modeling and Simulation International*, 2018, vol. 94, no. 12, p. 1053–1061. <https://doi.org/10.1177/0037549718761573>.
24. **Pochkhua G.** Computer modeling of political conflict resolution through economic cooperation. *Computer Sciences and Telecommunications*, 2019, no. 2, p. 22–35. <https://www.elibrary.ru/item.asp?id=45620381>.
25. **Goncharov N., Goncharov I., Dushkin A.** Assessment of the Conflict Stability of Information Systems Using Mathematical Modeling. 2019 International Multi-Conference on Industrial Engineering and Modern Technologies, FarEastCon 2019, Vladivostok, 01–04 October 2019. Vladivostok: Institute of Electrical and Electronics Engineers Inc, 2019, p. 8934011. <https://doi.org/10.1109/FarEastCon.2019.8934011>.
26. **Сысоев В. В.** Конфликт. Сотрудничество. Независимость. Системное взаимодействие в структурно-параметрическом представлении. М.: Московская академия экономики и права. 1999. 151 с.
27. **Светлов В. А.** Управление конфликтом. Новые технологии принятия решений в конфликтных ситуациях: учеб. пособие. Саратов: Ай Пи Эр Медиа, 2019. 136 с.
28. **Попов А. В.** Исследование структурных и конфликтных свойств систем с использованием знаковых графов // Актуальные проблемы прикладной математики, информатики и механики: сб. тр. Междунар. науч. конф. (г. Воронеж, 11–13 ноября 2019 г.). Воронеж:

- Научно-исследовательские публикации. 2020. С. 1050–1055. <https://www.elibrary.ru/item.asp?id=42493528>.
29. **Пьянков О. В.** Математическое моделирование информационно-аналитических систем органов внутренних дел: монография. Воронеж: Воронежский институт МВД России. 2013. 132 с.
 30. **Борисов В. В., Бычков И. А., Дементьев А. В., Соловьев А. П., Федулов А. С.** Компьютерная поддержка сложных организационно-технических систем. М.: Горячая линия – Телеком. 2002. 154 с.
 31. **Ananayev A. V., Ivannikov K. S.** Risk-model for communication networks operation stability assessment. *Journal of Physics: Conference Series: Current Problems*. Voronezh, 07–09 December 2020, Voronezh, 2021, p. 012028. <https://doi.org/10.1088/1742-6596/1902/1/012028>.
 32. **Thurley K., Wu L. F., Altschuler S. J.** Modeling Cell-to-Cell Communication Networks. Using Response-Time Distributions. *Cell Systems*, 2018, vol. 6, no. 3, p. 355–367. <https://doi.org/10.1016/j.cels.2018.01.016>.
 33. **Leung K. K., Massey W. A., Whitt W.** Traffic models for wireless communication networks. *IEEE Journal on Selected Areas in Communications*, 1994, vol. 12, no. 8, p. 1353–1364. <https://doi.org/10.1109/49.329340>.
 34. Официальный сайт компании Paessler. URL: <https://www.paessler.com/ru/prtg> (дата обращения 10.10.2022).
 35. **Боровиков В.** Statistica. Искусство анализа данных на компьютере: для профессионалов. 2-е изд. СПб.: Питер. 2003. 688 с.

References

1. **Alekseev V. V., Korystin S. I., Malyshev V. A., Sysoev V. V.** Modeling the information impact on the ergative element in ergotechnical systems. М.: Stensvil. 2003. 200 pp. (in Russ.)
2. **Menshikh V. V., Sysoev V. V.** Structural adaptation of control systems. М.: IPRZHR, 2002. 150 p. (in Russ.)
3. **Alekseev V. V., Zajcev A. V., Lysunkin P. S.** Methodology to improve the quality of the ergative element in ergotechnical systems based on artificial intelligence. *Nadezhnost' i kachestvo slozhnyh system*, 2018, no. 3, pp. 17–22. (in Russ.)
4. **Talah A. N., Alekseev V. V., Zhukov A. O., Kulakov A. V., Tyutyunnik V. M.** Mathematical model of information process quality management in ergotechnical systems. *Inzhenernaya fizika*, 2018, no. 2, pp. 25–30. (in Russ.)
5. **Emelin N. M.** Approaches to assessing the impact of error-free personnel actions on the reliability of ergative systems. *Izvestiya Instituta inzhenernoj fiziki*, 2022, no. 1, pp. 112–114. (in Russ.)
6. **Menshikh V. V., Zverev G. I.** Choice of functions of elements of ergative systems to ensure their functioning under destructive influences. *Modelirovanie, optimizaciya i informacionnye tekhnologii*, 2020, vol. 8, no. 3. (in Russ.)
7. **Asyaev I. Yu., Danilov A. M.** Modeling of ergatic systems. *Obrazovanie i nauka v sovremennom mire. Innovacii*, 2021, no. 5, pp. 68–76. (in Russ.)
8. **Zheleznev E. G., Komissarov P. V., Cymaj Yu. V.** Research on ergatic control systems. *Sovremennye naukoemkie tekhnologii*, 2021, no. 4, pp. 37–41. (in Russ.)
9. **Krasnyanskiy M. N., Obukhov A. D., Dedov D. L.** Problem formulation for determining the effectiveness of development and application of adaptive training systems for ergatic systems. *International Review on Modelling and Simulations*, 2020, vol. 13, no. 3, pp. 159–169.
10. **Obukhov A. D., Dedov D. L., Arkhipov A. E.** Development of structural model of adaptive training complex in ergatic systems for professional use. *IOP Conference Series: Materials Science and Engineering*, Tomsk, 04–06 December 2017. Tomsk: Institute of Physics Publishing. 2018. P. 022075.

11. 11. **Kulakov S. M., Kojnov R. S., Taraborina E. N.** On the functional structure of the ergative system of precedent management of a complex production object. *Izvestiya Saratovskogo universiteta. Novaya seriya. Seriya: Matematika. Mekhanika. Informatika*, 2022, vol. 22, pp. 241–249. (in Russ.)
12. 12. **Popov A. V.** Model of interaction between elements of an ergotechnical system by the example of a special-purpose communication network. *Matematicheskie metody v tekhnologiyah i tekhnike*, 2022, no. 3, pp. 48–51. (in Russ.)
13. 13. **Popov A. V.** Decomposition approach to constructing an ergotechnical system model on the example of a special purpose communication network. *Infokommunikacionnye tekhnologii*, 2022, no. 1, pp. 8–17. (in Russ.)
14. 14. **Mikoni S. V., Sokolov B. V., Yusupov R. M.** Qualimetry of models and polymodel complexes. M.: RAN, 2018. 314 p. (in Russ.)
15. 15. **Popov A. V.** Comparison of methods for assessing the processes of ergotechnical systems of internal affairs bodies. *Avtomatizatsiya processov upravleniya*, 2022, no. 3, pp. 67–76. (in Russ.)
16. 16. **Druzhinin V. V., Kontorov D. S., Kontorov M. D.** *Vvedenie v teoriyu konflikta* [Introduction to Conflict Theory]. M.: Radio i svyaz', 1989. 288 p. (in Russ.)
17. 17. **Raza S. A., Standing C.** A systemic model for managing and evaluating conflicts in organizational change. *Systemic Practice and Action Research*, 2011, vol. 24, no. 3, pp. 187–210.
18. 18. **Zhang A., Cheng J., Yu B., Xin L.** Conflict Analysis and Detection Based on Model Checking for Spatial Access Control Policy. *Tsinghua SCIENCE and technology*, 2017, vol. 22, no. 5, pp. 478–488.
19. 19. **Mitici M., Blom A. P. H.** Mathematical models for air traffic conflict and collision probability estimation. *IEEE Transactions on Intelligent Transportation Systems*, 2018, vol. 20, no. 3, pp. 1052–1068.
20. 20. **Bekesiene S., Meidute-Kavaliauskiene I., Samoilenko I., Nikitin A.** The Complex Systems for Conflict Interaction Modelling to Describe a Non-Trivial Epidemiological Situation. *Mathematics*, 2022, vol. 10, no. 4.
21. 21. **Kubiv S., Balanyuk Y.** Development of a mathematical model of conflict between the parties in the implementation of the offset transaction. *Technology audit and production reserves*, 2020, vol. 2, no. 52, pp. 28–31.
22. 22. **Kenetova R. O.** On the issue of mathematical modeling based on the theory of conflict. *Vestnik KRAUNC. Fiz.-Mat. Nauki*, 2021, vol. 36, no. 3, pp. 72–79.
23. 23. **Petukhov A. Y., Malkhanov A. O., Sandalov V. M., Petukhov Y. V.** Modeling conflict in a social system using diffusion equations. *Simulation: Transactions of the Society for Modeling and Simulation International*, 2018, vol. 94, no. 12, pp. 1053–1061.
24. 24. **Pochkhua G.** Computer modeling of political conflict resolution through economic cooperation. *Computer Sciences and Telecommunications*, 2019, no. 2, pp. 22–35.
25. **Goncharov N., Goncharov I., Dushkin A.** Assessment of the Conflict Stability of Information Systems Using Mathematical Modeling. 2019 International Multi-Conference on Industrial Engineering and Modern Technologies, FarEastCon 2019, Vladivostok, 01–04 October 2019. Vladivostok: Institute of Electrical and Electronics Engineers Inc., 2019. Art. 8934011.
26. **Sysoev V. V.** Conflict. Cooperation. Independence. System interaction in structural-parametric representation. M.: Moskovskaya akademiya ekonomiki i prava, 1999. 151 p. (in Russ.)
27. **Svetlov V. A.** Conflict Management. New Technologies of Decision-Making in Conflict Situations. Saratov: IPR Media, 2019. 136 p. (in Russ.)
28. **Popov A. V.** Study of structural and conflict properties of systems using symbolic graphs. Current Problems of Applied Mathematics, Computer Science, and Mechanics: Collected papers. Voronezh: Nauchno-issledovatel'skie publikacii. 2020, pp. 1050–1055. (in Russ.)
29. **Pyankov O. V.** Mathematical modeling of information-analytical systems of internal affairs bodies. Voronezh: Voronezhskij institut MVD Rossii, 2013. 132 p. (in Russ.)

30. **Borisov V. V., Bychkov I. A., Dement'ev A. V., Solov'ev A. P., Fedulov A. S.** Computer support for complex organizational and technical systems. M.: Goryachaya liniya – Telekom, 2002. 154 p. (in Russ.)
31. **Ananev A. V., Ivannikov K. S.** Risk-model for communication networks operation stability assessment. *Journal of Physics: Conference Series: Current Problems. Voronezh, 07–09 December 2020*, 2021, P. 012028.
32. **Thurley K., Wu L. F., Altschuler S. J.** Modeling Cell-to-Cell Communication Networks. Using Response-Time Distributions. *Cell Systems*, 2018, vol. 6, no. 3, pp. 355–367.
33. **Leung K. K., Massey W. A., Whitt W.** Traffic models for wireless communication networks. *IEEE Journal on Selected Areas in Communications*, 1994., vol. 12, no. 8, pp. 1353–1364.
34. Paessler official web site [Online]. URL: www.paessler.com/ru/prtg (accessed on: 10.07.2022). (in Russ.)
35. **Borovikov V.** Statistica. The Art of Computer Data Analysis: For Professionals. 2nd ed. SPb.: Piter. 2003. 688 p. (in Russ.)

Информация об авторах

Алексей Вячеславович Попов, адъюнкт

Олег Викторович Пьянков, доктор технических наук, доцент, заместитель начальника кафедры

Information about the Authors

Aleksey V. Popov, post-graduate, Voronezh Institute of the Ministry of Internal Affairs of Russia

Oleg V. Pyankov, Doctor of Engineering, Associate Professor, Deputy Chief of Department, Voronezh Institute of the Ministry of Internal Affairs of Russia

*Статья поступила в редакцию 09.12.2022;
одобрена после рецензирования 26.01.2023; принята к публикации 26.01.2023*

*The article was submitted 09.12.2022;
approved after reviewing 26.01.2023; accepted for publication 26.01.2023*

Метод стеганоанализа JPEG-изображений на основе цепей Маркова и его применение в сочетании с различными алгоритмами машинного обучения

Александра Владимировна Прокофьева¹
Алексей Николаевич Шниперов²

^{1,2} Сибирский федеральный университет
Красноярск, Россия

¹prokofe-aleksandra@yandex.ru, <https://orcid.org/0000-0001-5104-4511>

²ashnipervov@sfu-kras.ru, <https://orcid.org/0000-0003-4231-9805>

Аннотация

В работе предложен метод нахождения вектора характеристик изображений, позволяющий эффективно детектировать наличие скрытой информации в изображениях формата JPEG, встроенной различными популярными инструментами стеганографии. Данный метод основан на использовании матриц переходных вероятностей. Кроме того, в работе выполнена сравнительная оценка применения различных технологий машинного обучения для решения задачи стеганоанализа статических изображений формата JPEG, а именно: деревьев решений с градиентным бустингом, линейных моделей, метода k-ближайших соседей, метода опорных векторов, нейронных сетей и искусственных иммунных систем. Приведены результаты качества классификации каждым из вышеперечисленных методов. Сущность метода нахождения вектора характеристик изображения заключается в использовании матрицы переходных вероятностей и применении метода калибровки изображения для повышения точности стеганоанализа и уменьшения числа ложных срабатываний. Для каждого изображения из обучающей и тестовой выборки таким способом находится вектор его характеристик, число элементов которого составляет 324. Далее на полученных данных из обучающей выборки производилось обучение моделей каждым из вышеперечисленных методов машинного обучения отдельно. Тестирование качества построенной модели осуществлялось на данных тестовой выборки также для каждого алгоритма отдельно. Для оценки качества моделей использовались следующие метрики: точность, величина ошибки первого и второго рода результатов бинарной классификации, а также время классификации одного изображения. Для обучения и тестирования методов была использована выборка изображений IStego100K, состоящая из 208 тысяч изображений одинакового размера 1024×1024 с различными значениями качества JPEG из диапазона от 75 до 95. Для встраивания скрытого сообщения использовался один из трех алгоритмов стеганографии: J-UNIWARD, nsF5 и UERD. Результатом проведенного исследования является подтверждение того, что предложенный подход нахождения вектора характеристик изображения позволяет детектировать наличие скрытого вложения в изображениях, полученных в результате применения неадаптивных методов стеганографии (Steghide, OutGuess и nsF5) с очень высокой точностью, более 95 %. Для заполненных контейнеров, полученных в результате встраивания сообщения одним из адаптивных методов (J-UNIWARD, UERD), показатели точности обнаружения находятся в пределах 50–60 %. Практическая значимость заключается в экспериментальных данных, подтверждающих эффективность метода стеганоанализа в отношении детектирования скрытой информации в изображениях формата JPEG. Результаты работы могут быть полезны исследователям в области стеганографии и стеганоанализа для сравни-

тельного анализа применения технологий машинного обучения для решения задачи обнаружения наличия скрытого вложения в изображениях формата JPEG.

Ключевые слова

стеганоанализ, стеганография J-UNIWARD, nsF5, UERD, бинарная классификация, метод k-ближайших соседей, метод опорных векторов, нейронные сети, искусственные иммунные системы

Для цитирования

Прокофьева А. В., Шниперов А. Н. Метод стеганоанализа JPEG-изображений на основе цепей Маркова и его применение в сочетании с различными алгоритмами машинного обучения // Вестник НГУ. Серия: Информационные технологии. 2022. Т. 20, № 4. С. 61–75. DOI 10.25205/1818-7900-2022-20-4-61-75

A Markov Chain – Based Method for JPEG Image Steganalysis and Its Application in Combination with Various Machine Learning Algorithms

Aleksandra V. Prokofieva¹

Alexey N. Shniperov²

^{1,2}Siberian Federal University
Krasnoyarsk, Russian Federation

¹prokofe-aleksandra@yandex.ru, <https://orcid.org/0000-0001-5104-4511>

²ashniperov@sfu-kras.ru, <https://orcid.org/0000-0003-4231-9805>

Abstract

The paper proposes a method of extracting the feature vector of images, which makes it possible to effectively detect the presence of hidden information in JPEG images embedded by various popular steganography tools. This method is based on the usage of the transition probability matrix. The essence of the method for extracting the feature vector of the image is to use the transition probability matrix and apply the image calibration method to improve the accuracy of steganalysis and reduce the number of false positives. For each image from the training and test sets a feature vector is found in this way, the number of elements is 324. Further, the models were trained on the training dataset by each of machine learning methods separately: decision trees with gradient boosting, linear models, k-nearest neighbors, support vector machines, neural networks, and artificial immune systems. To assess the capacity of the models the following metrics were used: accuracy, the rate of the false positive and false negative errors, and the confusion matrix. The results of classification by each of the above methods are given. For training and testing a dataset IStego100K was used, which consists of 208 thousand images of the same size 1024×1024 with different quality values in the range from 75 to 95. One of the J-UNIWARD, nsF5, and UERD steganography algorithms was used to embed a hidden message. As a result, we can observe that the proposed approach to extracting the feature vector makes it possible to detect the presence of hidden information embedded by non-adaptive steganography (Steghide, OutGuess and nsF5) in static JPEG images with high accuracy (more than 95%). However, for adaptive steganography methods (J-UNIWARD, UERD) the accuracy is less (about 50-60%).

Keywords

stegananalysis, steganography, J-UNIWARD, nsF5, UERD, binary classification, k-nearest neighbors, support vector machines, neural networks, artificial immune systems

For citation

Prokofieva A. V., Shniperov A. N. A Markov Chain – Based Method for JPEG Image Steganalysis and Its Application in Combination with Various Machine Learning Algorithms. *Vestnik NSU. Series: Information Technologies*, 2022, vol. 20, no. 4, pp. 61–75. (in Russ.) DOI 10.25205/1818-7900-2022-20-4-61-75

Введение

Большая часть последних исследований в области стеганографии и стеганоанализа сосредоточилась на изображениях формата JPEG ввиду широкого использования этого формата

для передачи изображений в сети Интернет. Высокая популярность данного формата изображений привела к появлению множества новых методов скрытой передачи информации, в которых в качестве стеганоконтейнера выступает изображение. Такие стеганоконтейнеры могут содержать в себе противоправную информацию (к примеру, экстремистского или криминального характера), также с помощью таких методов возможно обойти мониторинг средств защиты информации, таких как средств антивирусной защиты и средств предотвращения утечек конфиденциальной информации (DLP-систем). Ввиду всего вышесказанного можно говорить о высокой актуальности задачи стеганоанализа изображений формата JPEG, развитии новых методов обнаружения скрытых каналов передачи информации.

На сегодняшний день существует большое количество методов стеганоанализа, которые различаются по анализируемым характеристикам изображения и методам встраивания, которым они противодействуют. Методы стеганоанализа традиционно разделяют на сигнатурные, статистические и эвристические. Также можно отдельно выделить четвертый тип: метод визуального поиска, который предполагает визуальный просмотр каждого изображения, а также различные преобразования изображения в пространственной области (преобразование яркости, цветовая коррекция и т. д.). Однако поскольку этот подход сильно зависит от того изображения, в которое производится встраивание скрытого сообщения, и требует непосредственного участия эксперта для анализа изображения, его сложно применять на практике в условиях большого трафика изображений. Сигнатурные методы стеганоанализа предназначены для работы с форматными методами скрытой передачи информации, которые встраивают скрытое сообщение в определенные структурой файла места (например, поле комментария формата файла JPEG) или в процессе встраивания оставляют специфические сигнатуры, по которым удается детектировать скрытое вложение. Статистические методы стеганоанализа основываются на анализе статистических характеристик исследуемого изображения, их корреляции с характеристиками пустых стеганоконтейнеров такого же типа. Наиболее известными статистическими методами являются RS-стеганоанализ и WS-стеганоанализ [1], гистограммный [2] и другие подходы. Данные методы могут показывать очень высокие показатели точности обнаружения заполненных стеганоконтейнеров, а также по определению объема скрытого вложения, однако их точность в значительной степени зависит от метода стеганографии, которым они противодействуют (они могут отлично работать на некоторых методах и совершенно не работать против любых других). Эвристические методы стеганоанализа представляют большой интерес для исследователей, поскольку они не привязаны к какому-то определенному алгоритму встраивания скрытой информации, хоть и в целом не достигают таких показателей, как статистические. Однако для использования на практике, когда алгоритм скрытой передачи информации не известен, они применимы в большей степени. В основном данные методы базируются на решении задачи бинарной классификации с применением методов машинного обучения, например, методы, предложенные в работах [3; 4; 5]. Одной из последних значимых работ является атака обратной совместимости [6], которая основывается на том факте, что изменения в квантованных коэффициентах дискретного косинусного преобразования (ДКП), появившиеся в результате встраивания скрытого сообщения, увеличивают дисперсию гауссова распределения ошибок округления в пространственной области. Этот метод позволяет надежно обнаруживать заполненные стеганоконтейнеры, полученные любым методом стеганографии, даже при небольшом размере скрытого вложения. Однако данный метод работает только на изображениях JPEG с качеством 99 и 100.

В работе [7] нами уже был представлен метод стеганоанализа изображений, основанный на применении искусственных иммунных систем (ИИС), в котором на этапе предобработки изображения вектор характеристик вычислялся с помощью вейвлет-преобразования Хаара, длина такого вектора характеристик для квадратных изображений была равна 36, для прямоугольных – 54. Однако этот метод обладал некоторыми недостатками:

– недостаточная точность классификации изображений (75–80 %). Для применения в реальных информационных системах на практике данный показатель недостаточен, поскольку приводит к высокой вероятности ложных срабатываний и пропусков события передачи изображения, содержащего скрытое вложение;

– продолжительное время обучения искусственной иммунной системы в следующих условиях работы: для обучения и тестирования ИИС использовалась небольшая база изображений, состоящая всего из 7,5 тыс. изображений. Осуществлялось только 10 итераций алгоритма клональной селекции на этапе обучения, в программной реализации не использовались никакие методы распараллеливания вычислений. Поскольку в данной работе для построения метода классификации изображений (после предобработки) использовался только метод на основе искусственных иммунных систем, мы не можем получить полное представление о его эффективности и влиянии на результаты обнаружения изображений – заполненных контейнеров, а также влиянии на время классификации одного изображения.

Поэтому основной целью настоящей работы является исправление существующих недостатков метода стеганоанализа на основе ИИС, а именно: повышение точности классификации изображения путем модификации метода предобработки и получения вектора характеристик изображения, а также проведение сравнительной оценки применения различных иных технологий машинного обучения, благодаря чему можно будет сделать выводы об эффективности применения метода на основе ИИС.

Помимо этого, среди опубликованных статей довольно редко встречаются работы, в которых проводится сравнение методов стеганоанализа на основе машинного обучения. В большинстве случаев оценка эффективности метода и результирующие показатели точности представляются непосредственно авторами статей. Проведение сравнительного анализа различных методов стеганоанализа на базе предоставленных авторами показателей довольно затруднительно, поскольку для оценки эффективности методов разными авторами используются различные выборки изображений, методы встраивания скрытого сообщения, различные метрики эффективности алгоритмов, и т. д. Таким образом, проведение сравнения эффективности различных методов стеганоанализа в одинаковых условиях может быть полезно для коллег, занимающихся исследованиями в данной области.

Описание методики проведения экспериментов

Итак, для начала сформулируем общую задачу метода стеганоанализа. Пусть $I = C \cup S$ – множество объектов заданного типа (изображений формата JPEG), S – множество заполненных стеганоконтейнеров, каждый из которых содержит скрытую информацию, C – множество пустых стеганоконтейнеров, не содержащих скрытой информации, полагаем $S \cap C = \emptyset$. Каждый из объектов $img \in I$ представлен вектором $\vec{D}$ его характеристик. Общая постановка задачи стеганоанализа изображения $img \in I$ заключается в решении задачи бинарной классификации наличия/отсутствия скрытого вложения.

Для обучения и тестирования методов классификации нами была использована база изображений IStego100K [8], состоящая из 208 104 изображений одинакового размера 1024×1024 , все изображения из обучающей и тестовой выборки взяты из одного источника. Среди них 200 тысяч изображений составляют обучающую выборку, а оставшиеся 8 104 – тестовую выборку (значения приведены суммарно для всех алгоритмов и значений коэффициента качества изображения, т. е. для разных алгоритмов используются разные изображения). Для каждого изображения в IStego100K коэффициенты качества JPEG различаются и находятся в диапазоне от 75 до 95. Для встраивания скрытого сообщения использовался один из трех алгоритмов стеганографии: J-UNIWARD, nsF5 и UERD. В базе изображений IStego100K есть различные наборы, отличающиеся величиной стеганографического встраивания: 0,1, 0,2, 0,3, 0,4 bpnz (бит на не-

нулевой АС-коэффициент ДКП), в данной работе для экспериментов использовались заполненные контейнеры, величина стеганографического воздействия в которых составляет 0,4 bpnz.

Алгоритм nsF5 [9] считается неадаптивным к содержимому изображения методом стеганографии, т. е. встраивание скрытого сообщения в изображение происходит последовательно. Но в отличие от других подобных методов, таких как JSteg, встраивание не так сильно влияет на изменение гистограмм коэффициентов ДКП, поэтому данный метод не подвержен основным статистическим атакам, и, тем самым, менее подвержен стеганоанализу.

Алгоритмы J-UNIWARD [10] и UERD [11] являются на данный момент самыми надежными, современными, а также не поддающимися детектированию обычными статистическими методами стеганоанализа. Стеганография с использованием минимальных искажений является наиболее успешной моделью адаптивной стеганографии, при которой сообщение встраивается в наиболее зашумлённые места изображения, и встраивание зависит от сложности текстуры. Алгоритм J-UNIWARD обладает выдающимися характеристиками резистивности к стеганоанализу, в то время как алгоритм UERD имеет меньшую сложность для обнаружения.

Получение вектора характеристик изображения

Этап получения векторов характеристик изображений является одним из ключевых в задаче стеганоанализа и напрямую может определять эффективность того или иного метода.

Поскольку метод, основанный на вычислении вейвлет-преобразования Хаара на практике показывал недостаточную точность классификации, как нам хотелось бы видеть для применения системы в реальных условиях, нами был модифицирован метод, основанный на нахождении матрицы переходных вероятностей изображения (subtractive pixel adjacency matrix), использовавшийся авторами в статьях [5; 9]. Вектор характеристик в таких методах представляет собой марковский процесс как разницу между абсолютными значениями соседних коэффициентов ДКП (дискретного косинусного преобразования), которому, как известно, подвергается каждый блок изображения при сжатии JPEG.

Итак, суть предлагаемого метода заключается в том, что на основе матрицы коэффициентов ДКП изображения img размером $N \times M$, согласно формулам (1), находятся четыре разностные матрицы соседних значений коэффициентов ДКП. На рис. 1 представлен принцип получения разностных матриц для горизонтального направления. На этом этапе они будут вычислены для четырех направлений: горизонтального, вертикального и диагонального направлений и направления побочной диагонали, которые обозначаются соответственно $F_h(u, v)$, $F_v(u, v)$, $F_d(u, v)$ и $F_m(u, v)$.

$$F_h(u, v) = F(u, v) - F(u + 1, v), \quad F_v(u, v) = F(u, v) - F(u, v + 1) \quad (1)$$

$$F_d(u, v) = F(u, v) - F(u + 1, v + 1), \quad F_m(u, v) = F(u + 1, v) - F(u, v + 1)$$

где $u \in [1, N - 1]$, $v \in [1, M - 1]$.

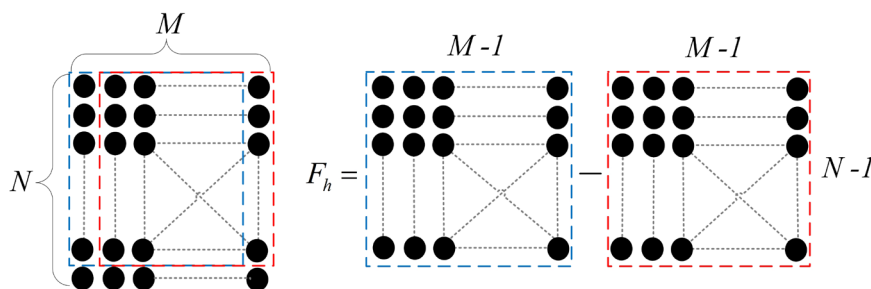


Рис. 1. Принцип получения разностной матрицы $F_h(u, v)$ в горизонтальном направлении

Fig.1. The principle of obtaining a difference matrix $F_h(u, v)$ in the horizontal direction

Далее на основе этих разностных матриц вычисляются матрицы переходных вероятностей, согласно формулам (2). Значение коэффициентов ДКП изображения прогнозируется исходя из значений соседних коэффициентов, а величина ошибки прогнозирования вычисляется путем вычитания значения предсказания из значения исходного коэффициента. Затем происходит сравнение с заданным пороговым значением.

Поэтому, прежде чем найти матрицы переходных вероятностей, сначала нужно задать порог T (примем $T = 4$, поскольку в результате экспериментов это значение позволяет получить лучшие результаты детектирования заполненных стеганоконтейнеров). Значения F_h , F_v , F_d и F_m округляются таким образом, чтобы они попадали в диапазон $[-T, T]$. По полученным значениям F_h , F_v , F_d и F_m рассчитываются матрицы вероятностей перехода $M_h(i, j)$, $M_v(i, j)$, $M_d(i, j)$ и $M_m(i, j)$ следующим образом:

$$M_h(i, j) = \frac{\sum_{v=1}^{N-1} \sum_{u=1}^{M-1} \delta[F_h(u, v) - i, F_h(u+1, v) - j]}{\sum_{v=1}^{N-1} \sum_{u=1}^{M-1} \delta[F_h(u, v) - i]}, \quad M_v(i, j) = \frac{\sum_{v=1}^{N-1} \sum_{u=1}^{M-1} \delta[F_v(u, v) - i, F_v(u, v+1) - j]}{\sum_{v=1}^{N-1} \sum_{u=1}^{M-1} \delta[F_v(u, v) - i]},$$

$$M_d(i, j) = \frac{\sum_{v=1}^{N-1} \sum_{u=1}^{M-1} \delta[F_d(u, v) - i, F_d(u+1, v+1) - j]}{\sum_{v=1}^{N-1} \sum_{u=1}^{M-1} \delta[F_d(u, v) - i]}, \quad M_m(i, j) = \frac{\sum_{v=1}^{N-1} \sum_{u=1}^{M-1} \delta[F_m(u+1, v) - i, F_m(u, v+1) - j]}{\sum_{v=1}^{N-1} \sum_{u=1}^{M-1} \delta[F_m(u+1, v) - i]},$$

где N, M – размеры изображения img , $i, j \in [-T, T]$; δ – дельта (символ) Кронекера (функция двух целых переменных, которая равна 1, если они равны, и 0 в противном случае).

Число элементов вектора характеристик составляет $4 \times (2T + 1)^2$. В нашем случае оно составляет 324, поскольку T выбрано равным 4.

Для повышения эффективности работы метода мы предлагаем использовать метод калибровки – преобразование изображения, позволяющее получить приблизительное отражение статистических свойств пустого стеганоконтейнера для анализируемого изображения. Анализируемое изображение img переводится в пространственную область (обратным дискретным косинусным преобразованием). Далее производится его обрезка на четыре пикселя с двух сторон и повторное сжатие полученного изображения с использованием матрицы квантования изображения img . В результате получается изображение img' формата JPEG, которое будет использоваться в качестве изображения – пустого контейнера. Далее для этого изображения img' находится вектор характеристик способом, рассмотренным выше для изображения img .

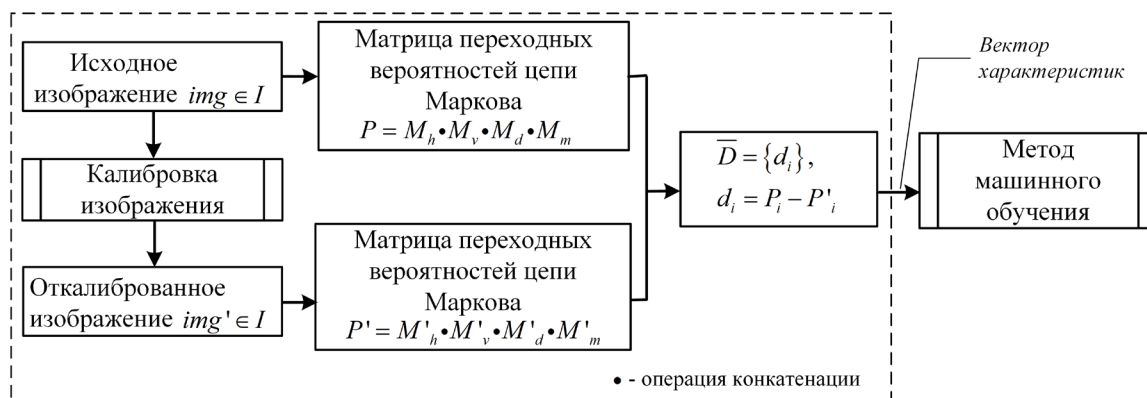


Рис. 2. Принцип получения вектора характеристик изображения

Fig. 2. The principle of obtaining the vector of image characteristics

Таблица 1

Результаты сравнения точности классификации для методов на основе вейвлет-преобразования Хаара и нахождении матрицы переходных вероятностей

Table 1

Results of Comparison of Classification Accuracy for Methods Based on the Haar's Wavelet Transform and Finding the Transition Probability Matrix

Алгоритм стеганографии Метод нахождения вектора характеристик	Steghide	OutGuess	nsF5
Точность обнаружения, %			
Вейвлет-преобразование Хаара [7]	71,1	74,5	74,9
Матрица переходных вероятностей	96,88	97,91	95,57
Величина ошибки I рода, %			
Вейвлет-преобразование Хаара [7]	26,3	16,5	17,1
Матрица переходных вероятностей	3,9	2,3	4,6
Величина ошибки II рода, %			
Вейвлет-преобразование Хаара [7]	31,5	34,5	32,9
Матрица переходных вероятностей	1,92	1,23	3,7

На следующем шаге каждая компонента итогового вектора характеристик изображения находится как разность соответствующих компонент векторов откалиброванного и исходного изображений, поскольку это позволяет снизить зависимость точности бинарной классификации от обучающей выборки изображений.

Итак, на рис. 2 приведена общая схема получения вектора характеристик изображения.

В результате проведенных экспериментов (приведены в табл. 1 и 2) мы можем наблюдать, что данный подход в выделение вектора характеристик изображения позволяет детектировать наличие скрытого вложения в лучшем из неадаптивных методов стеганографии (nsF5) с очень высокой точностью. Результаты сравнения работы данного метода нахождения векторов характеристик изображения и метода, основанного на нахождении вейвлет-преобразования Хаара (предложенного нами в предыдущей работе [7]), приведены в табл. 1. Мы можем видеть, насколько более эффективно данный метод стал работать на неадаптивных алгоритмах стеганографии, таких как nsF5, Steghide, OutGuess: точность обнаружения выросла более чем на 20 %.

Однако на адаптивных методах стеганографии и особенно на J-UNIWARD (который, к слову, хуже всех на данный момент поддается стеганоанализу) результаты пока что оставляют желать лучшего. И над увеличением точности детектирования адаптивных методов стеганографии мы работаем дальше.

Описание рассматриваемых методов машинного обучения

При решении задачи классификации часто возникает сложность с выбором методов машинного обучения, так как их существует достаточно много, а проверить качество работы каждого в отдельности не представляется возможным. При этом каждый из методов обладает своими слабыми и сильными сторонами, имеет свою область применимости.

В данной работе мы провели сравнительную оценку эффективности в задаче стеганоанализа самых популярных алгоритмов машинного обучения: деревьев решений с градиентным

бустингом, линейных моделей (реализованных в библиотеке LightAutoML [13]), метода k -ближайших соседей, метода опорных векторов (support vector machines – SVM), нейронных сетей, и искусственных иммунных систем [7].

В общем случае для каждого изображения из обучающей и тестовой выборки находится вектор характеристик (способом, описанным в предыдущем подразделе), состоящий из 324 значений. Далее для каждого алгоритма стеганографии, использованного для получения множества заполненных контейнеров выборки (J-UNIWARD, nsF5 и UERD), отдельно обучается модель (одним из вышеперечисленных методов машинного обучения) на векторах характеристик изображений из обучающей выборки. Завершающим этапом является тестирование качества построенной модели (точность классификации) на данных тестовой выборки.

Опишем подробнее, каким образом настраивались, подбирались параметры каждого из представленных методов в приведенных конкретных условиях и в контексте решения задачи стеганоанализа изображений.

Метод k -ближайших соседей

Как известно, метод k -ближайших соседей – алгоритм классификации, где принадлежность объекта к классу определяется за счет наименьшего расстояния между группой объектов известного класса. Ближайшие соседи находятся исходя из множества обучающей выборки, и по ключевому для данного метода значению k высчитывается, какой класс наиболее многочислен среди них. За счет вычисления расстояния между всеми объектами выполнение данного алгоритма занимает весьма продолжительное время и требует значительных ресурсов по используемой памяти. Однако данный алгоритм устойчив к выбросам, а единственным неизвестным параметром является k .

Для практической реализации данного метода использовалась библиотека scikit-learn на языке Python. Параметр k был определен с помощью метода локтя (elbow rule) – для этого мы построили график, отражающий зависимость между значениями k и величиной, равной проценту ошибочно классифицированных объектов (сумма величины ошибок первого и второго рода). Точка, где эта величина является минимальной, и считается оптимальным значением k . В нашем случае оптимальное значение $k = 2$ (см. рис. 3).

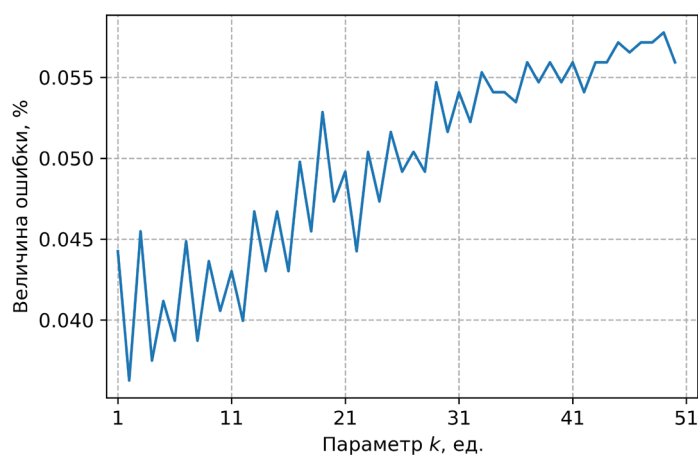


Рис. 3. Зависимость между значением k и величиной ошибочно классифицированных объектов тестовой выборки
Fig. 3. The dependence between the value of k and the value of the incorrectly classified objects of the test dataset

Деревья решений с градиентным бустингом, линейные модели (библиотека LightAutoML)

LightAutoML – это библиотека на языке Python, написанная сотрудниками лаборатории Artificial Intelligence Сбербанка, позволяющая автоматически построить модели машинного

обучения [13]. Данная библиотека обеспечивает оптимальный и быстрый поиск гиперпараметров с использованием методов итеративной оптимизации и байесовских методов. У библиотеки LightAutoML в арсенале два метода машинного обучения, которые она моделирует: деревья решений с градиентным бустингом (GBMs) и линейные модели. Разработчиками были выбраны именно эти алгоритмы машинного обучения, потому что несмотря на тенденцию в развитии нейронных сетей для разных областей, методы, основанные на GBMs, показывают высокие результаты производительности на табличных данных и превосходят другие подходы во многих тестах и соревнованиях. Кроме того, линейные модели быстры и могут повысить производительность моделей на основе деревьев решений в ансамблях [13].

Для обучения модели потребовалось выполнить несколько настроек. На первом шаге необходимо указать тип модели, в нашем случае это `binary` – бинарная классификация, поскольку для каждого изображения $img \in I$ требуется определить, относится ли оно к множеству заполненных стеганоконтейнеров S или множеству пустых стеганоконтейнеров C . На втором шаге необходимо импортировать *.csv файлы с векторами характеристик обучающей и тестовой выборки и добавить к ним колонку со значением принадлежности изображения к классу заполненных контейнеров (0 – если является пустым стеганоконтейнером, 1 – если заполненным). И на следующем шаге запустить функцию построения модели, в качестве параметров выбрано использование сочетания обоих вышеперечисленных алгоритмов. После чего остается только протестировать полученную модель на тестовой выборке изображений и оценить результаты (результаты приведены в табл. 2).

Метод опорных векторов (SVM)

Метод (машина) опорных векторов (SVM) – это расширение классификатора опорных векторов, которое является результатом определенного расширения пространства объектов с использованием ядер. Функции «ядра» позволяют вычислять различные разделяющие гиперплоскости, где под «ядром» подразумевается преобразование данных. Данный метод также был реализован с помощью библиотеки `sklearn.svm` на языке Python. Для экспериментов использовались несколько функций «ядра»: линейное ядро, радиальная базисная функция (RBF), полиномиальное ядро различных степеней, сигмоид-ядро. Наилучшие показатели точности для задачи стеганоанализа позволило получить использование радиальной базисной функции:

$$K(D_i, D'_i) = e^{-\gamma \sum_{n=1}^L (D_{ij} - D'_{ij})^2},$$

где D_i, D'_i – векторы характеристик двух изображений, L – длина вектора характеристик (в нашем случае $L = 324$), γ – некоторая положительная константа (для экспериментов использовалась $\gamma = 100$).

Нейронные сети

Нами была сконфигурирована нейронная сеть с помощью библиотеки Keras для Python, которая позволяет максимально ускорить и упростить процесс создания нейронных сетей, путем перебора подбирая наилучшую конфигурацию и параметры нейронной сети.

На первом шаге были импортированы данные обучающей и тестовой выборки и определили последовательный тип модели. Далее протестировали несколько конфигураций. Здесь приведем следующую, которая позволила достичь лучших показателей на тестовой выборке изображений: нейросеть с двумя скрытыми слоями: входной слой состоит из 480 нейронов, и каждый скрытый слой – из 112 нейронов, на выходном слое всего 1 нейрон, поскольку в нашем случае решается задача бинарной классификации.

В качестве функции активации на скрытых слоях использовалась `relu` (Rectified Linear Unit), на выходном слое использовали сигмоидную функцию, которая выполняет перенормировку значений в диапазоне от 0 до 1.

На этапе компилирования модели использовали следующие параметры: в качестве оптимизатора использовали стохастический градиентный спуск (SGD), так как это позволяет сократить задействованные вычислительные ресурсы и помогает достичь более высокой скорости обучения. В качестве функции потерь используем бинарную кросс-энтропию, в качестве метрики оценки – точность бинарной классификации.

Далее на этапе обучения использовались следующие параметры: фиксированный размер подмножества обучающей выборки (размер пакета, англ. batch size) равный 64 и двадцать эпох обучения (размер пакета определяет количество элементов, которые будут распространяться по сети). Большее количество эпох обучения приводило к переобучению модели и соответственно к худшим результатам на тестовой выборке.

Искусственные иммунные системы

В основе искусственных иммунных систем лежит факт обнаружения и нейтрализация чужеродных объектов (антигенов) иммунной системой. Антигены провоцируют иммунный ответ организма, который начинает вырабатывать защитные клетки иммунной системы – антитела различных видов, пока не найдется такое антитело, которое специфично связывается с антигеном и нейтрализует его. Совокупность сформированных в течение жизни антител формируют иммунитет организма [14]. Применение ИИС для решения задач стеганоанализа является сравнительно новым подходом, однако за последние несколько лет уже опубликован ряд работ в этой области (для стеганоанализа в области JPEG – работа [15]).

Для построения искусственной иммунной системы было принято решение использовать сочетание двух подходов: алгоритма отрицательного отбора, применяющегося при инициализации ИИС, и клонального отбора, применяющегося на этапе обучения ИИС [14], поскольку так мы можем создать систему, которая будет отличать чужеродные клетки от своих, а далее на этапе обучения увеличить относительный размер популяции тех антител, которые доказали свою ценность при распознавании. На данном этапе для обучения использовалось 10 поколений мутации антител. А для получения первоначального набора антител из обучающей выборки выделяются только изображения, являющиеся заполненными контейнерами, дополнительно генерируется небольшое количество антител случайным образом.

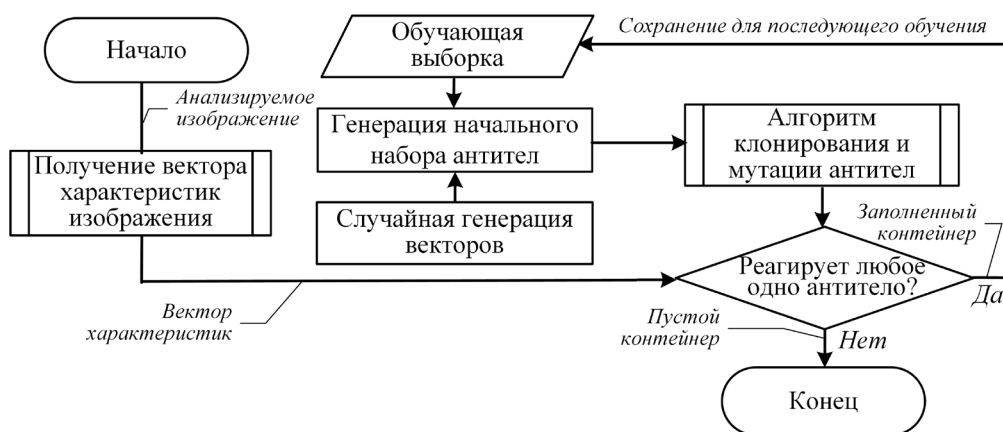


Рис. 4. Общая структурно-функциональная схема предлагаемого метода

Fig. 4. General structural and functional scheme of the proposed method

После получения итогового набора антител переходим к фазе тестирования полученной системы, основа работы которой строится на том, что если вектор характеристик анализируемого изображения попадает в окрестность хотя бы одного антитела, то такое изображение бу-

дет отнесено к множеству заполненных стеганоконтейнеров. В противном случае, к множеству пустых.

В общем виде структурно-функциональная схема предлагаемой искусственной иммунной системы для решения задачи стеганоанализа изображений представлена на рис. 4.

Анализ полученных результатов и оценка эффективности методов

Обучение и тестирование ИИС проводились на основе базы изображений IStego100K, состоящей из изображений одинакового размера 1024×1024 с различными значениями качества JPEG из диапазона от 75 до 95, как уже говорилось ранее, встраивание скрытого сообщения производилось с помощью алгоритмов стеганографии J-UNIWARD, nsF5 и UERD.

В задачах машинного обучения для оценки качества моделей и сравнения различных алгоритмов используются различные метрики. В этой статье, поскольку решается задача бинарной классификации, мы будем использовать следующие критерии качества модели: точность (ассурагу), величину ошибки первого и второго рода результатов бинарной классификации. Точность показывает количество правильно классифицированных элементов (истинно положительных и истинно отрицательных) от общего числа элементов тестовой выборки. Помимо этого, для оценки качества модели мы будем использовать время, затраченное на классификацию одного изображения на этапе тестирования, поскольку для применения метода на практике в качестве составной части средств защиты информации оно имеет критическое значение (если обработка изображения в высоконагруженной корпоративной сети будет занимать слишком большое время, то, вероятнее всего, от данного функционала системы защиты информации откажутся, ввиду нарушения доступности). Также можно сравнивать время, затраченное на обучение модели, но делать это не совсем корректно, поскольку оно зависит от используемой программной реализации (библиотек или собственноручно написанного кода) и использования параллельных вычислений в данной реализации. Также обучение модели выполняется единожды и не столь критично влияет на выбор алгоритма классификации.

Итак, время классификации одного изображения включает в себя нахождение вектора характеристик цветного изображения размером 1024×1024 пикселей описанным выше методом и непосредственно его классификацию одним из алгоритмов классификации.

В табл. 2 представлены результаты метрик точности и величины ошибки первого и второго рода для каждого алгоритма машинного обучения, представленного в данной статье отдельно для каждого алгоритма стеганографии, который использовался для построения выборки изображений.

По результатам проведенных экспериментов можно сделать следующие выводы. Метод k -ближайших соседей показывает довольно высокую эффективность, но требует наибольших вычислительных затрат при классификации изображения, следовательно, увеличивает время классификации изображения, что особенно важно при работе в реальных системах с изображениями большого размера.

Метод опорных векторов является очень точным классификатором, устойчив к шуму и наименее предрасположен к переобучению. Нейронные сети применяются в сложных областях и характеризуются большим количеством параметров и показывают очень хорошие результаты. Однако общим недостатком указанных методов является следующий факт: достижение высоких результатов требует серьезных временных затрат на обучение модели.

Метод на основе искусственных иммунных систем также показывает достаточно высокие результаты точности классификации, сравнимые с остальными развитыми методами машинного обучения (такими как нейронные сети, например), однако процесс репродукции антител на этапе обучения также является вычислительно затратным.

Таблица 2

Результаты сравнительной оценки применения предложенного метода
нахождения вектора характеристик изображения в сочетании
с различными алгоритмами машинного обучения

Table 2

Results of Comparative Evaluation of the Application of the Proposed Method of Finding
the Vector of Image Characteristics in Combination with Various Machine Learning Algorithms

	Точность, %	Величина ошибки I рода, %	Величина ошибки II рода, %	Время классификации одного изображения, мс
Алгоритм встраивания скрытого сообщения nsF5				
<i>k</i> -ближайших соседей	95,88	4,9	0,92	240,52
Деревья решений с градиентным бустингом (GBMs) + линейные модели (LightAutoML)	97,91	2,3	1,23	210,46
Метод опорных векторов	97,91	2,23	1,54	206,1
Нейронная сеть с 2 скрытыми слоями	97,97	0	3,1	206,39
Искусственные иммунные системы (AIS)	95,57	4,6	3,7	213,5
Алгоритм встраивания скрытого сообщения J-UNIWARD				
<i>k</i> -ближайших соседей	60,68	31,7	69,5	240,52
Деревья решений с градиентным бустингом (GBMs) + линейные модели (LightAutoML)	46,2	52,5	59,07	210,46
Метод опорных векторов	52,76	46,73	49,23	206,1
Нейронная сеть с 2 скрытыми слоями	25,6	90,9	8	206,39
Искусственные иммунные системы (AIS)	52,82	45,12	55,6	213,5
Алгоритм встраивания скрытого сообщения UERD				
<i>k</i> -ближайших соседей	64,65	29,7	56,89	240,52
Деревья решений с градиентным бустингом (GBMs) + линейные модели (LightAutoML)	33,33	65,6	70,6	210,46
Метод опорных векторов	65,14	36,22	29,62	206,1
Нейронная сеть с 2 скрытыми слоями	34	81,7	4,1	206,39
Искусственные иммунные системы (AIS)	63,2	38,1	31,3	213,5

Заключение

В результате данной работы были улучшены показатели точности классификации изображения путем применения метода нахождения матрицы переходных вероятностей и метода

калибровки изображения, необходимого для оценки вектора характеристик изображения, являющегося пустым стеганоконтейнером. Предложенный подход нахождения вектора характеристик изображения позволяет детектировать наличие скрытого вложения в изображениях, полученных в результате применения неадаптивных методов стеганографии (Steghide, OutGuess и nsF5) с очень высокой точностью более 95 %. Таким образом, точность обнаружения выросла более чем на 20 % по сравнению с методом, использовавшимся нами в предыдущей работе. Для заполненных контейнеров, полученных в результате встраивания сообщения одним из адаптивных методов (J-UNIWARD, UERD), показатели точности обнаружения пока находятся в пределах 50–60 %, что требует доработки в дальнейших исследованиях. Также данный метод нахождения вектора характеристик позволяет классифицировать изображение за очень короткий промежуток времени (до 250 мс), что является вполне приемлемым для использования на практике.

Кроме того, в данной работе была проведена сравнительная оценка применения различных технологий машинного обучения для решения задачи стеганоанализа статических изображений формата JPEG при использовании метода нахождения вектора характеристик на основе использования матрицы переходных вероятностей. Таким образом, метод, основанный на применении искусственных иммунных систем, показывает близкие по точности результаты детектирования в сравнении с классическими методами классификации (методом k -ближайших соседей и методом опорных векторов) и нейронными сетями. По сравнению с ними деревья решений с градиентным бустингом и линейные модели показывают худшие результаты точности детектирования адаптивных методов стеганографии в силу того, что они наиболее склонны к переобучению. Следует отметить, что метод k -ближайших соседей требует больших вычислительных затрат при классификации одного изображения, что неприменимо для реальных систем стеганоанализа. Метод опорных векторов требует определенного количества времени на обучение модели, так же как и методы на основе искусственных иммунных систем и нейросетей.

Список литературы

1. **Gulášová M., Jókay M.** Steganalysis of stegostorage library // Tatra Mountains Mathematical Publications. 2016. Vol. 67, № 1. P. 99–116. DOI 10.1515/tmmp-2016-0034
2. **Fridrich J.J., Goljan M., Hoge D.** Steganalysis of JPEG Images: Breaking the F5 Algorithm // 5th International Workshop on Information Hiding. 2002. DOI 10.1007/3-540-36415-3
3. **Hendrych J., Ličev L.** Advanced methods of detection of the steganography content // Lecture Notes in Electrical Engineering. 2020. Vol. 554. P. 484–493. DOI 10.1007/978-3-030-14907-9_47
4. **Yousfi Y., Butora J., Fridrich J., Giboulot Q.** Breaking Alaska: Color separation for steganalysis in JPEG domain // IH and MMSEC 2019 - Proceedings of the ACM Workshop on Information Hiding and Multimedia Security. 2019. P. 138–149. DOI 10.1145/3335203.3335727
5. **Saito T., Zhao Q., Naito H.** Second Level Steganalysis - Embedding Location Detection Using Machine Learning // 2019 IEEE 10th International Conference on Awareness Science and Technology, iCAST2019 - Proceedings. IEEE, 2019. P. 1–6. DOI 10.1109/ICAwST.2019.8923205.
6. **Butora J., Fridrich J.** Reverse JPEG Compatibility Attack // IEEE Transactions on Information Forensics and Security. IEEE, 2020. Vol. 15, № c. P. 1444–1454. DOI 10.1109/TIFS.2019.2940904
7. **Shniperov A. N., Prokofieva A. V.** Steganalysis Method of Static JPEG Images Based on Artificial Immune System // Automatic Control and Computer Sciences. 2020. Vol. 54, № 5. DOI 10.3103/S0146411620050077
8. **Yang Z., Wang K., Ma S., Huang Y., Kang X., Zhao X.** IStego100K: Large-scale Image Steganalysis Dataset // Digital Forensics and Watermarking. IWDW 2019. Lecture Notes in Computer Science. 2019. Vol. 12022. DOI 10.1007/978-3-030-43575-2_29

9. **Fridrich J., Pevný T., Kodovský J.** Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities // MM and Sec'07 - Proceedings of the Multimedia and Security Workshop 2007. 2007. DOI 10.1145/1288869.1288872
10. **Holub V., Fridrich J., Denemark T.** Universal distortion function for steganography in an arbitrary domain // Eurasip Journal on Information Security. 2014. Vol. 2014. DOI 10.1186/1687-417X-2014-1
11. **Guo L., Ni J., Su W., Tang C., Shi Y.** Using Statistical Image Model for JPEG Steganography: Uniform Embedding Revisited // IEEE Transactions on Information Forensics and Security. 2015. Vol. 10, № 12, DOI 10.1109/TIFS.2015.2473815
12. **Pevný T., Fridrich J.** Merging Markov and DCT features for multi-class JPEG steganalysis. 2007. P. 650503, DOI 10.1117/12.696774
13. **Vakhrushev A. et al.** LightAutoML: AutoML Solution for a Large Financial Services Ecosystem. 2021. [Электронный ресурс] URL: https://www.researchgate.net/publication/354379217_LightAutoML_AutoML_Solution_for_a_Large_Financial_Services_Ecosystem (дата обращения: 01.11.2021).
14. **Дасгупта Д.** Искусственные иммунные системы и их применение / Под ред. Романюха А. ФИЗМАТЛИТ, 2006. 344 с.
15. **Pérez J.D.J.S., Rosales M.S., Cruz-Cortés N.** Universal steganography detector based on an artificial immune system for JPEG images // Proceedings - 15th IEEE International Conference on Trust, Security and Privacy in Computing and Communications. 2017. Pp. 1896–1903, DOI 10.1109/TrustCom.2016.0290

References

1. **Gulášová M., Jókay M.** Steganalysis of stegostorage library. *Tatra Mountains Mathematical Publications*, 2016, vol. 67, no. 1, pp. 99–116. DOI 10.1515/tmmp-2016-0034
2. **Fridrich J. J., Goljan M., Hoge D.** Steganalysis of JPEG Images: Breaking the F5 Algorithm. 5th International Workshop on Information Hiding, 2002. DOI 10.1007/3-540-36415-3
3. **Hendrych J., Ličev L.** Advanced methods of detection of the steganography content. Lecture Notes in Electrical Engineering, 2020, vol. 554, pp. 484–493. DOI 10.1007/978-3-030-14907-9_47
4. **Yousfi Y., Butora J., Fridrich J., Giboulot Q.** Breaking Alaska: Color separation for steganalysis in JPEG domain. IH and MMSec 2019 - Proceedings of the ACM Workshop on Information Hiding and Multimedia Security, 2019, pp. 138–149. DOI 10.1145/3335203.3335727
5. **Saito T., Zhao Q., Naito H.** Second Level Steganalysis - Embedding Location Detection Using Machine Learning. 2019 IEEE 10th International Conference on Awareness Science and Technology, iCAST 2019 - Proceedings. IEEE, 2019. Pp. 1–6. DOI 10.1109/ICAwST.2019.8923205
6. **Butora J., Fridrich J.** Reverse JPEG Compatibility Attack. IEEE Transactions on Information Forensics and Security. IEEE, 2020, vol. 15, no. c, pp. 1444–1454. DOI 10.1109/TIFS.2019.2940904
7. **Shniperov A. N., Prokofieva A. V.** Steganalysis Method of Static JPEG Images Based on Artificial Immune System. *Automatic Control and Computer Sciences*, 2020, vol. 54, no. 5. DOI 10.3103/S0146411620050077
8. **Yang Z., Wang K., Ma S., Huang Y., Kang X., Zhao X.** IStego100K: Large-scale Image Steganalysis Dataset. Digital Forensics and Watermarking. IWDW 2019. Lecture Notes in Computer Science, 2019, vol. 12022. DOI 10.1007/978-3-030-43575-2_29
9. **Fridrich J., Pevný T., Kodovský J.** Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities. MM and Sec'07 - Proceedings of the Multimedia and Security Workshop 2007, 2007. DOI 10.1145/1288869.1288872

10. **Holub V., Fridrich J., Denemark T.** Universal distortion function for steganography in an arbitrary domain. *Eurasip Journal on Information Security*. 2014, vol. 2014. DOI 10.1186/1687-417X-2014-1
11. **Guo L., Ni J., Su W., Tang C., Shi Y.** Using Statistical Image Model for JPEG Steganography: Uniform Embedding Revisited. *IEEE Transactions on Information Forensics and Security*, 2015, vol. 10, no. 12, DOI 10.1109/TIFS.2015.2473815
12. **Pevny T., Fridrich J.** Merging Markov and DCT features for multi-class JPEG steganalysis. 2007. P. 650503, doi: 10.1117/12.696774
13. **Vakhrushev A. et al.** LightAutoML: AutoML Solution for a Large Financial Services Ecosystem [Online]. 2021. URL: https://www.researchgate.net/publication/354379217_LightAutoML_AutoML_Solution_for_a_Large_Financial_Services_Ecosystem (01.11.2021).
14. **Dasgupta D.** *Iskusstvennye immunnye sistemy i ih primeneniye*; Ed. A. Romanyuha. FIZMATLIT, 2006. 344 p. (in Russ.).
15. **Pérez J. D. J. S., Rosales M. S., Cruz-Cortés N.** Universal steganography detector based on an artificial immune system for JPEG images. *Proceedings - 15th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*. 2017. Pp. 1896–1903, DOI 10.1109/TrustCom.2016.0290

Информация об авторах

Александра Владимировна Прокофьева, аспирант кафедры прикладной математики и компьютерной безопасности

Алексей Николаевич Шниперов, кандидат технических наук, доцент кафедры прикладной математики и компьютерной безопасности

Information about the Authors

Aleksandra V. Prokofieva, postgraduate student of the Department of Applied Mathematics and Computer Security, Siberian Federal University

Alexey N. Shniperov, Candidate of Sciences in Technology, assistant Professor of the Department of Applied Mathematics and Computer Security, Siberian Federal University

Статья поступила в редакцию 27.11.2022;

одобрена после рецензирования 26.01.2023; принята к публикации 26.01.2023

The article was submitted 27.11.2022;

approved after reviewing 26.01.2023; accepted for publication 26.01.2023

Научная статья

УДК 004.942

DOI 10.25205/1818-7900-2022-20-4-76-99

Поиск подходящей архитектуры для разработки цифрового двойника гибридных энергетических систем в изолированных от сетевых энергосистем средах с использованием ТРИЗ-эволюционного подхода

Георгий Андреевич Тимофеев

Комсомольский-на-Амуре государственный университет
Комсомольск-на-Амуре, Россия

gtimjob@gmail.com, <https://orcid.org/0009-0008-7648-7878>

Аннотация

Данная работа посвящена исследованию архитектуры цифровых двойников. Разумеется, неизбежно, что технологиям создания цифровых двойников, заимствуемым из машиностроительных отраслей, присущи специфические недостатки, такие как необходимость в массивных затратных программных инструментах и высококвалифицированном персонале, хорошо заметная в немногочисленных примерах цифровых двойников энергосистем. Не хватает однозначно интерпретируемых достоверных данных в стандартных машиночитаемых форматах, адекватных математических моделей, приборного оснащения. В статье предложены подходы к преодолению этих недостатков при проектировании и эксплуатации систем цифровых двойников.

Ключевые слова

цифровые двойники, ТРИЗ-эволюция, гибридные энергетические системы, архитектура цифровых двойников, изолированные среды, промышленные энергетические системы, виртуальное и физическое прототипирование, интеграция

Для цитирования

Тимофеев Г. А. Поиск подходящей архитектуры для разработки цифрового двойника гибридных энергетических систем в изолированных от сетевых энергосистем средах с использованием ТРИЗ-эволюционного подхода // Вестник НГУ. Серия: Информационные технологии. 2022. Т. 20, № 4. С. 76–99. DOI 10.25205/1818-7900-2022-20-4-76-99

Searching for an Appropriate Architecture for the Development of Digital Twins for Hybrid Energy Systems in Isolated Environments Using the TRIZ Evolutionary Approach

Georgy A. Timofeev

Komsomolsk-on-Amur State University
Komsomolsk-on-Amur, Russian Federation

gtimjob@gmail.com, <https://orcid.org/0009-0008-7648-7878>

Abstract

This work is devoted to the study of the architecture of digital twins. Of course, it is inevitable that the technology of creating digital twins, borrowed from engineering, realizing specific turnovers, such as the need for massive costly soft-

© Тимофеев Г. А. , 2022

ware tools and highly qualified personnel, is clearly visible in the few examples of digital dual power systems. There is a lack of unambiguously interpreted data values in standard machine-readable formats, adequate mathematical models, and instrumentation. The article proposed approaches to overcome the shortcomings in the design and operation of digital twin systems.

Keywords

digital twins, TRIZ-evolution, hybrid energy systems; architecture of digital twins, isolated environments, industrial energy systems, virtual and physical prototyping, integration

For citation

Timofeev G. A. Searching for an Appropriate Architecture for the Development of Digital Twins for Hybrid Energy Systems in Isolated Environments Using the TRIZ Evolutionary Approach. *Vestnik NSU. Series: Information Technologies*, 2022, vol. 20, no. 4, pp. 76–99. (in Russ.) DOI 10.25205/1818-7900-2022-20-4-76-99

Введение

В настоящее время в России и мире всё активнее исследуются и решаются вопросы замены ископаемого топлива на возобновляемые источники энергии. В некоторых населенных пунктах, не имеющих доступа к централизованным энергосетям [1], могут быть использованы автономные гибридные энергетические системы (АГЭС). Различные условия потребления энергии в таких населенных пунктах могут создавать трудности в оценке энергообеспеченности. Также большое количество исследователей отмечают, что основные проблемы АГЭС связаны с накопителями энергии (электрические аккумуляторы, суперконденсаторы, водородные накопители, гидронакопители) [10–13, 50, 51].

На данный момент существует множество вариантов решения проблем, связанных с заменой ископаемого топлива на возобновляемые источники энергии и энергообеспеченностью автономных гибридных энергетических систем. Одним из наиболее эффективных методов является использование технологии цифровых двойников (ЦД). Эта технология позволяет детально изучить работу системы, оценить ее эффективность и сделать необходимые коррективы для оптимизации ее работы. Также использование цифрового двойника позволяет проводить моделирование ситуаций и исследование их влияния на работу системы, что помогает предотвратить непредвиденные ситуации и снизить риски.

Для систематизации знаний о цифровых двойниках предлагается использовать ТРИЗ-эволюционный подход. Развернувшаяся в последние десятилетия информационная революция обострила потребность в создании и развитии новых методов извлечения и систематизации знаний, которые обеспечивали бы возможность изучения и структурирования огромного объема информации за ограниченное время. Разрешение данного противоречия может быть достигнуто с помощью ТРИЗ-эволюционного подхода. ТРИЗ (теория решения изобретательских задач) – это область знаний, которая исследует механизмы развития искусственных систем с целью создания практических методов решения инновационных задач [3, 4]. Благодаря своей высокой эффективности и универсальности ТРИЗ завоевала международное признание и успешно применяется и развивается в множестве (118) областей человеческой деятельности, в первую очередь в промышленном производстве, науке и образовании [5, 6].

Цель данного исследования – систематизация знаний о существующих механизмах реализации цифровых двойников и определение наиболее подходящей архитектуры для реализации цифрового двойника для автономных гибридных энергетических систем, которые не имеют централизованной энергосети.

В связи с этим были решены следующие задачи:

- 1) выявить определение, что такое цифровой двойник;
- 2) выявить основные концепции архитектуры для реализации цифровых двойников;
- 3) определить критерии оптимальности для достижения оптимальных конфигураций использования источников энергии.

- 4) определить главные противоречия, которые стали «движущими силами» каждой новой архитектуры;
- 5) установить приемы разрешения этих противоречия на основе ТРИЗ-инструментов;
- 6) сформировать ТРИЗ-эволюционная карту для цифровых двойников;
- 7) определить вариант цифровых двойников для гибридных энергетических систем, предназначенных для поселений Крайнего Севера, изолированных от сетевых энергосистем.

1. Определение и основные концепции

Цифровой двойник – система, состоящая из цифровой модели изделия и двусторонних информационных связей с изделием (при наличии изделия) и (или) его составными частями. Объединяя различные математические и компьютерные модели в единую систему, можно получить новую сущность – цифровой двойник, которая позволяет всесторонне описать изделие и системно подойти к разработке, производству и эксплуатации изделий [7]. По структуре цифровой двойник – это сложный комплекс взаимосвязанных компьютерных моделей, способных представить объект-оригинал, его состояние и поведение при разных параметрах объекта или среды. В литературе по цифровым двойникам не хватает однозначно интерпретируемых достоверных данных в стандартных машиночитаемых форматах, адекватных математических моделей, приборного оснащения для автономных гибридных энергетических систем, лишенных обеспечения центральной системы энергосети. В рамках этого раздела будут рассмотрены различные варианты, которые упоминаются в научной литературе.

В рамках данного исследования был сформирован поисковый запрос digital twin Energy system и «цифровой двойник энергетических систем» для наукометрических баз данных, что привело к формированию реестра из 269 статей. Далее, для формирования морфологической таблицы, реестр был отфильтрован в соответствии с тематикой исследования. На основе ключевых слов, аннотаций и содержания статей были выявлены факторы «Классификация источника энергии», «Программная платформа» и «Архитектура цифрового двойника». Для каждого фактора в статьях были идентифицированы варианты и определены их упоминания в источниках. Результатом данного этапа является морфологическая таблица (см. табл.).

Изучая базовые концепции цифровых двойников, можно выделить архитектуру, описанную в научной статье [16]. В основной базовой концепции ЦД были выделены три основных аспекта: физическое пространство, виртуальное пространство и их связь для обмена данными и информацией. Эта концепция известна в промышленной сфере как киберфизическая система (CPS) или, более конкретно, киберфизическая производственная система (CPPS). CPS описывает как автономные, так и кооперативные элементы и подсистемы на всех уровнях производства, которые могут взаимодействовать друг с другом в зависимости от ситуации. Целью CPS является создание элементов, которые могут получать и обрабатывать данные, что позволяет им самостоятельно выполнять определенные задачи и взаимодействовать с людьми. CPS можно охарактеризовать физическим активом и его кибернетическим аналогом, что означает, что цифровой двойник (DT) может быть рассмотрен только как цифровая модель внутри CPS

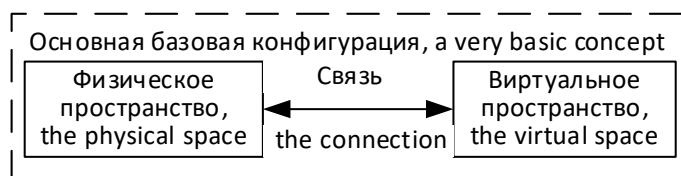


Рис. 1. Основная базовая конфигурация ЦД

Fig. 1. Basic concept

Морфологическая таблица

Morphological Table

Факторы	Var. 01	Var. 02	Var. 03	Var. 04	Var. 05	Var. 06
Классификация источника энергии	Солнечная энергия	Ветровая энергия	Ядерная энергия	Водородная энергия	—	—
	[17, 22, 23, 31, 38, 43] всего 6	[18, 32, 33, 35, 36, 37] всего 6	[30, 42] всего 2	[40, 41] всего 2	—	—
Программная платформа	RT MATLAB/Simulink	Amazon Web Services	Predix	Prod-romos	CFD ANSYS Fluent	Plant Simulation
	[15, 19, 20, 25, 29, 34, 57] всего 7	[20] всего 1	[17] всего 1	[19] всего 1	[21] всего 1	[25] всего 1
Архитектура цифрового двойника	5D-DT, GDTA, Intelligent Digital Twin	Эталонная структура для цифровых двойников	COGNITWIN, IoT	—	—	—
	[15] всего 1	[16, 24, 26, 27, 28] всего 5	[17, 18, 39, 44, 45, 46, 47, 48] всего 8	—	—	—

[16]. Понимание этой концепции важно для определения наиболее подходящей архитектуры для реализации цифрового двойника для автономных гибридных энергетических систем, оставшихся без центральной системы энергосетей. Требования общества: увеличить, точность и надежность данных.

2. Первый этап ТРИЗ

Идеальный конечный результат: цифровая копия физического объекта или процесса с различными подсистемами, самовзаимодействующими друг с другом с целью получения точных и надежных данных от физического объекта.

Техническое противоречие №1: при увеличении надежности данных недопустимо увеличиваются потери информации.

Главный параметр №1: надежность данных.

Главный параметр №2: потери информации.

В рамках **первого этапа эволюции** был применен принцип **предварительного действия**. Это позволило создать процесс преобразования данных в информацию между виртуальным и физическим объектом, а также другие необходимые уровни для функционирования системы. На рис. 2 представлена схема данного этапа ТРИ-эволюции.

Концепция киберфизической системы (CPS) или киберфизической производственной системы (CPPS) широко известна в промышленной сфере. Для реализации CPS была предложена пятиуровневая архитектура, известная как 5C architecture, состоящая из следующих уровней:

- 1) интеллектуального соединения;
- 2) преобразования данных в информацию;

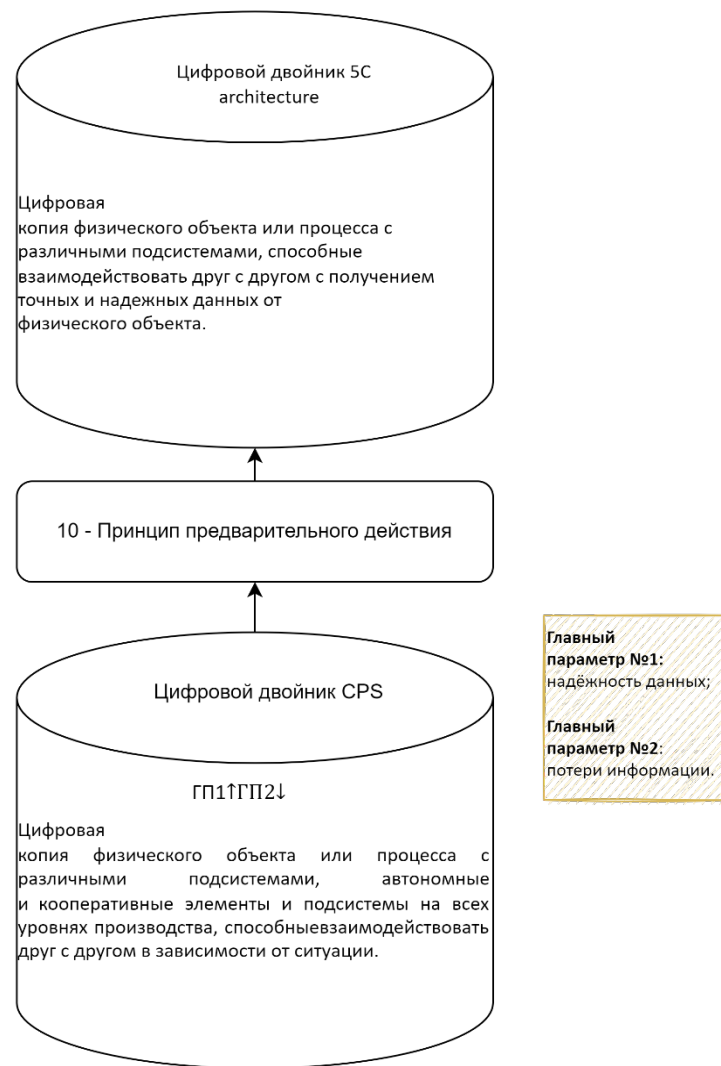


Рис. 2. ТРИЗ-эволюция модели CPS в 5C architecture
Fig. 2. TRIZ evolution of the CPS model into 5C architecture

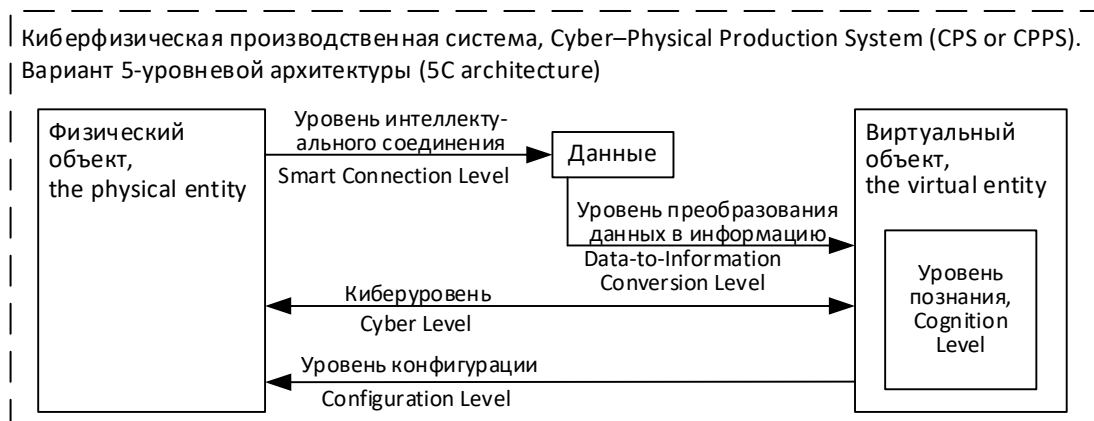


Рис. 3. Пятиуровневая архитектура для CPS/ CPPS (5C architecture)
Fig. 3. Five-level architecture for CPS/CPPS (5C architecture)

- 3) киберуровень;
- 4) познания;
- 5) конфигурации. [15]

Требование общества: разработка архитектуры, которая позволит уменьшить потерю информации и добиться оптимального соответствия между виртуальным и физическим объектом. В рамках данного этапа была разработана концепция, которая описывает архитектуру, имеющую возможность получать и обрабатывать данные, позволяющую самостоятельно выполнять определенные задачи и взаимодействовать. В соответствии с требованиями общества к минимизации потери информации и достижению оптимального соответствия виртуального и физического объекта на этом этапе были разработаны архитектуры, которые позволяют системам получать и обрабатывать данные самостоятельно, но они не рассматривают взаимодействие со сложными системами и требуют улучшения.

3. Второй этап ТРИЗ

Идеальный конечный результат: цифровая копия физического объекта или процесса с различными подсистемами, с платформой виртуальных сущностей, состоящих из множества моделей виртуальных сущностей, организующая информацию для зеркального отображения определенного аспекта физической сущности.

Техническое противоречие №2: при уменьшении потери информации недопустимо увеличиваются потери времени.

Главный параметр №2: потеря информации.

Главный параметр №3: потери времени.

Приемы разрешения:

- 1) ПРИНЦИП ПОСРЕДНИКА;
- 2) ПРИНЦИП КОПИРОВАНИЯ.

В рамках второго этапа эволюции **принцип посредника** был реализован путем внедрения промежуточных объектов между **виртуальным** и **физическим** объектом. Это позволило реализовать передачу действия между ними. В этом контексте были внедрены дополнительные посредники, такие как платформа управления данными и платформа услуг. Схема этого этапа ТРИЗ-эволюции представлена на рис. 4.

Эталонная структура для цифровых двойников является ключевым компонентом в контексте архитектуры 5C CPS. Для предлагаемой системы были идентифицированы основные строительные блоки ЦД, включая их свойства (структуру и взаимосвязь). Компоненты архитектуры на верхнем уровне включают платформы:

- 1) физических сущностей;
- 2) виртуальных сущностей;
- 3) управления данными;
- 4) услуг.

«Платформа виртуальных сущностей» представляет собой коллекцию моделей виртуальных сущностей, включающую информацию для отображения специфических аспектов физической сущности. Она отвечает за создание и обслуживание «семантических моделей», таких как геометрические, физические, поведенческие, модели правил и модели процессов, физической сущности, для создания ее виртуального представления.

«Платформа управления данными» осуществляет функции сбора, управления и хранения данных, включая сбор, передачу, интеграцию, обработку, очистку, анализ и извлечение информации. «Платформа услуг» состоит из моделей услуг и слоя управления услугами, которые организуют услуги для конкретных приложений [16]. На рис. 5 представлена схема этой архитектуры.

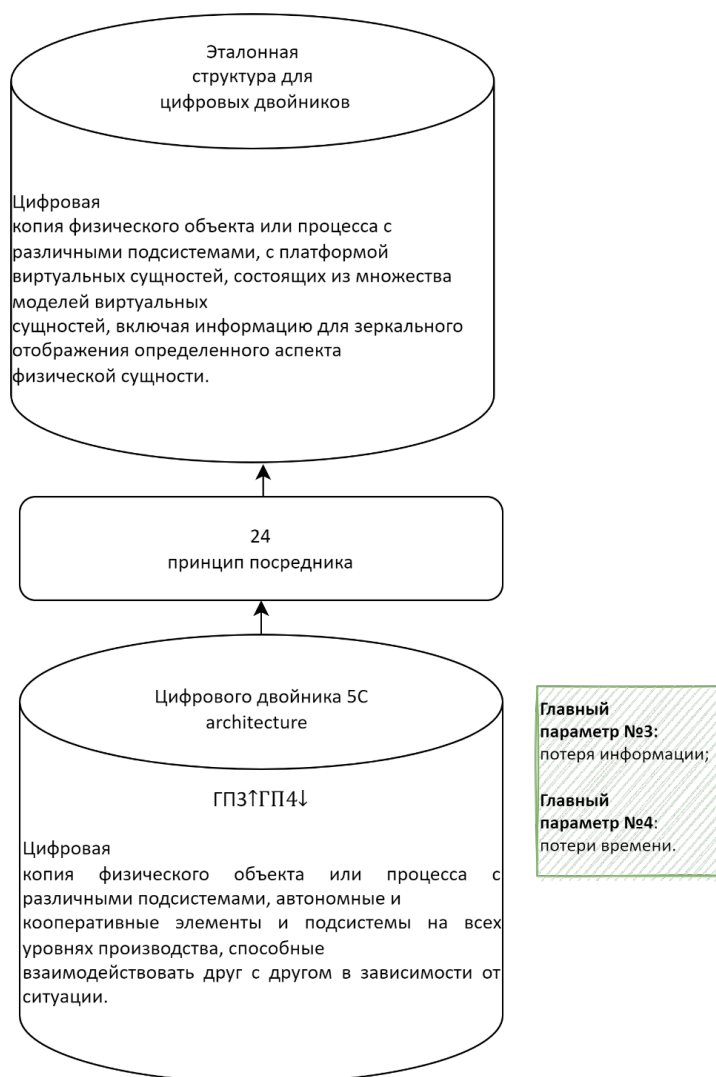


Рис. 4. ТРИЗ-эволюция модели из 5C architecture в «Эталонная структура для цифровых двойников»

Fig. 4. TRIZ model evolution from 5C architecture to “Reference structure for digital twins”

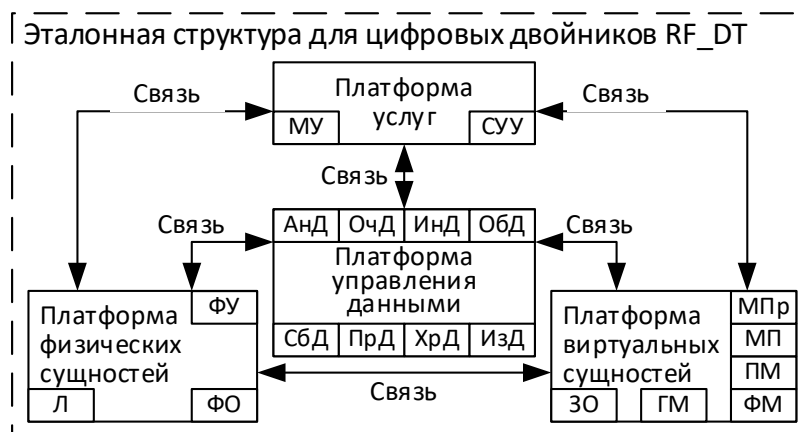


Рис. 5. Архитектура пятимерного цифрового двойника (5D-DT)

Fig. 5. Architecture of a 5D digital twin (5D-DT)

На рис. 6 представлена схема преобразования архитектуры 5C CPS в модель цифрового двойника, а также отображена взаимосвязь между физическим объектом и его цифровой моделью. Эта модель представляет собой альтернативное видение этапа ТРИЗ-эволюции.

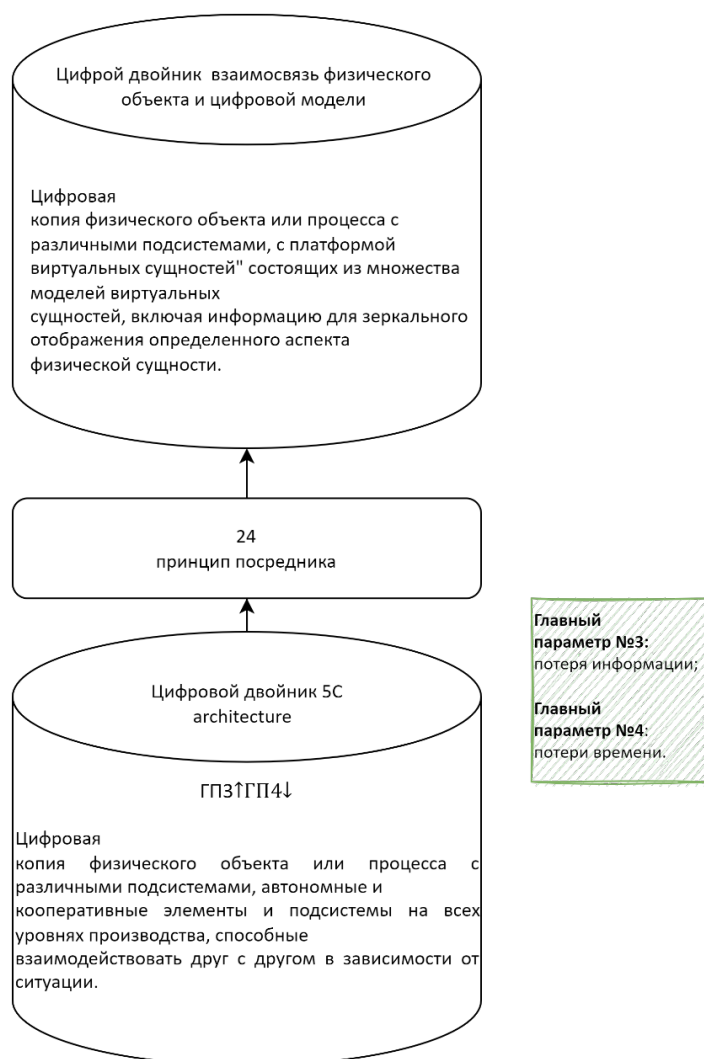


Рис. 6. ТРИЗ-эволюция модели цифрового двойника и взаимосвязь физического объекта и цифровой модели
Fig. 6. TRIZ evolution of the digital twin model and the relationship between a physical object and a digital model

Следующая модель является альтернативным видением данного этапа ТРИЗ-эволюции.

Модель выражается через пять вспомогательных компонентов (рис. 7) [26]:

- 1) датчики и исполнительные механизмы из физического мира;
- 2) интеграция;
- 3) данные;
- 4) аналитика;
- 5) постоянно обновляемое приложение цифрового двойника.

Датчики, распределенные по всему производственному процессу, создают сигналы, которые позволяют ЦД собирать оперативные данные и данные об окружающей среде, касающиеся физического процесса в реальном мире. Реальные эксплуатационные данные и данные об окружающей среде, полученные от датчиков, объединяются с данными компании, такими как спецификация (BOM), корпоративные системы и проектные спецификации.

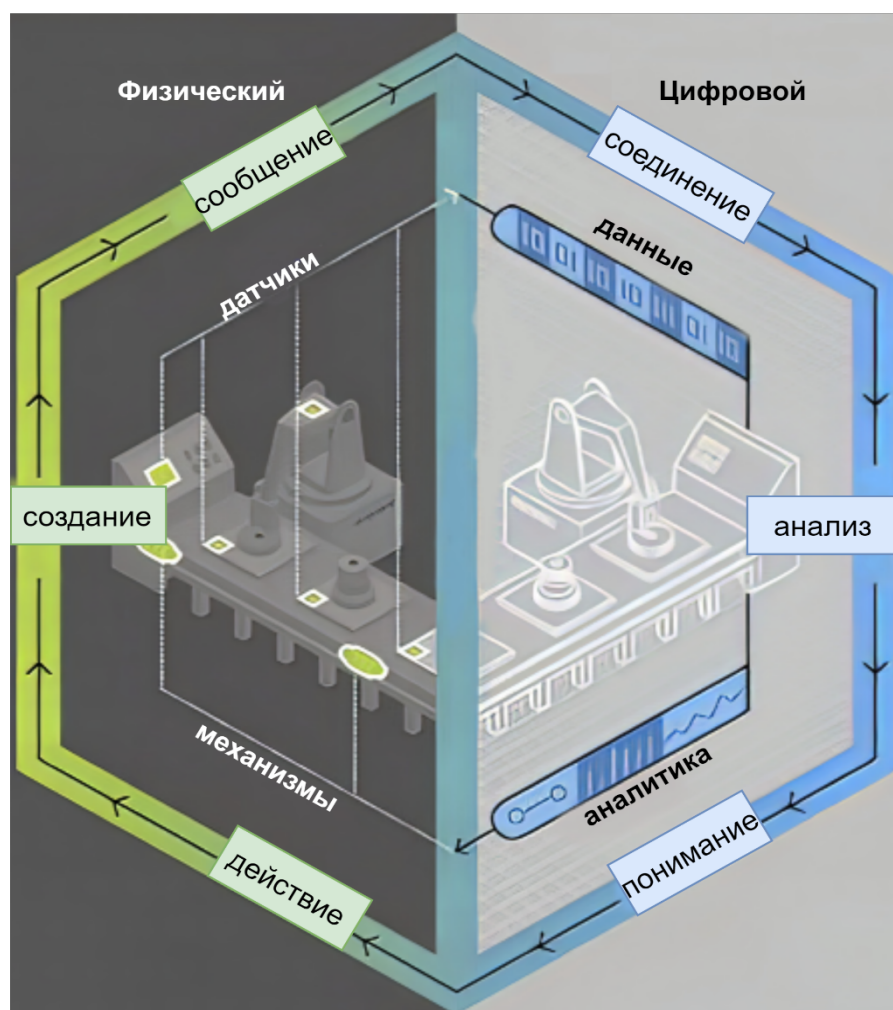


Рис. 7. Модель цифрового двойника и взаимосвязь физического объекта и цифровой модели

Fig. 7. Digital twin model and relationship between physical object and digital model

На рис. 8 представлена схема преобразования модели из архитектуры 5С в модель цифрового двойника системы на основе данных, полученных из системы диспетчерского управления предприятием. Данная модель представляет собой альтернативный вариант представления этапа ТРИЗ-эволюции.

Предложенная методология предназначена для создания цифровых дубликатов холодильной системы на основе данных, которые хранятся в системе диспетчерского управления предприятием. Методология использует теорию нечетких множеств, и алгоритм автоматического извлечения правил дает возможность определить зависимости между измерениями холодильной системы и ее поведением в реальном мире. Это позволяет создать более точную модель цифрового двойника системы водяного охлаждения [24], как показано на рис. 9. Архитектура получения цифровых двойников системы водяного охлаждения, основанная на использовании теории нечетких множеств и алгоритма автоматического извлечения правил, является перспективным подходом для повышения эффективности и оптимизации реальных физических систем.

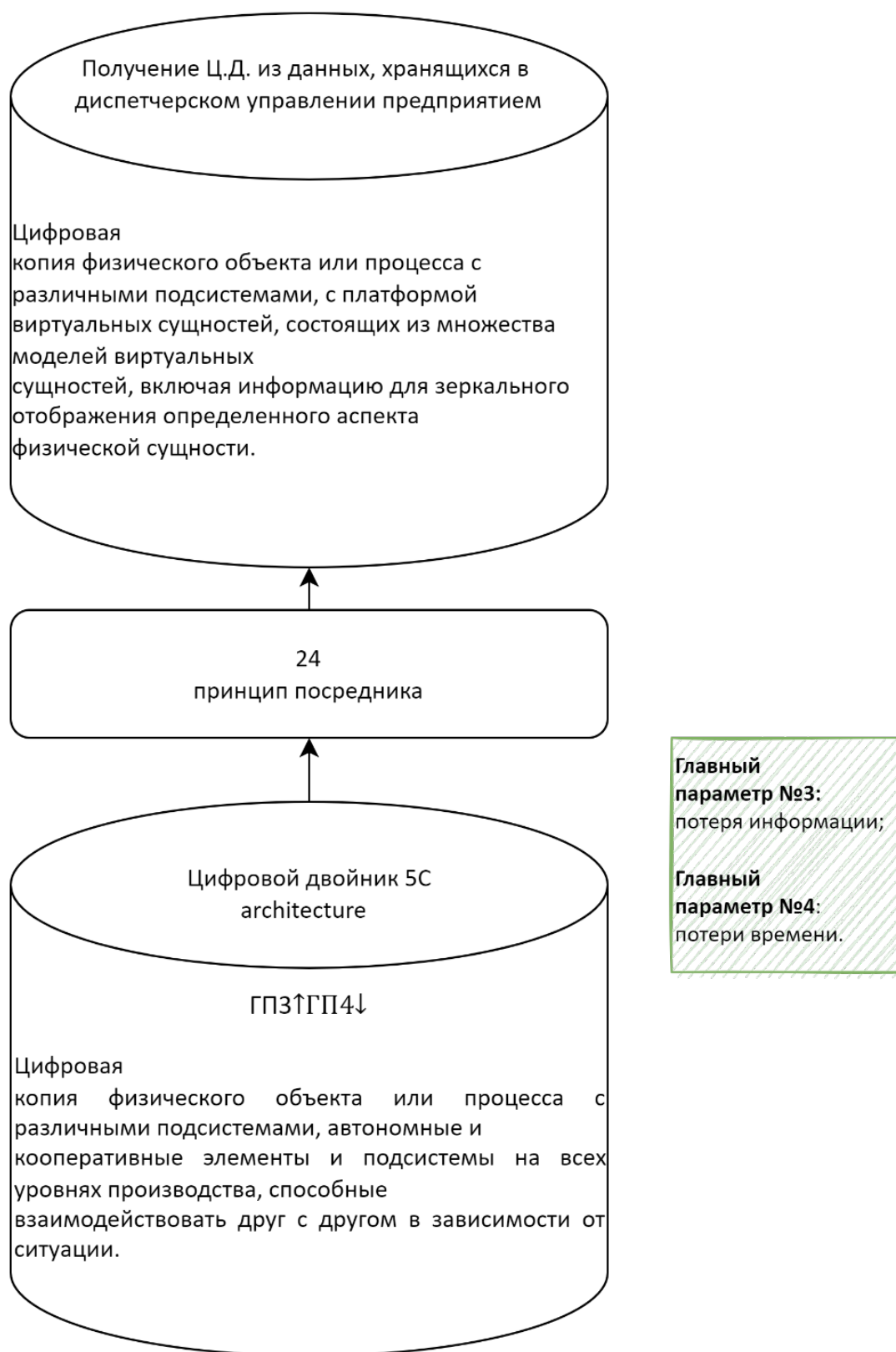


Рис. 8. ТРИЗ-эволюция модели в модель получения цифровых двойников системы из данных

Fig. 8. TRIZ evolution into a model of receiving system's digital twins from data

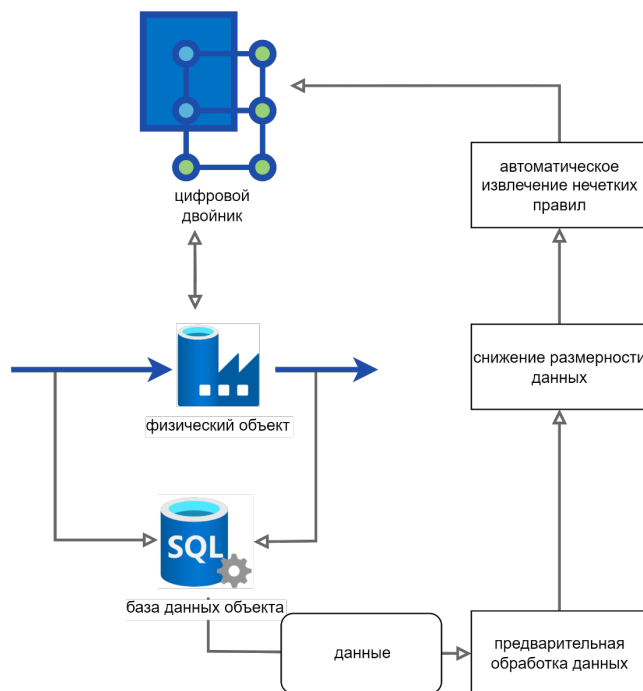


Рис. 9. Модели получения цифровых двойников системы из данных
Fig. 9. Models of receiving digital twin of a system from data

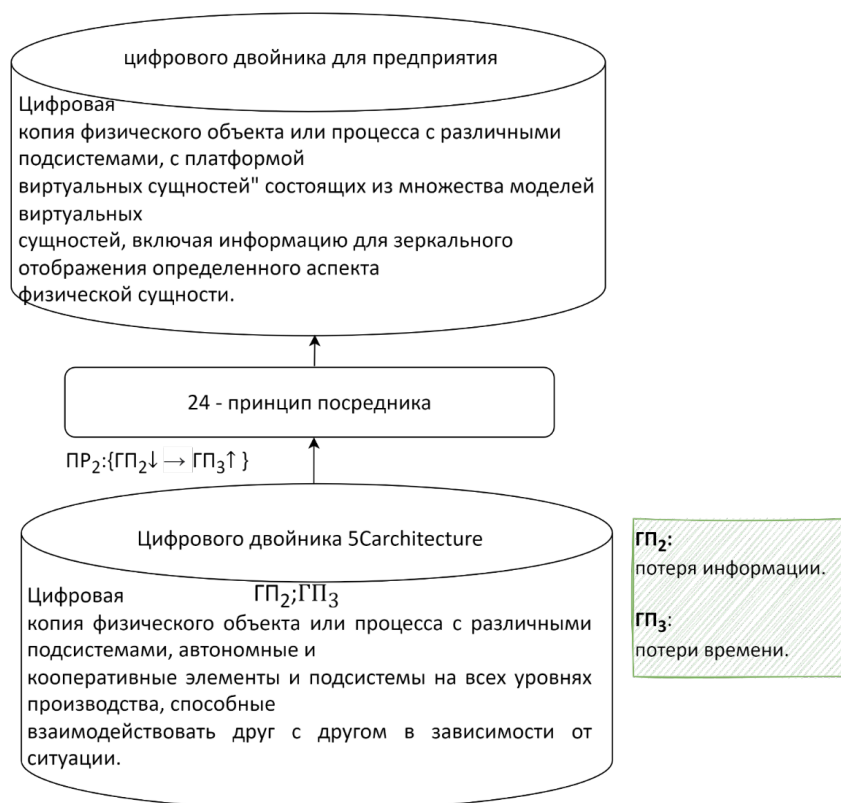


Рис. 10. ТРИЗ-эволюция из 5C architecture в модель цифрового двойника предприятия холодильной системы
Fig. 10. TRIZ evolution from 5C architecture to a digital twin model of the refrigeration system enterprise

На рис. 10 представлена схема ТРИЗ-эволюции из 5C architecture в модель цифрового двойника предприятия холодильной системы. Данная модель представляет собой альтернативное представление этапа ТРИЗ-эволюции.

На рис.11 представлена схема ТРИЗ-эволюции модели из 5C architecture в модель цифрового двойника предприятия холодильной системы. В рамках предлагаемой методологии цифровой двойник предприятия может быть создан для охвата всех процессов организации, включая закупку товаров для производства, логистику и услуги. Он включает в себя все данные о технологических процессах, эффективности методов управления качеством и планировании последующих операций на производстве. Авторы считают, что технология цифрового двойника предоставляет огромное количество возможностей для моделирования самых разных ситуаций, которые могут возникнуть на производстве, вплоть до критических, не подвергая риску само производство [28].

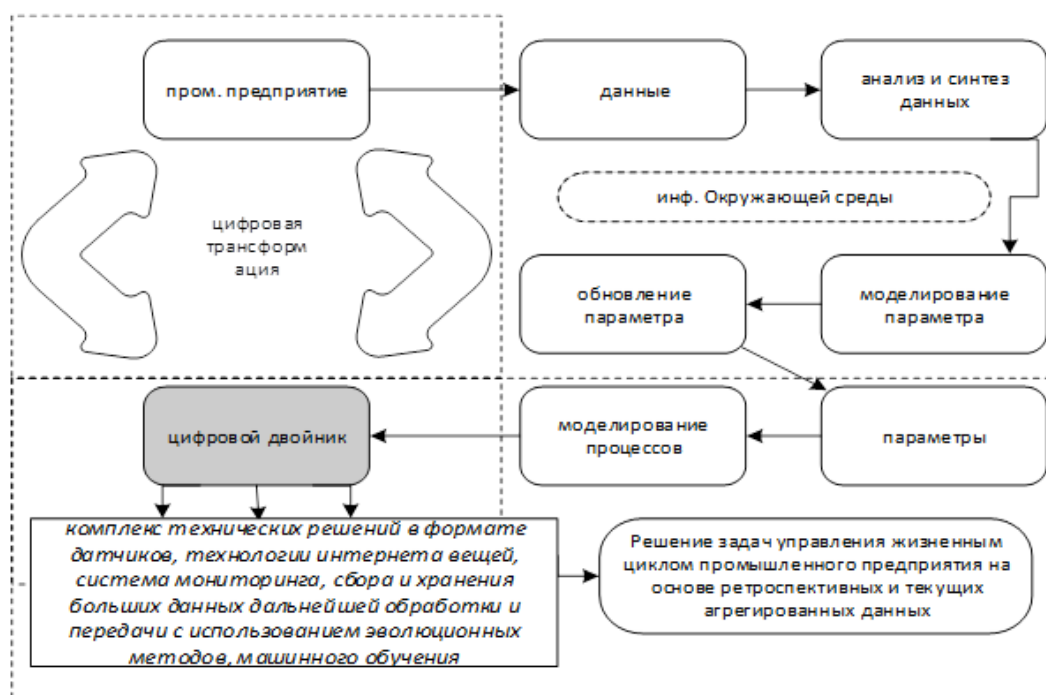


Рис. 11. Модель цифрового двойника предприятия

Fig. 11. Enterprise digital twin model

На данном этапе концепции описывают возможность интеграции различных функциональных компонентов в модель цифрового двойника [16; 24]. Для описания математической модели есть примеры, где используется теория нечетких множеств и алгоритм автоматического извлечения правил [24]. В соответствии с концепцией цифрового двойника (ЦД) может быть разработана модель, которая интегрирует несколько функциональных компонентов в один цифровой образ. В частности, ЦД предприятия может охватывать все процессы, которые происходят в организации, от закупки товаров для производства до логистики и услуг [28]. Данные архитектуры цифрового двойника могут быть расширены и применены для гибридных энергетических систем, которые предназначены для изолированных от сетевых энергосистем поселений Крайнего Севера. Это подтверждает гибкость концепции ЦД и ее возможность адаптации к различным системам, в том числе и к такой крупной системе, как система предприятия [28].

Требования общества: обеспечить архитектуру ЦД оболочкой для визуализации.

3. Третий этап ТРИЗ

Идеальный конечный результат: цифровая копия физического объекта или процесса, которая инкорпорирует различные подсистемы, включающие платформу виртуальных сущностей, которые могут взаимодействовать друг с другом, а также включающие визуализацию модели.

Техническое противоречие №3: увеличивая удобство эксплуатации, недопустимо ухудшается адаптация, универсальность.

Главный параметр №4: удобство эксплуатации.

Главный параметр №5: адаптация, универсальность.

Приемы разрешения: 24. ПРИНЦИП ПОСРЕДНИКА;

26. ПРИНЦИП КОПИРОВАНИЯ.

Принцип дробления был применен на 3 этапе эволюции: между виртуальным и физическим объектом были интегрированы промежуточные объекты, которые выполняют дополнительные функции, чтобы увеличить удобство эксплуатации. На рис. 12 представлена схема ТРИЗ-эволюции модели 2 этапа в модель COGNITWIN. Эта модель представляет собой альтернативное представление этапа ТРИЗ-эволюции.

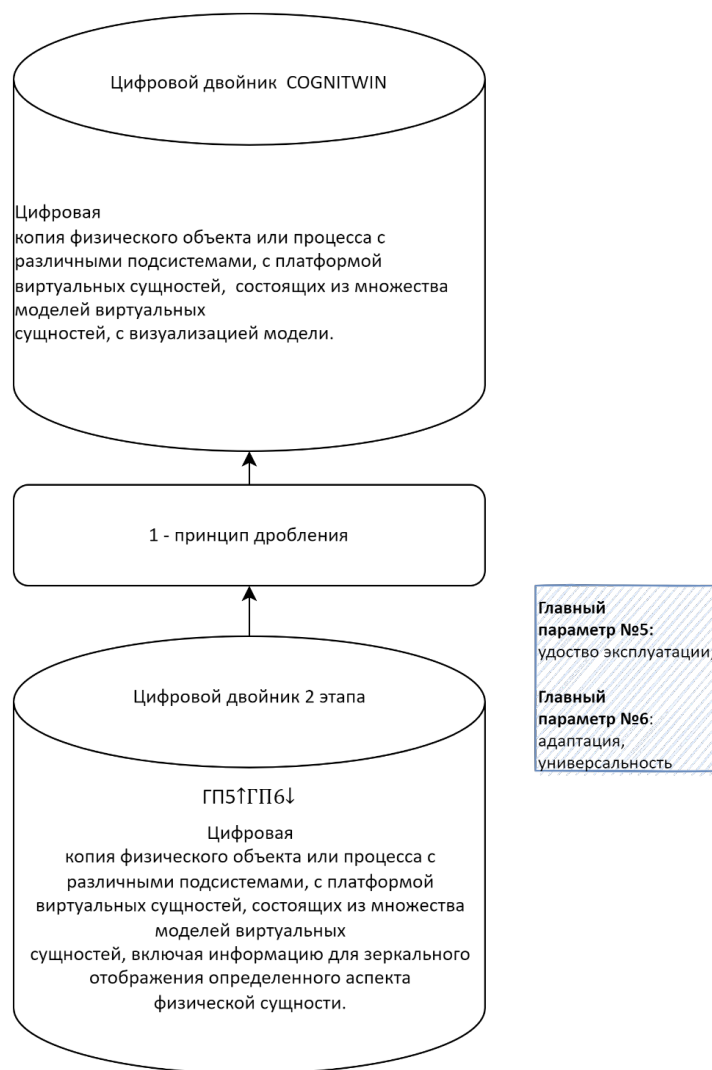


Рис. 12. ТРИЗ-эволюция модели 2-го этапа в модель COGNITWIN
Fig. 12. TRIZ evolution of the 2nd stage model into the COGNITWIN model

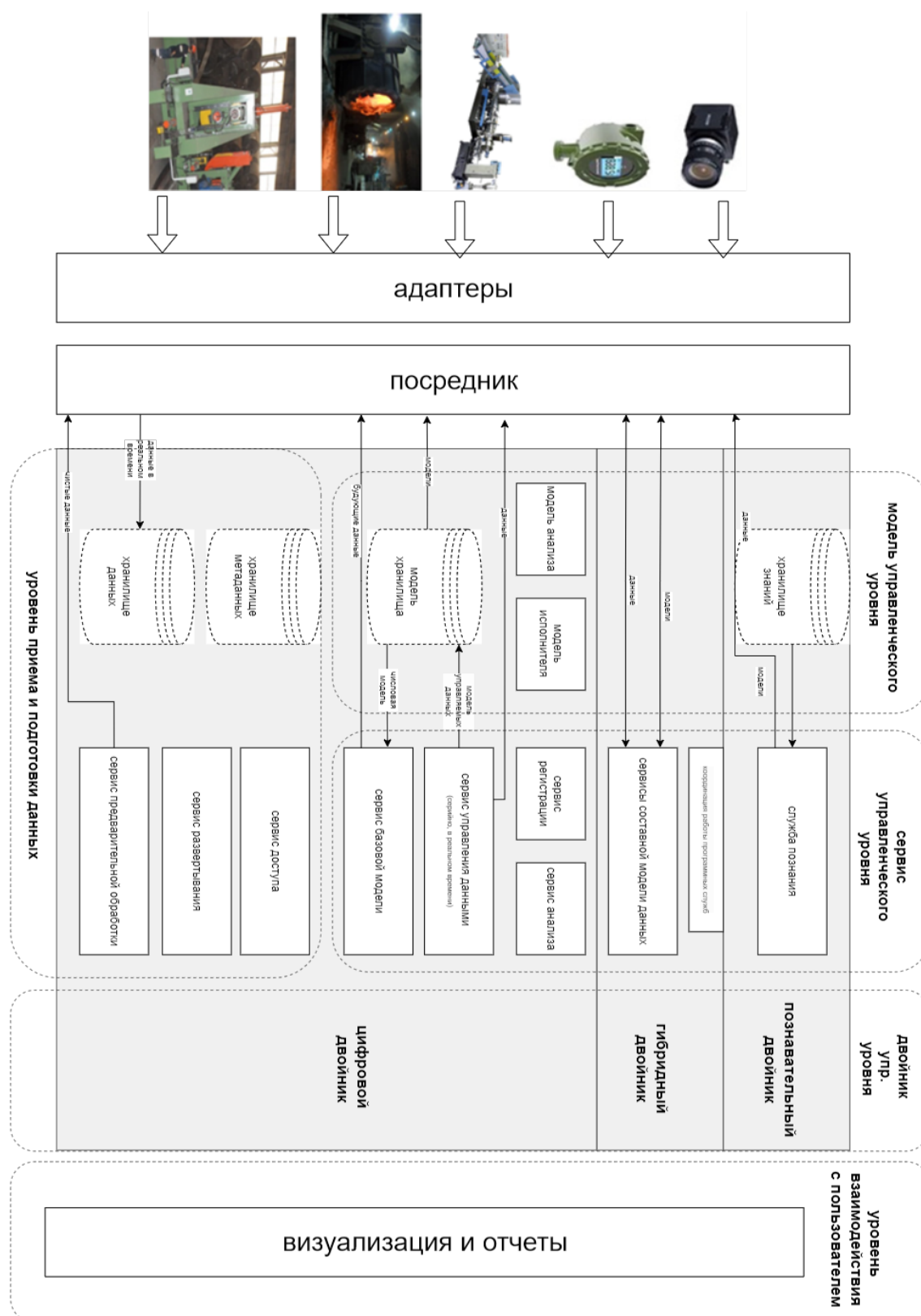


Рис. 13. Схема модели Cognitive Twin Toolbox
Fig. 13. Diagram of the Cognitive Twin Toolbox Model

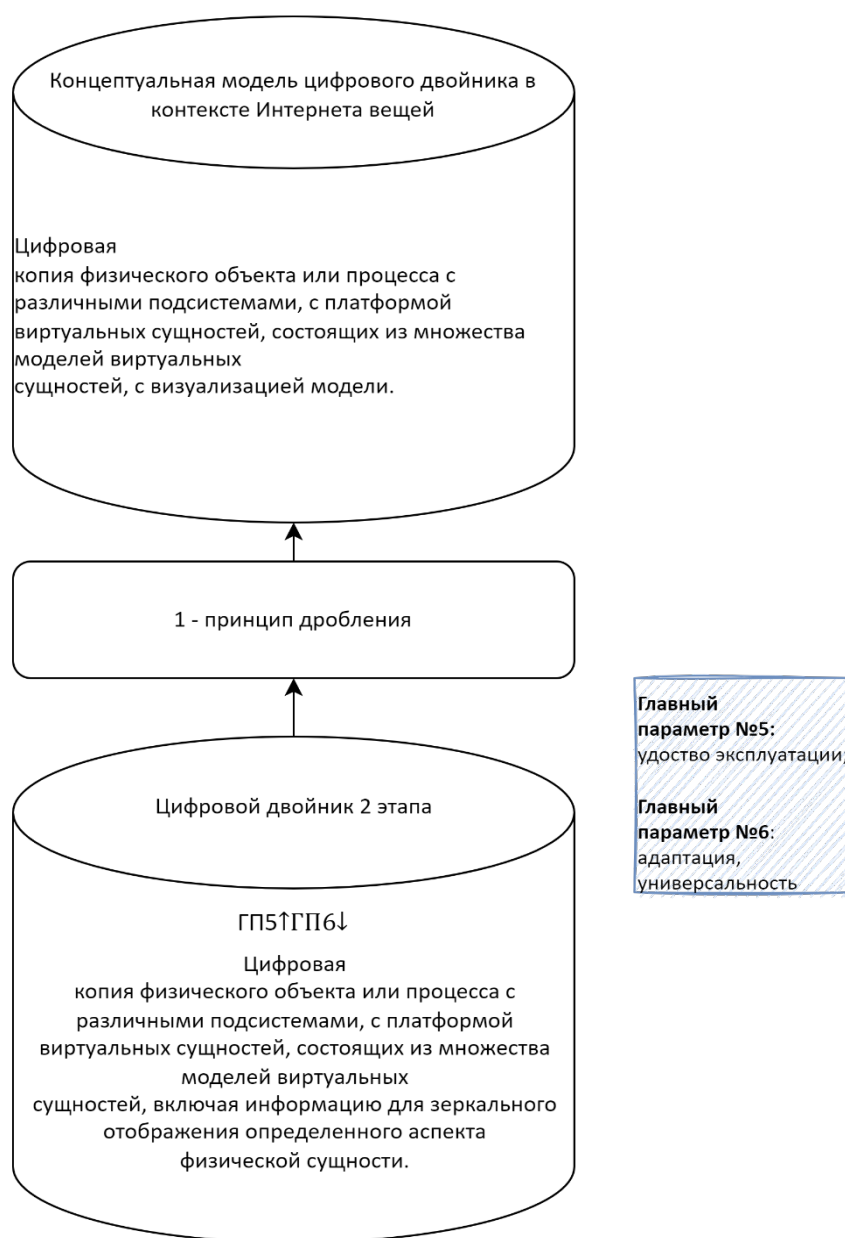


Рис. 14. ТРИЗ-эволюция модели 2-го этапа в модель «Концептуальная модель цифрового двойника в контексте интернета вещей»

Fig. 14. TRIZ evolution of the 2nd stage model into the model “Digital Twin Model in the Context of the Internet of Things”

Еще одна модель Cognitive Twin Toolbox (COGNITWIN) с особым акцентом на **обрабатывающую промышленность** (рис. 13) [46]. Имеет три уровня двойников:

- 1) цифровой двойник;
- 2) гибридный двойник;
- 3) когнитивный двойник.

Когнитивный двойник предлагает **пять уровней**:

- 1) управления моделями;
- 2) сбора и подготовки данных;

- 3) управления сервисами;
- 4) управления двойниками;
- 5) взаимодействия с пользователем.

На рис. 14 представлена диаграмма ТРИЗ-эволюции модели 2 этапа в концептуальную модель цифрового двойника в контексте интернета вещей. Эта модель представляет собой альтернативное представление данного этапа ТРИЗ-эволюции.

В качестве альтернативного видения данного этапа ТРИЗ-эволюции предлагается использование «концептуальной модели цифрового двойника». Данная модель описывает структуру на пяти уровнях, которые состоят из различных компонентов, обеспечивающих эффективное функционирование системы:

- 1) физическое пространство;
- 2) коммуникационная сеть;
- 3) виртуальное пространство;
- 4) аналитика и визуализация данных;
- 5) приложение.

В концепции описывается система соединения различных предметов и устройств в Интернете, позволяющая им обмениваться данными и управляться удаленно. Может включать в себя различные виды устройств – от смартфонов и компьютеров до датчиков и других маленьких устройств, которые могут быть встроены в окружающую среду. Задачи различных слоев очень похожи на уже упомянутые слои в других фреймворках. Описаны два концептуальных варианта использования в области автомобилестроения и умного здравоохранения, но не представлена реальная реализация предложенной модели (рис. 15).

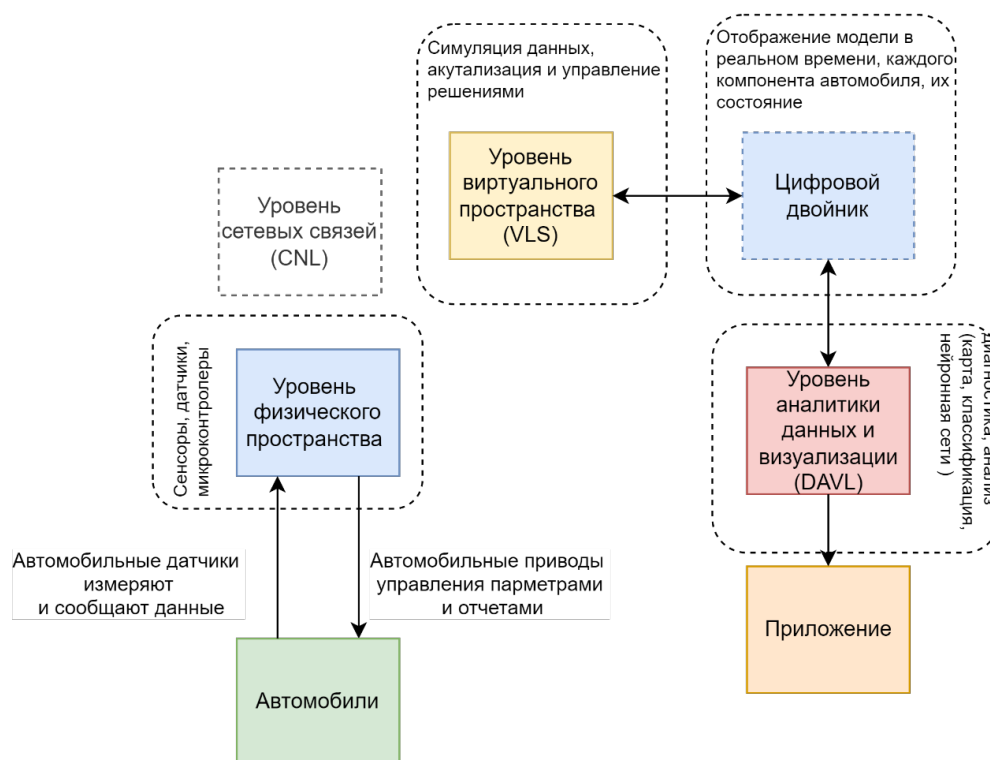


Рис. 15. Схема концептуальной модели ЦД
Fig. 15. Scheme of the conceptual model of CD

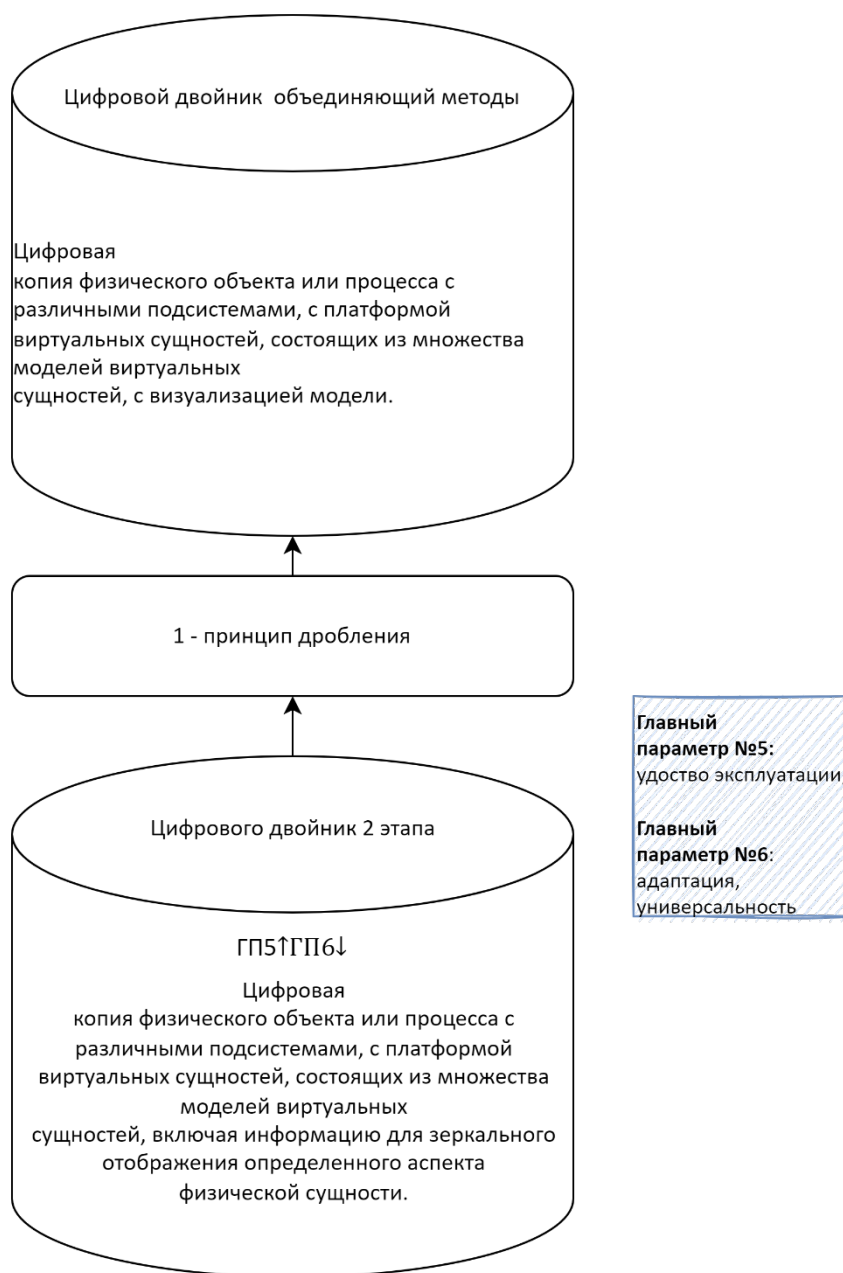


Рис. 16. ТРИЗ-эволюция модели 2 этапа в модель ЦД, объединяющего методы
 Fig. 16. TRIZ evolution of the 2nd stage model into the model of CD that combines methods

На рис. 16 представлена схема ТРИЗ-эволюции, отображающая процесс перехода от модели 2 этапа к концептуальной модели цифрового двойника в контексте интернета вещей. Эта модель представляет собой альтернативное представление данного этапа ТРИЗ-эволюции. Предложенная концепция создания цифровых двойников основана на использовании комплексного метода, который объединяет технологии имитационного моделирования, модули оптимизации (например, генетические алгоритмы), базу данных и подсистему анализа многомерных данных (например, онлайн-аналитическую обработку). Особенностью предложенного метода является обеспечение непрерывной интеграции рассматриваемых подсистем с поддержкой функциони-

рования всех программных модулей в режиме реального времени (рис. 17) [39]. Предлагаются следующие взаимодействующие подсистемы:

- 1) подсистема имитационного моделирования, предназначенная для расчета значений множественных характеристик предприятия при заданных сценарных условиях;
- 2) модуль оптимизации (генетический алгоритм), который агрегирует через целевые функции с имитационной моделью производственной компании и обеспечивает возможность поиска наилучших (рациональных, субоптимальных) решений при существующих ограничениях;
- 3) база данных, играющая ключевую роль в сборе и обработке актуальных данных по предприятию, а также исходных данных для имитационных моделей. Она также служит для хранения результатов имитационного моделирования;
- 4) подсистема визуализации и управления имитационными моделями, доступ к различным функциям имитационных моделей, например, вызов оптимизационных экспериментов, сохранение результатов расчетов;
- 5) подсистема оперативной аналитической обработки (OLAP), позволяющая анализировать результаты моделирования и оптимизации с детализацией соответствующих агрегированных данных;
- 6) подсистема интеллектуального анализа данных, обеспечивающая анализ взаимосвязей между ключевыми характеристиками информационной модели производства, а также обновление значений всех влияющих факторов и коэффициентов значимости. Результаты анализа сохраняются в базе данных системы.

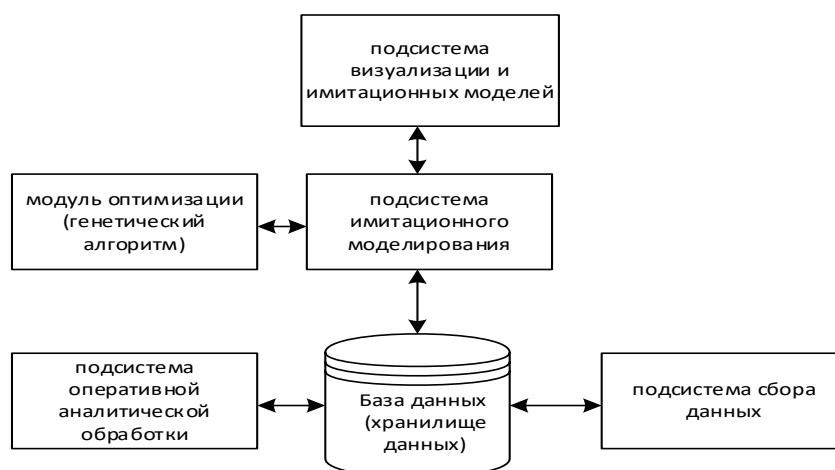


Рис. 17. Схема модели ЦД на использовании комплексного метода

Fig. 17. Scheme of the CD model based on the complex method

В данном исследовании предлагается архитектура устойчивого интеллектуального производства, управляемого цифровыми двойниками, которая представлена на рис. 18 [47]. Эта структура состоит из:

- 1) базовой платформы;
- 2) производственного оборудования;
- 3) производственной системы;
- 4) производственного сервиса.

Данные базовой платформы поступают из уровня оборудования платформы, который включает в себя оборудование, агрегат, производственную линию и производственный цех. После того как платформа получает данные с уровня устройств, она объединяет облачные вычисления, искусственный интеллект, интернет вещей и другие технологии. Затем всесторон-

не рассматривает экологические, экономические и социальные факторы и объединяет людей, оборудование и технологии для предоставления данных для виртуального и физического прототипирования. Устойчивое интеллектуальное производственное оборудование сопоставляет с виртуальными и физическими прототипами.

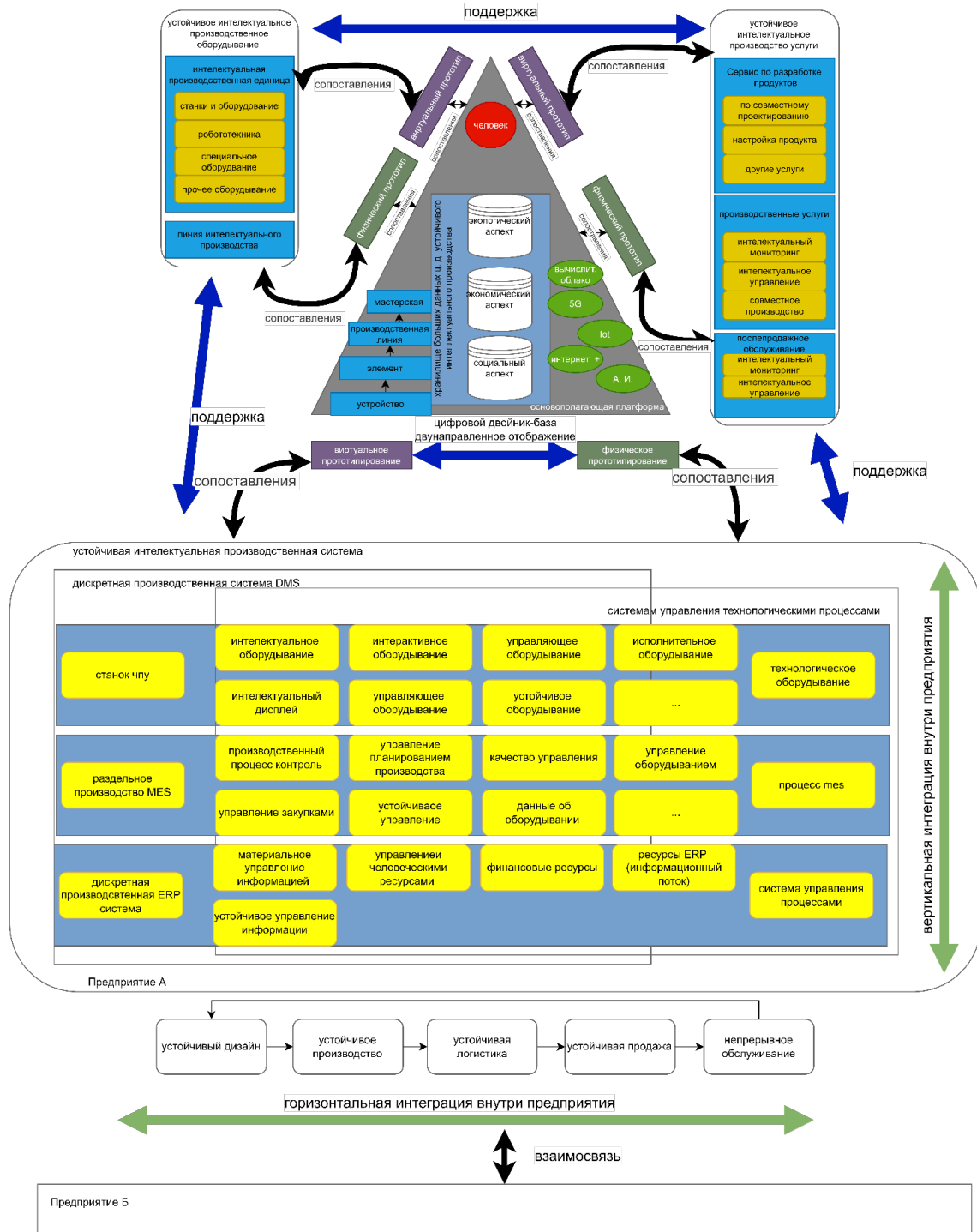


Рис. 18. Архитектура структуры устойчивого интеллектуального производства, управляемого цифровыми двойниками

Fig. 18. Architecture of the structure of sustainable intelligent production managed by digital twins

Выбрав фактор «архитектура цифрового двойника», как наиболее значимый, была составлена ТРИЗ-эволюция цифровых двойников (рис. 19).

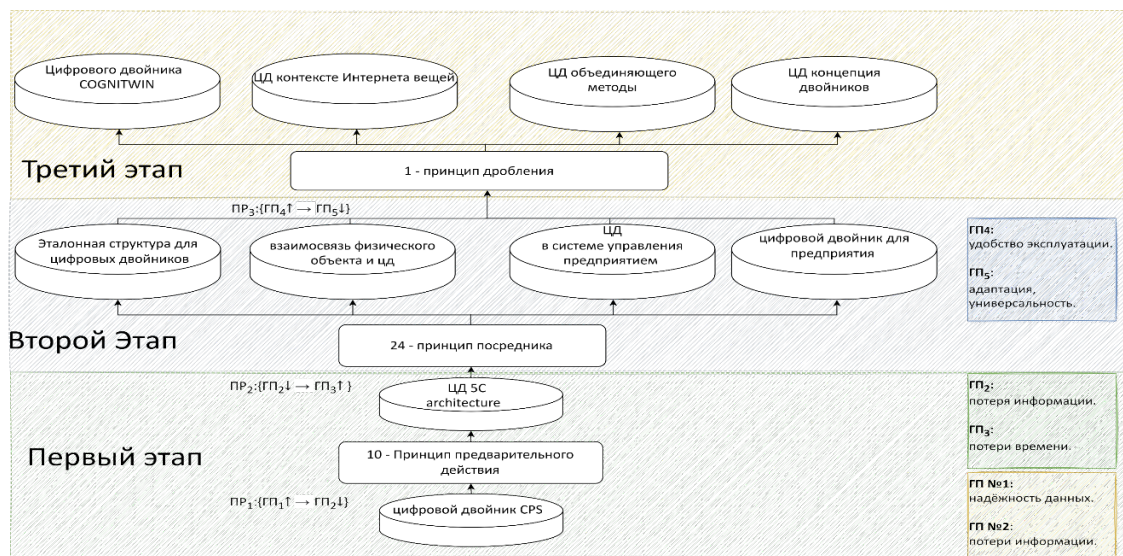


Рис. 19. ТРИЗ-эволюция цифровых двойников
Fig. 19. TRIZ evolution of digital twins

Она позволяет идентифицировать ключевые технологические проблемы и найти наиболее эффективные решения для их решения. Использование ТРИЗ-эволюционной карты может помочь в разработке новых и более инновационных продуктов и технологий.

Заключение

В рамках этой статьи были систематизированы знания о существующих механизмах реализации цифровых двойников и их архитектуры. При помощи ТРИЗ-эволюционного подхода проанализировано развитие цифровых двойников. Определены главные **противоречия**:

- 1) при увеличении надежности данных недопустимо увеличиваются потери информации;
- 2) при уменьшении потери информации недопустимо увеличиваются потери времени;
- 3) при увеличении удобства эксплуатации недопустимо ухудшаются адаптация, универсальность.

Также были выявлены приемы, которые стали «движущими силами» каждой новой модели, предложены приемы разрешения этих противоречий на основе **ТРИЗ-инструментов**:

- 1) предварительного действия;
- 2) принципа посредника;
- 3) принципа дробления.

Была создана ТРИЗ-эволюционная карта (рис. 19), отражающая развитие моделей цифровых двойников в различных альтернативных решениях. Для создания цифровых двойников для гибридных энергетических систем, предназначенных для поселений Крайнего Севера, был выбран вариант архитектуры концепции создания цифровых двойников комплексного подхода [39]. Он был выбран потому, что включает все необходимые модули, а также, что особенно важно, обеспечивает непрерывную интеграцию рассматриваемых подсистем с поддержкой функционирования всех программных модулей в режиме реального времени. Архитектура была выбрана из-за наличия всех необходимых модулей и обеспечения непрерывной интегра-

ции рассматриваемых подсистем, а также поддержки функционирования всех программных модулей в режиме реального времени.

Список литературы

1. Об утверждении Программы комплексного развития систем коммунальной инфраструктуры городского округа «Город Хабаровск» на период до 2035 года // Администрация города Хабаровска URL: <https://clck.ru/32D7W8> (дата обращения: 01.10.2022).
2. **Васильев Г. В., Бердонос В. Д.** Эффективная методика поиска источников в наукометрических базах данных // Актуальные проблемы информационно-телекоммуникационных технологий и математического моделирования в современной науке и промышленности: Материалы I Международной научно-практической конференции молодых ученых, Комсомольск-на-Амуре, 20–25 марта 2021 года. Комсомольск-на-Амуре: Комсомольский-на-Амуре государственный университет, 2021. С. 194–198. DOI 10.17084/978-5-7765-1488-3-2021-194. – EDN MJZWZD.
3. **Животова А. А., Бердонос В. Д.** Перспективные направления развития систем машинного перевода // Информатика и системы управления. 2022. № 2(72). С. 116–128. DOI 10.22250/18142400_2022_72_2_116. – EDN TZDCVP.
4. **Altshuller G.** The innovation algorithm: TRIZ, systematic innovation, and technical creativity. Worcester, Massachusetts: Technical Innovation Center, 1999.
5. **Gadd K.** TRIZ for Engineers: Enabling Inventive Problem Solving. John Wiley & Sons, 2011.
6. Creativity, learning techniques and TRIZ / T. Bertoncelli, O. Mayer, M. Lynass // Procedia CIRP. 2016. Vol. 39. P. 191–196.
7. ГОСТ Р 57700.37 – 2021 Компьютерные модели и моделирование Цифровые двойники изделий. Утв. и введен в действие Приказом Федер. агентства по техн. регулированию и метрологии от 16 сентября 2021 г. №9 979-ст. Москва: Стандартинформ, 2021. V, 11 с.
8. Отчет | Предложения Счетной палаты // Счетная палата Российской Федерации URL: <https://ach.gov.ru/upload/iblock/b99/b998773313b87e724ed09f287754d180.pdf> (дата обращения: 01.06.2022).
9. Цифровой двойник // Моделирование и цифровые двойники URL: <https://www.cadfem-cis.ru/service/digital-twin/#:~:text=Цифровой%20двойник%20%28Digital%20Twin%29%20-,%20воздействий%20помех%20и%20окружающей%20среды> URL: (дата обращения: 13.06.2022).
10. **Безруких П.** Возобновляемая энергетика: сегодня – реальность, завтра – необходимость. М.: Лесная страна, 2007. 120 с.
11. **Риполь-Сарагоси Т. Л., Кууск А. Б.** Возобновляемые и нетрадиционные источники энергии: учебное пособие. Ростов н/Д: Рост. гос. ун-т. путей сообщения, 2019. 122 с.
12. **Рудаков А. И., Хузятова Р. Р.** Гибридные источники энергии для электроснабжения энергетических систем // Мировая наука. 2019. № 6(27). С. 394–396.
13. **Лукутин Б. В., Суржикова О. А., Шандарова Е. Б.** Возобновляемая энергетика в децентрализованном электроснабжении: монография. М.: Энергоатомиздат, 2008. 231 с.
14. **Куприяновский В. П., Климов А. А., Воропаев Ю. Н. и др.** Цифровые двойники на базе развития технологий BIM, связанные онтологиями, 5G, IoT и смешанной реальностью для использования в инфраструктурных проектах и IFRABIM // International Journal of Open Information Technologies. 2020. Т. 8. № 3. С. 55–74.
15. **Steindl G., Stagl M., Kasper L., Kastner W., Hofmann R.** 2020. Ge-neric Digital Twin Architecture for Industrial Energy Systems. DOI 10.3390/app10248903
16. **Snijders R., Pileggi P., Broekhuijsen J., Verriet J., Wiering M., Kok K.** 2020. Machine Learning for Digital Twins to Predict Responsiveness of Cyber-Physical Energy Systems.

17. **Sultanov M. M., Arakelyan E. K., Boldyrev I. A., Lunenko V. S., Menshikov P. D.** Digital twins application in control systems for distributed generation of heat and electric energy. 2021. Pp. 89–101 DOI 10.24425/ather.2021.137555
18. **Darbali–Zamora R., Johnson J., Summers A., Jones C. B., Hansen C., Showalter C.** State Estimation–Based Distributed Energy Resource Optimization for Distribution Voltage Regulation in Telemetry–Sparse Environments Using a Real–Time Digital Twin. 2021. DOI 10.3390/en14030774
19. **Ahn E. Y., Lee J., Bae J., Kim J. M.** Analysis of Emerging Geotechnologies and Markets Focusing on Digital Twin and Environmental Monitoring in Response to Digital and Green New Deal. Vol. 20. P. 609–617 DOI 10.9719/EEG.2020.53.5.609
20. **Lydon G. P., Caranovic S., Hischier I., Schlueter A.** Coupled simulation of thermally active building systems to support a digital twin. 2019. DOI 10.1016/j.enbuild.2019.07.015
21. **Arafet K., Berlanga R.** Digital Twins in Solar Farms: An Approach through Time Series and Deep Learning. 2021. DOI 10.3390/a14050156
22. **Bai L. C., Zhang Y.T., Wei H. Q., Dong J. B., Tian W.** Digital Twin Modeling of a Solar Car Based on the Hybrid Model Method with Data–Driven and Mechanistic .2021. DOI 10.3390/app11146399
23. **Junior C. A A. D., Villanueva J. M. M., Almeida R. J. S. D., de Medeiros I. E. A.** Digital Twins of the Water Cooling System in a Power Plant Based on Fuzzy Logic. 2021. DOI 10.3390/s21206737
24. **Mawson V. J., Hughes B.** The development of modelling tools to improve energy efficiency in manufacturing processes and systems. 2019. P. 95–105. DOI 10.1016/j.jmsy.2019.04.008
25. **Jakobi M., Kunath L., Witzig A.** BIM Use–Case: Model–Based Performance Optimization. 2018. P. 1395–1402. DOI 10.18086/eurosun2018.11.07
26. **Parrott A.** Industry 4.0 and the digital twin manufacturing meets its match A. Parrott, L. Warshaw. [Электронный ресурс]. URL: <https://www2.deloitte.com/us/en/insights/focus/industry–4–0/digital–twin–technologysmart–factory.html>.
27. **Массель Л. В., Ворожцова Т. Н.** Онтологический подход к построению цифровых двойников объектов и систем энергетики // Онтология проектирования. 2020. Т. 10, № 3(37). С. 327–337. DOI 10.18287/2223-9537-2020-10-3-327-337
28. **Golovina T., Polyanin A., Adamenko A., Kheday E., Schepinin V.** Digital Twins as a New Paradigm of an Industrial Enterprise. 2020. P. 1115–1124. DOI 10.14716/ijtech.v11i6.4427
29. **Garcia H. E., Aumeier S. E., Al–Rashdan A. Y., Rolston B. L.** Secure embedded intelligence in nuclear systems: Framework and methods. 2020. DOI 10.1016/j.anucene.2019.107261
30. **Jain P., Poon J., Singh J. P., Spanos C., Sanders S. R., Panda S. K.** A Digital Twin Approach for Fault Diagnosis in Distributed Photovoltaic Systems. 2020. P. 940–956. DOI 10.1109/TPEL.2019.2911594
31. **Tutelea L. N., Deaconu S. I., Popa G. N.** Reduced Cost Low Speed Wind or Hydro Energy Conversion System with Twin Stator Windings Induction Generator. 2014. P. 317–324.
32. **Wang M. M., Wang C. Y., Hnydiuk–Stefan A., Feng S. Z., Atilla I., Li Z.** Recent progress on reliability analysis of offshore wind turbine support structures considering digital twin solutions. 2021. DOI 10.1016/j.oceaneng.2021.109168
33. **Gong H. J., Rooney T., Akeyo O. M., Branecky B., Ionel D. M.** Equivalent Electric and Heat–Pump Water Heater Models for Aggregated Community–Level Demand Response Virtual Power Plant Controls. 2021. P. 141233–141244. DOI 10.1109/ACCESS.2021.3119581
34. **Ganesh S., Perilla A., Torres J. R., Palensky P., van der Meijden M.** Validation of EMT Digital Twin Models for Dynamic Voltage Performance Assessment of 66 kV Offshore Transmission Network. 2021. DOI 10.3390/app11010244
35. **Etxegarai A., Valverde V., Eguia P., Perea E.** Analysis of Loss of Life of Dry–Type WTSU Transformers in Offshore Wind Farms. 2020. P.448–455.

36. **Alsmadi Y. M., Xu L. Y., Wang A. M.** Development of a Computer Twins–Based Wind Farm Testbed. 2015. P. 1005–1012.
37. **Trancossi M., Cannistraro G., Pascoa J.** Thermoelectric and solar heat pump use toward self sufficient buildings: The case of a container house. 2020. DOI 10.1016/j.tsep.2020.100509
38. **Boikov A. V., Savelev R. V., Payor V. A., Erokhina O. O.** THE CONTROL METHOD CONCEPT OF THE BULK MATERIAL BEHAVIOR IN THE PELLETIZING DRUM FOR IMPROVING THE RESULTS OF DEM–MODELING. 2019. Art. 41548. DOI 10.17580/cisistr.2019.01.02
39. **Gandzha S., Aminov D., Sogrin A.** Development of Water Submersible Gyro Generator of Combined Excitation for Energy Development of Small and Medium Rivers. 2020.
40. **Behling N., Williams M. C., Managi S.** Fuel cells and the hydrogen revolution: Analysis of a strategic plan in Japan. 2015. P. 204–221. DOI 10.1016/j.eap.2015.10.002
41. **Oti A. H., Farrell P., Abanda F. H., McMahon P., Mahamadu A. M., Mzyece D., Akintola A. A., Prinja N.** A BIM–driven framework for integrating rules and regulations in the decommissioning of nuclear power plants. DOI 10.1108/CI-11-2020-0186
42. **Manzoor M. T., Lenci G., Tetreault–Friend M.** Convection in volumetrically absorbing solar thermal receivers: A theoretical study. 2021. P. 1358–1368. DOI 10.1016/j.solener.2021.06.048
43. **Glaessgen E. H.** The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. 2012.
44. **Радионо́в А. А., Кара́ндаев А. С., Логи́нов Б. М., Гаси́ярова О.** Концептуальные направления создания цифровых двойников электротехнических систем агрегатов прокатного производства.
45. **Kořinek M., Tázlar O., Štekerová K.** Digital Twin Models: BIM Meets NetLogo. 2021.
46. **Abburu S., Berre A. J., Jacoby M., Roman D., Stojanovic L., Stojanovic N.** COGNITWIN – Hybrid and Cognitive Digital Twins for the Process Industry. 2020. Pp. 1–8. DOI 10.1109/ICE/ITMC49519.2020.9198403
47. **He B., Bai K.–J.** Digital twin–based sustainable intelligent manufacturing: a review. 2021. DOI 10.1007/s40436-020-00302-5
48. **Ghosh A. K., Ullah S., Teti R., Kubo A.** Developing sensor signal–based digital twins for intelligent machine tools.
49. **Васильев Г. В., Бердоносков В. Д.** Разработка и сравнение моделей прогнозирования потребления электроэнергии с помощью рекуррентных нейросетей с долгой краткосрочной памятью // Наука, инновации и технологии: от идей к внедрению: Материалы Международной научно-практической конференции, Комсомольск-на-Амуре, 07–11 февраля 2022 года. Комсомольск-на-Амуре: Комсомольский-на-Амуре государственный университет, 2022. С. 203–206. EDN FDYOTE.
50. **Васильев Г. В., Бердоносков В. Д.** Анализ моделей потоков энергий гибридных энергетических систем // Актуальные проблемы информационно-телекоммуникационных технологий и математического моделирования в современной науке и промышленности: Материалы I Международной научно-практической конференции молодых ученых, Комсомольск-на-Амуре, 20–25 марта 2021 года. Комсомольск-на-Амуре: Комсомольский-на-Амуре государственный университет, 2021. С. 191–194. DOI 10.17084/978-5-7765-1488-3-2021-191

Информация об авторе

Георгий Андреевич Тимофеев, ассистент, кафедра «Проектирование, управление и разработка информационных систем», Комсомольский-на-Амуре государственный университет

Information about the Author

Georgy A. Timofeev, assistant, Department “Projecting, Management and Development of Information Systems”, Komsomolsk-on-Amur State University

*Статья поступила в редакцию 17.12.2022;
одобрена после рецензирования 05.03.2023; принята к публикации 05.03.2023
The article was submitted 17.12.2022;
approved after reviewing 05.03.2023; accepted for publication 05.03.2023*

Правила оформления текста рукописи

Авторы представляют статьи на русском или английском языке объемом от 0,5 авторского листа (20 тыс. знаков) до 1 авторского листа (40 тыс. знаков), включая иллюстрации (1 иллюстрация форматом 190×270 мм = 1/6 авторского листа, или 6,7 тыс. знаков). Публикации, превышающие указанный объем, допускаются к рассмотрению только после индивидуального согласования с редакцией журнала.

Текст рукописи должен быть представлен в редколлегию в виде файла MS Word (.doc, .docx). Гарнитура Times New Roman, размер шрифта 11, межстрочный интервал 1, размеры полей – стандартные значения текстового редактора. Форматирование – выравнивание по ширине страницы, переносы слов включены, каждый новый абзац начинается с красной строки. Не допускается ручное форматирование абзацев (пробелами, лишними переводами строк, разрывами страниц).

Структура статьи

- Индекс УДК (универсальной десятичной классификации). Выравнивание по левому краю
- Название статьи. Выравнивание по центру, полужирный шрифт
- ФИО авторов (полностью). Выравнивание по центру, полужирный шрифт
- Места работы всех авторов. Выравнивание по центру, курсив
- Адреса электронной почты, ORCID авторов
- Аннотация статьи
- Ключевые слова, не более 10
- Благодарности, сведения о финансовой поддержке
- Название статьи **на английском языке**. Выравнивание по центру, полужирный шрифт
- ФИО авторов **на английском языке** (полностью). Выравнивание по центру, полужирный шрифт
- Места работы авторов **на английском языке**. Выравнивание по центру, курсив
- Аннотация статьи **на английском языке (Abstract)**, 200–250 слов
- Ключевые слова **на английском языке (Keywords)**, не более 10
- Благодарности, сведения о финансовой поддержке **на английском языке**, если есть соответствующий раздел на русском языке (**Acknowledgements**)
- Основной текст
- Список литературы / **References**
- Сведения об авторах

Требования к оформлению основного текста и иллюстративных материалов

Основной текст должен быть представлен в структурированном виде, рекомендуется использовать подзаголовки – например: Введение, Методика..., Выводы, Результаты, Заключение.

Подзаголовки отделяются и набираются полужирным шрифтом. В целях выделения частей текста и отдельных слов и словосочетаний допускается использование курсива или полужирного шрифта. Подчеркивание, разрядка, изменение основного кегля и выделение цветом не используются.

Иллюстрации к рукописи статьи должны быть приложены в виде отдельных файлов. При этом в тексте должно содержаться включенное изображение с указанием имени файла. Все иллюстрации, содержащие схемы, графики, алгоритмы и т. п., должны быть представлены в векторном виде (.ai, .eps, .cdr). Скриншоты и другие растровые изображения должны быть представлены в максимально высоком качестве, без каких-либо потерь и искажений (.jpg, .tif). Все иллюстрации должны иметь подрисуночную подпись – свое название. Надписи к таблицам и подписи к иллюстрациям приводятся **на двух языках (русском и английском)**.

Примеры:

Рис. 1. Диаграмма производительности...

Fig. 1. Performance diagram...

Таблица 1

Сравнение алгоритмов...

Table 1

Comparison of algorithms...

Нумерация последовательная и неразрывная от начала статьи. Не допускается использование других наименований, кроме «Рис.» / «Fig.», «Таблица» / «Table», и усложнение нумерации (например, «Рис. 3.2.»). Ссылка на иллюстрацию в тексте должна быть приведена в круглых скобках, например: (рис. 1), (табл. 1).

Формулы должны быть набраны с использованием редактора MathType либо встроенного редактора формул MS Word. Кегль основных символов – 11, греческие символы набираются прямым шрифтом, латинские – курсивом. Нумеруются только те формулы, на которые автор ссылается в тексте.

Abstract

Аннотация статьи на английском языке (Abstract) не должна быть дословным переводом русскоязычной аннотации. Раздел Abstract, как и основной текст, должен быть структурирован, в нем должно содержаться описание цели работы, методов исследования, научной значимости, выводов / результатов. Требуется качественный перевод на английский язык (при необходимости просим авторов обращаться к профессиональным переводчикам). **Объем Abstract 200–250 слов.**

Список литературы / References

Список литературы и список литературы на английском языке (References) размещаются в общем разделе. Рекомендуемое количество цитируемых в статье источников – не менее 10, в список желательно включать ссылки на актуальные работы по теме исследования, особенно в иностранных периодических изданиях.

В тексте статьи ссылки на литературу указываются цифрами в квадратных скобках, при необходимости указываются номера страниц, например: [2; 3. С. 15].

Список литературы нумеруется в порядке цитирования и оформляется в соответствии с ГОСТ Р 7.0.5-2008 на библиографическое описание (знаки тире в описании опускаются). Ссылки на неопубликованные работы, а также на Интернет-ресурсы (кроме электронных изданий, поддающихся библиографическому описанию) оформляются в виде сноски.

В Список литературы ссылки на источники следует включать на оригинальном языке опубликования. Каждый источник должен быть также оформлен на английском языке (References) по международному стандарту для публикаций в области информатики IEEE Style со следующими отличиями:

- инициалы авторов указываются после фамилии;
- название статьи не берется в кавычки, отделяется точкой;

- отсутствует союз «and» перед фамилией последнего автора;
- в диапазоне страниц – удвоенная «р» (например, «pp. 2–9»);
- год издания указывается после места издания (для книг) и сразу после названия журнала (для периодики).
- Перевод источника на английский язык:
- если источник имеет выходные данные на английском языке, то для формирования References **следует использовать именно эти данные**;
- если оригинальная публикация не содержит выходных данных на английском языке, то допускается транслитерация названия материала на латинский алфавит в сочетании с переводом на английский язык в квадратных скобках. В конце описания указывается, на каком языке написана эта работа, например, (in Russ.). При транслитерации можно воспользоваться Интернет-ресурсом <http://ru.translit.ru/>, рекомендуется выбрать стандарт BSI. Место издания не транслитерируется, указывается полностью на английском языке, например: Moscow. Название издательства / издателя, как правило, транслитерируется. Для журналов, у которых есть официальное название на английском языке, – использовать его (проверить на сайте журнала, или, например, в библиотеке WorldCat), если названия на английском языке нет, использовать транслитерацию по системе BSI. Не следует самостоятельно переводить названия журналов.

Если у цитируемого источника есть **цифровой идентификатор DOI** (<https://search.crossref.org/>), его требуется обязательно указывать в конце библиографической ссылки.

Примеры оформления ссылок. Каждый источник в том же пункте дублируется на английском языке (References).

Источник на русском языке, перевод на английский доступен в метаданных статьи

1. Журавлев С. С., Рудометов С. В., Окольников В. В., Шакиров С. Р. Применение модельно-ориентированного проектирования к созданию АСУ ТП опасных промышленных объектов // Вестник НГУ. Серия: Информационные технологии. 2018. Т. 16, № 4. С. 56–67. DOI 10.25205/1818-7900-2018-16-4-56-67

Zhuravlev S. S., Rudometov S. V., Okolnishnikov V. V., Shakirov S. R. Model-Based Design Approach for Development Process Control Systems of Hazardous Industrial Facilities. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, pp. 56–67. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-56-67

Источник на английском языке. Оформляем согласно требованиям для References. Приводим только 1 раз.

2. Telnov V. I. Optimization of the Beam Crossing Angle at the ILC for E + e- and yy Collisions. *Journal of Instrumentation*, 2018, vol. 13, no. 03, pp. P03020–P03020. DOI 10.1088/1748-0221/13/03/p03020

Метаданные источника доступны только на русском языке

3. Жижимов О. Л., Федотов А. М., Шокин Ю. И. Технологическая платформа массовой интеграции гетерогенных данных // Вестник НГУ. Серия: Информационные технологии. 2013. Т. 11, вып. 1. С. 24–41.

Zhizhimov O. L., Fedotov A. M., Shokin Yu. I. Tekhnologicheskaya platforma massovoi integratsii geterogennykh dannykh [Technology Platform for the Mass Integration of Heterogeneous Data]. *Vestnik NSU. Series: Information Technologies*, 2013, vol. 11, no. 1, pp. 24–41. (in Russ.)

Сведения об авторах

Последний раздел статьи – информация об авторе / авторах **на русском и английском языках**:

- ФИО полностью, ученая степень, ученое звание;
- идентификаторы автора, такие как ResearcherID (всем авторам рекомендуется использовать данные сервисы для ведения актуального списка своих публикаций);
- контактный телефон (не публикуется).

Если статья представляется на английском языке, необходимо приложить перевод на русский язык названия, аннотации, ключевых слов, сведений об авторе.

Доставка материалов

Материалы предоставляются в редакцию по электронной почте inftech@vestnik.nsu.ru.

Порядок рецензирования

Все статьи сначала проходят проверку на заимствование и только после этого отправляются на рецензирование. Редакционный совет не допускает к публикации материал, если имеется достаточно оснований полагать, что он является плагиатом.

Тип рецензирования статей – двухуровневое, одностороннее анонимное («слепое»).

Для каждой статьи редколлегией выбираются рецензенты, научная деятельность которых связана с темой представленного материала. Ответственный секретарь журнала обращается к ним с просьбой дать экспертную оценку статье либо помочь организовать рецензирование.

Рецензии для журнала «Вестник НГУ. Серия: Информационные технологии» составляются по единой схеме и подразумевают оценку по следующим критериям: соответствие тематике журнала, оригинальность и значимость результатов, качество изложения материала.

Заполненный бланк рецензии высылается на электронный адрес редакции. В зависимости от экспертных заключений статья может быть принята редакционным советом к опубликованию, рекомендована автору к доработке (с последующим повторным рецензированием либо без него) или отклонена (с предоставлением автору мотивированного отказа). Автору на электронный адрес высылается текст рецензии без указания ФИО рецензента и его контактных данных.

Все рецензии хранятся в редакции журнала не менее 5 лет. Редколлегия журнала обязуется при поступлении соответствующего запроса направлять копии рецензий в Министерство науки и высшего образования Российской Федерации.