

ISSN 2413 - 0133

Scientific journal

# ИНФОРМАЦИОННЫЕ И МАТЕМАТИЧЕСКИЕ ТЕХНОЛОГИИ В НАУКЕ И УПРАВЛЕНИИ

№4(20)/ 2020

6+

**ИНФОРМАЦИОННЫЕ И МАТЕМАТИЧЕСКИЕ  
ТЕХНОЛОГИИ В НАУКЕ И УПРАВЛЕНИИ**

**Научный журнал**

**№ 4 (20)**



## EDITORIAL BOARD

Ablameyko S.V.  
 Arshinskiy L.V.  
 Berestneva O.G.  
 Boukhanovsky A.V.  
 Bychkov I.V.  
 Eliseev S.V.  
 Gornov A.Y.  
 Gribova V.V.  
 Groumpos P.  
 Hodashinsky I.A.  
 Zorina T.G.  
 Kalimoldaev M.N.  
 Karpenko A.P.  
 Kazakov A.L.  
 Khamisov O.V.  
 Komendantova N.P.  
 Kureichik V.V.  
 Lis R.  
 Massel L.V.  
 Mokhor V.V.  
 Moskvichev V.V.  
 Ovtcharova J.  
 Popov G.T.  
 Smirnov S.V.  
 Stylios C.  
 Taratukhin V.V.  
 Voevodin V.V.  
 Voropai N.I.  
 Woern H.  
 Wolfengagen V.E.  
 Yusupova N.I.  
 Chubarov L.B.

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ ЖУРНАЛА

Абламейко С.В., академик НАН Беларуси, Минск, БГУ  
 Аршинский Л.В., д.т.н., Иркутск, ИрГУПС  
 Берестнева О.Г., д.т.н., Томск, ТПУ  
 Бухановский А.В., д.т.н., Санкт-Петербург, НИУ ИТМО  
 Бычков И.В., академик РАН, Иркутск, ИДСТУ СО РАН  
 Елисеев С.В., д.т.н., Иркутск, ИрГУПС  
 Горнов А.Ю., д.т.н., Иркутск, ИДСТУ СО РАН  
 Грибова В.В., д.т.н., Владивосток, ИАПУ ДВО РАН  
 Грумπος Π., Греция, University of Patras  
 Ходашинский И.А., д.т.н., Томск, ТУСУР  
 Зорина Т.Г., д.т.н., Республика Беларусь, Институт энергетики НАН Беларуси  
 Калимолдаев М.Н., академик НАН РК, Республика Казахстан, ИИВТ  
 Карпенко А.П., д.ф.-м.н., Москва, МГТУ им. Баумана  
 Казаков А.Л., д.ф.-м.н., Иркутск, ИДСТУ СО РАН  
 Хамисов О.В., д.ф.-м.н., Иркутск, ИСЭМ СО РАН  
 Комендантова Н.П., PhD, Австрия, Лаксенбург, ПАСА  
 Курейчик В.В., д.т.н., профессор ЮФУ, Таганрог  
 Лис Р., Польша, Wrocław University of Science and Technology  
 Массель Л.В., д.т.н., Иркутск, ИСЭМ СО РАН  
 Мохор В.В., д.т.н., Киев, ИПМЭ им. Г.Е. Пухова НАН Украины  
 Москвичев В.В., д.т.н., Красноярск, СКТБ «Наука» СО РАН  
 Овчарова Ж., Германия, Karlsruhe Institute of Technology (KIT)  
 Попов Г.Т., Болгария, г. София, Технический университет  
 Смирнов С.В., д.т.н., Самара, ИПУСС РАН  
 Стилос Х., Греция, Technological Educational Institute of Epirus  
 Таратухин В.В., Германия, ERCIS, University of Muenster  
 Воеводин В.В., чл.-корр. РАН, Москва, НИВЦ МГУ  
 Воропай Н.И., чл.-корр. РАН, Иркутск, ИСЭМ СО РАН  
 Вёрн Х., Германия, Karlsruhe Institute of Technology (KIT)  
 Вольфенгаген В.Э., д.т.н., Москва, МИФИ  
 Юсупова Н.И., д.т.н., Уфа, УГАТУ  
 Чубаров Л.Б., д.т.н., Новосибирск, ИВТ СО РАН

## EXECUTIVE EDITORIAL

Chief Editor Massel L.V.  
 Executive Editor  
 Bakhvalova Z.A.  
 Editor Kopaigorodsky A.N.  
 Editor Massel A.G.  
 Designer Pesterev D.V.

## ИСПОЛНИТЕЛЬНАЯ РЕДАКЦИЯ ЖУРНАЛА

Главный редактор	Массель Л.В.	д.т.н., Иркутск, ИСЭМ СО РАН
Выпускающий редактор	Бахвалова З.А.	к.т.н., Иркутск, ИСЭМ СО РАН
Редактор	Копайгородский А.Н.	к.т.н., Иркутск, ИСЭМ СО РАН
Редактор	Массель А.Г.	к.т.н., Иркутск, ИСЭМ СО РАН
Дизайнер	Пестерев Д.В.	Иркутск, ИСЭМ СО РАН

## Адрес учредителя, издателя и редакции

Федеральное государственное бюджетное учреждение науки  
 Институт систем энергетики им. Л.А. Мелентьева  
 Сибирского отделения Российской академии наук (ИСЭМ СО РАН)  
 664033 г. Иркутск, ул. Лермонтова, 130  
 Тел: (3952) 42-47-00 Факс: (3952) 42-67-96

Раб. тел.: 8 (3952) 500-646 доп. 441

Массель Л.В. e-mail: [massel@isem.irk.ru](mailto:massel@isem.irk.ru)

Раб. тел.: 8 (3952) 500-646 доп. 440

Бахвалова З.А. e-mail: [zinand@isem.irk.ru](mailto:zinand@isem.irk.ru)

Сайт журнала и конференции ИМТ - <https://www.imt-journal.ru/>

Журнал включен в Российский индекс научного цитирования (РИНЦ). Номер контракта 202-04/2016  
 Журнал зарегистрирован в Роскомнадзоре. Регистрационный номер ПИ № ФС 77 – 73539.

Отпечатано в полиграфическом участке ИСЭМ СО РАН

© Все права принадлежат авторам публикуемых статей.

Дата выхода 18.12.2020 г. Тираж 100 экз.

© Издательство ИСЭМ СО РАН  
 Цена свободная. (6+)

**Онтологическое и когнитивное моделирование**

- Захарова А.А., Подвесовский А.Г., Исаев Р.А.** Нечеткие когнитивные модели в управлении слабоструктурированными социально-экономическими системами 5
- Массель А.Г., Пяткова Н.И.** Применение методов когнитивного моделирования для анализа угроз энергетической безопасности 24
- Грибова В. В., Шалфеева Е.А.** Комплекс средств поддержки процессов разработки и сопровождения решателей для систем с онтологическими базами знаний 34
- Кононенко И.С., Ахмадеева И.Р., Сидорова Е. А.** Лингвистические аспекты исследования аргументации на основе онтологии 44

**Методы и средства обеспечения кибербезопасности**

- Колосок И.Н., Гурина Л.А.** Обработка синхронизированных векторных измерений методами нечеткой логики при кибератаках на электро-энергетическую систему 56
- Гаськова Д.А.** Метод определения уровня киберситуационной осведомленности энергетических объектов 64
- Ходашинский И.А., Бардамова М. Б.** Модификации алгоритма прыгающих лягушек для отбора признаков в нечетком классификаторе при аутентификации пользователя по рукописной подписи 75
- Саломатин А.А., Исхаков А. Ю.** Применение интегрированного показателя отпечатков браузера в задаче адаптивной аутентификации субъектов доступа 84

**Информационные и математические технологии**

- Елисеев А. В.** Системные представления динамических процессов в механических колебательных системах с особенностями 93
- Черкашин Е.А., Шигаров А., Христюк В.** Информационная инфраструктура для поддержки исследований микробиома Байкала 108
- Стойлова А.С., Машкина Л.В., Блам Ю.Ш.** Информационное согласование народнохозяйственной и отраслевой модели на примере лесного комплекса российской Федерации 124

**Технологии семантического анализа**

- Тучкова Н. П.** О семантической модели предметной области «Уравнения математической физики» 132
- Ворожцова Т. Н., Пестерев Д.В., Ан Г.В.** Семантические технологии управления знаниями для поддержки совместных энергетических и социо-экологических исследований 143
- Михеев А.В.** Анализ больших данных для обоснования решений по научно-технологическому развитию в энергетике 158
- Копайгородский А.Н., Мамедов Т.Г.** Архитектура интеллектуальной информационной системы для поддержки экспертных решений по стратегическому инновационному развитию энергетики 168

**Ontological and Cognitive Modeling**

- Zakharova A.A., Podvesovskii A.G., Isaev R.A.** Fuzzy cognitive models in management of semi-structured socio-economic systems 5
- Massel A. G., Pyatkova N. I.** Application of cognitive modeling methods for energy security threat analys 24
- Gribova V.V., Shalfeeva E.A.** A set of tools to support development and maintenance of solvers for systems with ontological knowledge bases 34
- Kononenko I. S., Akhmadeeva I. R., Sidorova E. A.** Linguistic aspects of ontology-based argumentation study 44

**Cybersecurity Methods and Tools**

- Kolosok I. N., Gurina L. A.** Processing of synchronous phasor measurements by fuzzy logic methods in the case of cyberattacks on information-communication infrastructure of a cyber-physical electric power system 56
- Gaskova D. A.** Method for determining the level of cyber situational awareness on energy facilities 64
- Hodashinsky I.A., Bardamova M. B.** Feature selection for fuzzy classifiers using the ranking and cross-validation 75
- Salomatin A. A., Iskhakov A. Yu.** Application of the integrated indicator of browser fingerprinting in the problem of adaptive authentication of access subjects 84

**Information and Mathematical Technologies**

- Eliseev A.V.** System representations of dynamic processes in mechanical oscillatory systems with special features 93
- Cherkashin E., Shigarov A., Khristyuk V.** Information infrastructure for supporting baikal microbiome research 108
- Stoylova A. S., Mashkina L. V., Blam Yu. Sh.** Information coordination of the national economic and industry models on the example of the Russian federation forestry 124

**Semantic Analysis Technologies**

- Tuchkova N.P.** On semantic model of subject domain "Mathematical physics equations" 132
- Vorozhtsova T.N., Pesterev D.V., An G.V.** Semantic technologies in knowledge management to support joint energy and socio-ecological research 143
- Mikheev A.V.** Big data analysis for innovation development decision making in energy sector 158
- Kopaygorodsky Alex N., Mamedov Timur G.** Architecture of the intellectual information system to support expert decisions on strategic innovative energy development 168

**НЕЧЕТКИЕ КОГНИТИВНЫЕ МОДЕЛИ В УПРАВЛЕНИИ  
СЛАБОСТРУКТУРИРОВАННЫМИ  
СОЦИАЛЬНО-ЭКОНОМИЧЕСКИМИ СИСТЕМАМИ**

**Захарова Алёна Александровна**

Д.т.н., доцент, профессор кафедры «Информатика и программное обеспечение»

e-mail: [zaa@tu-bryansk.ru](mailto:zaa@tu-bryansk.ru)

ORCID: 0000-0003-4221-7710,

**Подвесовский Александр Георгиевич**

К.т.н., доцент, заведующий кафедрой «Информатика и программное обеспечение»

e-mail: [apodv@tu-bryansk.ru](mailto:apodv@tu-bryansk.ru)

ORCID: 0000-0002-1118-3266,

**Исаев Руслан Александрович**

Ассистент кафедры «Информатика и программное обеспечение»,

e-mail: [ruslan-isaev-32@yandex.ru](mailto:ruslan-isaev-32@yandex.ru)

ORCID: 0000-0003-3263-4051,

Брянский государственный технический университет,  
241035, Россия, г. Брянск, бул. 50 лет Октября, д. 7,

**Аннотация.** В статье рассматривается проблематика моделирования слабоструктурированных социально-экономических систем на основе применения когнитивного подхода. Предлагается авторская информационная технология поддержки когнитивного моделирования таких систем, основанная на использовании нечетких когнитивных карт. Описывается опыт применения предложенной технологии и реализующей ее программной системы при решении ряда практических задач исследования стратегий управления социально-экономическими системами.

**Ключевые слова:** управление социально-экономическими системами, слабоструктурированная система, когнитивная модель, нечеткая когнитивная карта, структурно-целевой анализ, сценарный анализ.

**Цитирование:** Захарова А.А., Подвесовский А.Г., Исаев Р.А. Нечеткие когнитивные модели в управлении слабоструктурированными социально-экономическими системами // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 5-23. DOI: 10.38028/ESI.2020.20.4.001

**Введение.** В современной практике управления организационными и социально-экономическими системами лицу, принимающему решение (ЛПР), все чаще приходится иметь дело со слабоструктурированными задачами управления и ситуациями принятия решений. Одной из причин этого является недостаток информации о состоянии системы в условиях слабо контролируемой и изменяющейся внешней среды. Отсутствие достаточных знаний о системе, относительно которой принимаются управленческие решения, не является единственной неопределенностью, обусловленной субъективными причинами. Также можно выделить неопределенность, противоречивость, сложную структуру взаимосвязи целей

развития социально-экономических систем и, как следствие, сложность формирования критериев оценки управленческих решений. Как правило, неудовлетворенность текущим состоянием системы осознается субъектом управления, но его представления о причинах и возможных способах изменения ситуации размыты, нечетки и противоречивы. Выявление таких представлений и их формализация являются одними из главных задач, которые необходимо решать при разработке моделей и методов принятия решений в слабоструктурированных системах и ситуациях.

Одним из подходов, применяемых для разработки, анализа и обоснования решений в слабоструктурированных системах, является когнитивный подход. Этот подход ориентирован на структуризацию знаний о слабоструктурированной системе, для понимания основных процессов, протекающих в ней. При этом множество процессов описываются в виде модели экспертных знаний о законах и закономерностях функционирования системы, с преобладанием субъективных оценок и лингвистических значений. В рамках когнитивного подхода способом описания экспертных знаний об исследуемой системе и происходящих в ней процессах является когнитивная модель, которая допускает формальное представление в виде когнитивной карты. Когнитивной картой называется причинно-следственная сеть, которая отражает субъективное представление исследователя о системе (индивидуальное или коллективное) в виде множества семантических категорий, называемых факторами или концептами, и множества причинно-следственных связей между ними [1, 5, 9, 22].

Когнитивная модель представляет собой эффективный инструмент для разведочного, оценочного анализа ситуации. Она не позволяет получить точные, количественные характеристики исследуемой системы, но позволяет оценить тенденции и тренды, связанные с ее функционированием и развитием, и выявить значимые факторы, оказывающие наибольшее влияние на указанные процессы. Как отмечено в [14], считается, что знание грубой, возможно даже гипотетической модели системы позволяет прогнозировать сценарии развития ситуаций при различных управляющих воздействиях путем варьирования переменных модели. Благодаря этому появляется возможность поиска и генерации эффективных решений по управлению системой, а также выявления рисков и разработки стратегий их снижения.

В статье рассматривается предложенная авторами информационная технология поддержки когнитивного моделирования в управлении слабоструктурированными социально-экономическими системами, а также примеры ее применения при решении практических задач.

**1. Структура и особенности представления когнитивной модели на основе нечетких когнитивных карт.** В общем случае построение и анализ когнитивной модели предполагает решение следующих задач [9]:

1. Выявление множества факторов (концептов), характеризующих моделируемую систему и ее окружение.
2. Выявление причинно-следственных отношений, которые возникают между этими факторами в процессе функционирования и/или развития системы, а также влияний на систему со стороны внешних факторов.
3. Построение когнитивной карты системы в виде причинно-следственной сети, объединяющей факторы и отношения между ними.

4. Применение к построенной когнитивной карте методов аналитической обработки, направленных на исследование структуры причинно-следственных связей, выявление наиболее значимых факторов, циклов обратной связи и др., а также получение прогнозов развития ситуации при различных управляющих воздействиях на систему и возмущениях со стороны внешней среды.

В качестве математического аппарата, применяемого для представления когнитивных моделей и лежащего в основе методов их анализа, чаще всего используется нечеткая логика. Благодаря этому возник целый класс когнитивных моделей, основанный на так называемых нечетких когнитивных картах [5]. Одной из разновидностей нечетких когнитивных карт (НКК) являются так называемые НКК Силова, предложенные В.Б. Силовым в монографии [22]. Приведем более развернутое описание структуры когнитивной модели, основанной на НКК Силова, и общих принципов ее формального представления (описание частично базируется на материале работы [9]):

- Моделируемая система описывается конечным множеством концептов и причинно-следственных связей между ними.
- Концепты могут выражать как относительные (качественные), так и абсолютные (измеримые) характеристики системы.
- Когнитивная карта системы представляет собой взвешенный ориентированный граф, вершины которого соответствуют концептам, а дуги – причинно-следственным связям.
- С концептами связываются переменные состояния, которые могут принимать значения, выраженные в некоторой шкале, в пределах установленных ограничений.
- Причинно-следственные связи могут быть положительными или отрицательными. Увеличение значения переменной состояния концепта-причины приводит к увеличению значения переменной состояния концепта-следствия при положительной причинно-следственной связи, и к уменьшению этого значения при отрицательной.
- Причинно-следственные связи различаются по силе. Сила связи может быть постоянной либо переменной во времени. Чем она больше, тем сильнее изменение состояния концепта-причины влияет на состояние концепта-следствия.
- Среди концептов можно выделить: целевые (которые необходимо привести в заданное целевое состояние); управляемые (состояние которых поддается непосредственному управлению); наблюдаемые (состояние которых не может задаваться непосредственно, и определяется изменением состояний концептов-причин); внешние (состояние которых подвержено влиянию со стороны внешних факторов, не отраженных в модели).
- Текущее состояние системы описывается состояниями всех ее концептов.
- Исходное состояние системы задается вектором начальных состояний концептов.
- Существует некоторое целевое состояние, описываемое вектором состояний подмножества целевых концептов, в которое необходимо привести систему либо максимально к нему приблизиться.
- Под воздействием на концепт понимается допустимое изменение его состояния относительно текущего.
- Под управляющим воздействием на систему понимается вектор воздействий на подмножество управляемых концептов. Множество управляющих воздействий конечно,

поскольку число управляемых концептов конечно и обычно невелико, также ограничен диапазон возможных их значений и шаг изменения этих значений.

- Под внешним воздействием понимается вектор допустимых воздействий на подмножество внешних концептов.
- Сценарий описывает изменение состояния системы, вызванное управляющими и внешними воздействиями на нее.

Формализованное представление нечеткой когнитивной модели, основанное на перечисленных положениях, можно найти, например, в авторских работах [13, 18, 21].

К построенной когнитивной модели применяются два вида анализа: статический (или структурно-целевой) и динамический (или сценарный). В рамках структурно-целевого анализа решаются следующие задачи:

- Выявление концептов, оказывающих наибольшее влияние на целевые концепты и систему в целом, а также концептов, в наибольшей степени подверженных влиянию со стороны системы и других концептов. Это позволяет находить факторы, наиболее чувствительные по отношению к управляющим и внешним воздействиям, а также определять точки наиболее эффективного приложения управляющих воздействий.
- Выявление устойчивых подмножеств концептов, в совокупности оказывающих значительное влияние на систему в целом, что позволяет находить и исследовать комплексные управляющие воздействия, одновременно затрагивающие множество факторов.
- Нахождение циклов положительной и отрицательной обратной связи.

Сценарный анализ ориентирован на исследование сценариев развития системы и управления ею. Типовыми задачами сценарного анализа являются:

- моделирование сценариев саморазвития, т.е. сценариев поведения системы без вмешательства в нее извне либо при наличии только внешних воздействий;
- моделирование сценариев управляемого развития, т.е. сценариев поведения системы при различных управляющих воздействиях.

**2 Информационная технология поддержки когнитивного моделирования слабоструктурированных систем.** Для обеспечения программной поддержки когнитивного моделирования при управлении слабоструктурированными системами была разработана информационная технология, впервые предложенная и описанная в работе [18]. С течением времени данная технология совершенствовалась, в нее включались новые математические модели и методы построения и анализа когнитивных моделей. Схема предложенной технологии на текущем уровне ее развития показана на рис. 1.

В рамках этой технологии используется комплекс методов и алгоритмов, обеспечивающих реализацию ее основных этапов. Большинство методов являются авторскими. В табл. 1. приводится описание этапов технологии и реализующих эти этапы методов, со ссылками на источники, содержащими их подробное описание.

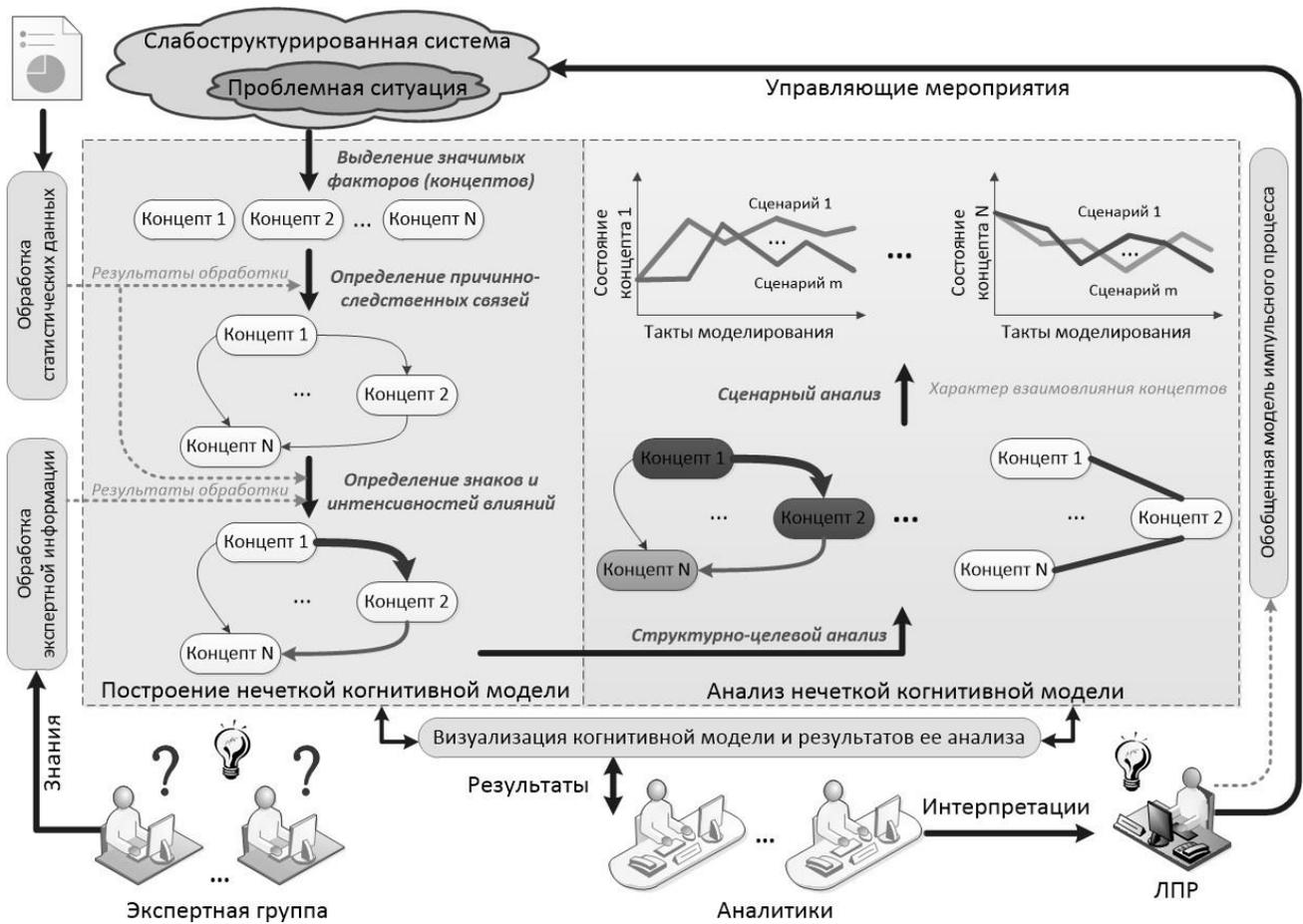


Рис. 1. Информационная технология поддержки когнитивного моделирования слабоструктурированных систем

Таблица 1. Реализация этапов технологии поддержки когнитивного моделирования

Этап когнитивного моделирования	Методы и алгоритмы, обеспечивающие реализацию этапа
Построение НКК	Методы и алгоритмы идентификации структуры и параметров когнитивной карты на основе совместного использования разнородных данных об исследуемой системе [17]
Структурно-целевой анализ НКК	Методы расчета системных показателей НКК и анализа альфа-срезов [13, 22]
Сценарный анализ НКК	Методы и алгоритмы прогнозирования состояния системы при различных воздействиях и поиска управляющих решений по приведению системы в целевое состояние [18, 19, 23]
Визуализация НКК	Методы и алгоритмы визуализации и интерактивного управления визуальным образом НКК [10, 16]

Предложенная информационная технология когнитивного моделирования и комплекс связанных с ней моделей, методов и алгоритмов построения, анализа и визуализации НКК

были реализованы в виде программной системы поддержки принятия решений «ИГЛА» (Интеллектуальный Генератор Лучших Альтернатив) [9].

**3. Опыт применения когнитивного подхода в моделировании и управлении слабоструктурированными социально-экономическими системами.** Разработанная информационная технология поддержки когнитивного моделирования и программная система «ИГЛА» применялись при построении и исследовании нечетких когнитивных моделей для ряда прикладных задач моделирования и управления слабоструктурированными социально-экономическими системами. Рассмотрим некоторые примеры задач.

### **3.1. Стратегическое управление инновационным механизмом предприятия**

Согласно определению [11], инновационный механизм представляет собой «организационно-экономическую форму осуществления инновационной деятельности и способствования ее проведению, поиска инновационных решений, а также рычаг стимулирования и регулирования этой деятельности». В работе [8] было показано, что инновационный механизм относится к числу слабоструктурированных социально-экономических систем, и была построена его нечеткая когнитивная модель. Задачей моделирования являлся поиск управляющих воздействий на инновационный механизм предприятия с целью повышения объема продаж новой продукции и ее рентабельности, сокращения времени разработки новой продукции, а также повышения конкурентоспособности предприятия в целом. Исследование проводилось на примере машиностроительного предприятия ЗАО «Термотрон-Завод» (г. Брянск).

С помощью ведущих специалистов предприятия и экспертов в области управления инновациями был выявлен ряд факторов, оказывающих существенное влияние на инновационную деятельность предприятия и его инновационный потенциал. По итогам согласования мнений экспертов (основной целью которого являлось устранение двусмысленностей и повторений, а также согласование используемой экспертами терминологии) был составлен список из 19 концептов, которые, по характеру и уровню влияния, могут быть разделены на 4 группы: персонал, маркетинг, НИОКР, предприятие в целом. Для определения связей между концептами и задания их интенсивности использовались экспертные методы. Построенная когнитивная карта показана на рис. 2. Сплошные дуги соответствуют положительным влияниям, пунктирные – отрицательным.

По результатам *структурно-целевого анализа* когнитивной модели были выявлены концепты, в наибольшей степени оказывающие влияние на систему. В терминах предметной области, на инновационный потенциал предприятия положительно влияют факторы, которым соответствуют концепты «Интенсивность мотивационных мер», «Уровень предварительных маркетинговых исследований», «Усилия по продвижению новой продукции на рынок». Именно эти концепты целесообразно выбирать в качестве управляемых, поскольку воздействие руководства предприятия на соответствующие факторы может дать наибольший эффект для роста его инновационного потенциала.

Наибольшему положительному влиянию со стороны системы подвержены концепты «Горизонт планирования стратегии развития предприятия», «Инвестиционная привлекательность», «Конкурентоспособность предприятия», «Резерв ресурсов», «Рентабельность новой продукции». Исходя из этого, с достаточно высокой достоверностью можно предположить, что путем контроля состояния соответствующих факторов можно предотвратить отрицательное воздействие на систему извне. При этом, если руководство

предприятия намерено оказывать долговременное управляющее воздействие на эти факторы, ему следует делать акцент на влияние опосредованное, через систему, путем воздействия на концепты, выбранные в качестве управляемых. В частности, на рентабельность новой продукции невозможно влиять напрямую, однако данный показатель зависит от таких факторов, как маркетинговые усилия, мотивация персонала и т.п.

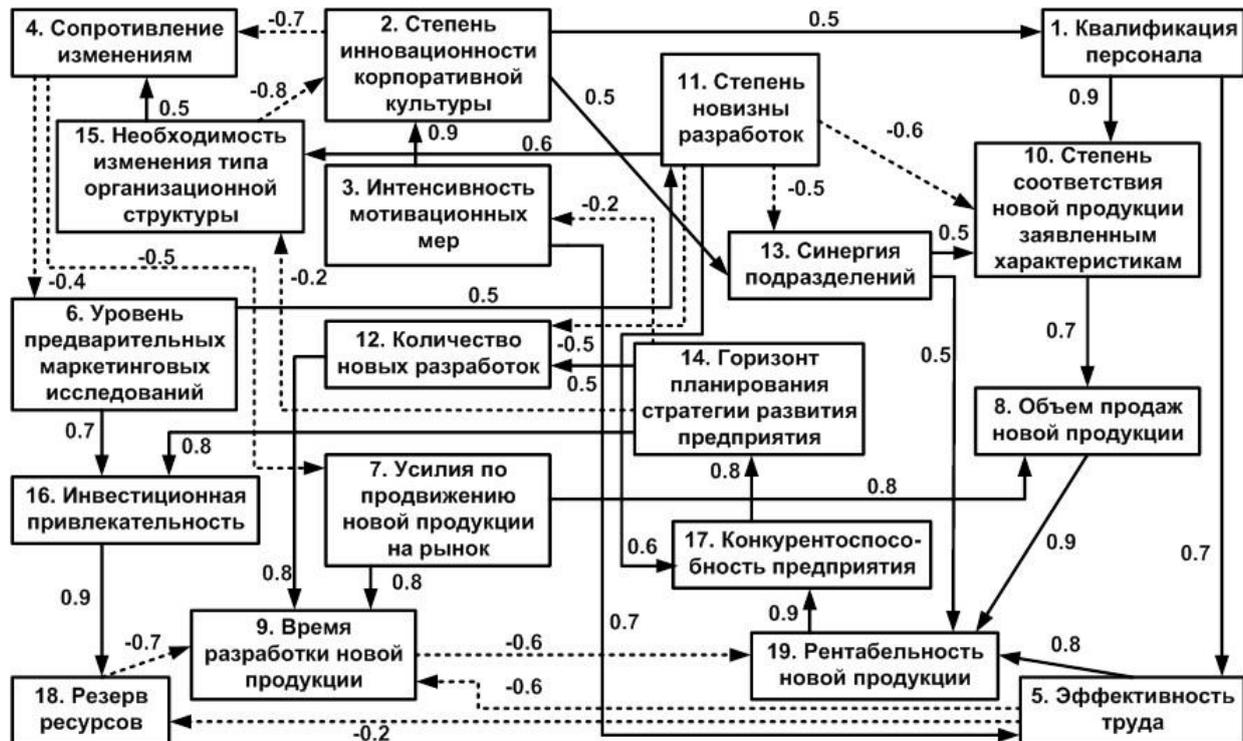


Рис. 2. Нечеткая когнитивная карта инновационного механизма предприятия (источник: [8])

Аналогично, в результате выявления концептов, оказывающих отрицательное влияние на систему в целом, а также концептов, подверженных отрицательному влиянию со стороны системы, был также сделан ряд выводов. В частности, ускорению инновационных процессов на предприятии будут способствовать такие меры, как противодействие росту сопротивления изменениям, сокращение сроков разработки новой продукции, а также постоянный мониторинг технологических и рыночных новшеств.

Кроме того, было обнаружено, что концепты «Степень инновационности корпоративной культуры» и «Время разработки новой продукции» присутствуют одновременно в узлах положительных и отрицательных взаимовлияний. Это позволяет предположить, что управляющее воздействие со стороны руководства на данные факторы целесообразно, но требует осторожности, поскольку ошибочные решения могут привести к ощутимым негативным последствиям для инновационного механизма предприятия в целом.

В ходе сценарного анализа когнитивной модели моделировались различные стратегии управления инновационным механизмом, и исследовалась динамика состояния целевых показателей при тех или иных управляющих воздействиях. В качестве целевых были выбраны концепты «Объем продаж новой продукции», «Время разработки новой продукции», «Степень соответствия новой продукции заявленным характеристикам»,

«Конкурентоспособность предприятия», «Рентабельность новой продукции». Моделировались следующие сценарии:

- оказание управляющего воздействия на один управляемый концепт (соответствует ситуации, когда руководство предприятия вкладывает основной ресурс только в один выбранный фактор, например, только в стимулирование мотивации сотрудников либо только в проведение предварительных маркетинговых исследований);
- оказание комплексного воздействия на некоторое подмножество либо на все управляемые концепты.

Пример представления результатов сценарного анализа показан на рис. 3. Ось абсцисс соответствует условному модельному времени (деления соответствуют тактам моделирования). Каждый график соответствует некоторой альтернативе, т.е. вектору управляющих воздействий на концепты, выбранные в качестве управляемых.

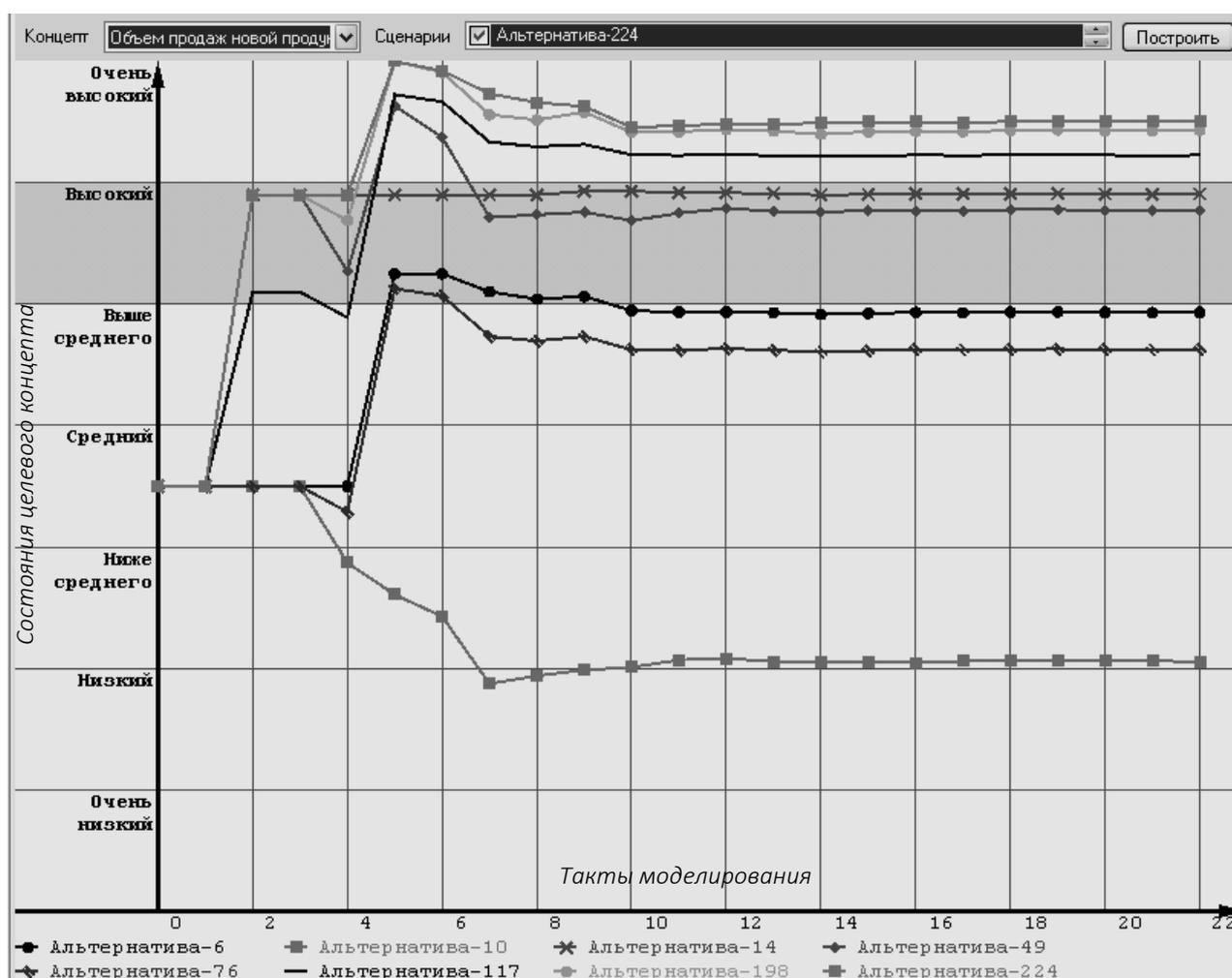


Рис. 3. Динамика изменения состояния целевого концепта «Объем продаж новой продукции» при различных воздействиях на управляемые концепты (источник: [8])

В результате сценарного анализа удалось выявить наиболее эффективные стратегии управления инновационным механизмом предприятия, которые способны в динамике обеспечить желаемые состояния большинства целевых факторов. Вместе с тем, по причине противоречивости поставленной цели (повышение объема продаж новой продукции, ее

рентабельности и степени соответствия заявленным характеристикам, при одновременном сокращении времени ее разработки), применение каждого сценария обеспечивает желаемые результаты лишь для некоторых целевых концептов. Например, усиление мотивационных мер наилучшим образом влияет на повышение объема продаж и рентабельности новой продукции, но в части сокращения сроков разработки новой продукции данная стратегия уступает такому решению, как увеличение объема предварительных маркетинговых исследований. Также результаты моделирования комплексных сценариев показали, что при попытке одновременно воздействовать на все управляемые факторы происходит распыление ресурсов, и результаты оказываются посредственными для всех целевых показателей.

**3.2. Моделирование стратегий управления комплексным развитием сельских территорий.** Исследование проводилось авторами совместно со специалистами ФГБНУ «Федеральный Ростовский аграрный научный центр», которые выступали в роли заказчиков исследования и экспертов соответствующей предметной области. Целью исследования являлось прогнозирование развития сельских территорий при различных начальных тенденциях и управляющих воздействиях, а также поиск эффективных управленческих решений для обеспечения их устойчивого развития [21].

Определение списка ключевых факторов, влияющих на развитие сельских территорий, выполнялось с привлечением группы экспертов, в результате чего были отобраны 11 наиболее значимых факторов взятых за основу для формирования множества концептов когнитивной модели. С предметной точки зрения, эти концепты можно разделить на 4 блока:

*Институциональный блок:*

- 1) развитие рыночной инфраструктуры (налоговая, кредитная, бюджетная, инновационная политика);
- 2) развитие сельского самоуправления.

*Социально-демографический блок:*

- 1) среднегодовая численность населения;
- 2) уровень безработицы;
- 3) развитие социальной сферы.

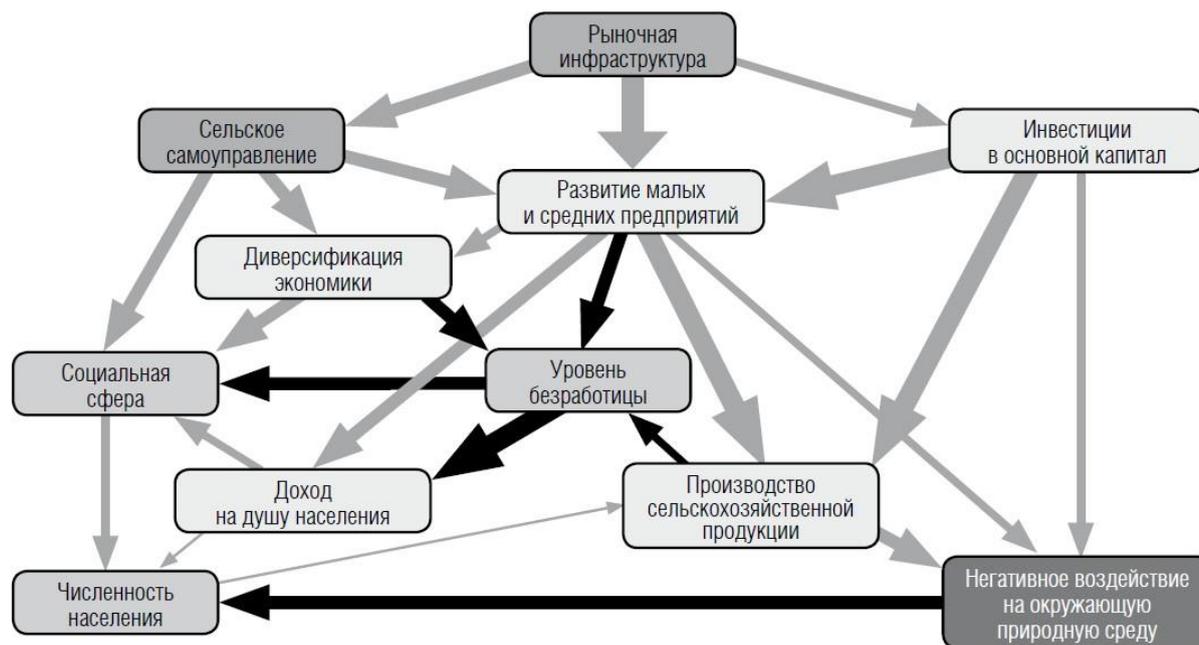
*Экономический блок:*

- 1) доходы на душу населения;
- 2) производство сельскохозяйственной продукции;
- 3) развитие малого и среднего предпринимательства;
- 4) инвестиции в основной капитал;
- 5) уровень диверсификации экономики.

*Экологический блок*

- 1) негативное воздействие на окружающую природную среду.

При построении когнитивной карты авторами была впервые применена новая авторская методика идентификации структуры и параметров нечеткой когнитивной модели, основанная на совместном использовании экспертной и статистической информации [17, 20]. В качестве последней использовались данные Росстата. Построенная когнитивная карта показана на рис 4. Толщина дуг пропорциональна интенсивности влияний между концептами, при этом светлый оттенок дуги соответствует положительному влиянию, а темный – отрицательному.



**Рис. 4.** Нечеткая когнитивная карта для модели управления комплексным развитием сельских территорий (источник: [20])

В результате структурно-целевого анализа построенной когнитивной карты было выявлено, что наиболее существенное положительное влияние на развитие сельских территорий оказывают такие факторы, как «Развитие рыночной инфраструктуры», «Развитие сельского самоуправления» и «Инвестиции в основной капитал». Таким образом, управляющие воздействия следует в первую очередь оказывать на эти факторы, т.е. выбирать соответствующие концепты в качестве управляемых. В свою очередь, наибольшее отрицательное влияние на систему оказывает фактор «Уровень безработицы», при этом он также влияет на целевые факторы «Доход на душу населения» и «Развитие социальной сферы», что свидетельствует о необходимости его контроля при моделировании управляющих воздействий.

Также было замечено, что значение консонанса влияния концептов на систему в большинстве случаев превышало 0,9, что свидетельствует о высоком уровне доверия к знаку и интенсивности указанных влияний.

Было сформулировано желаемое целевое состояние моделируемой системы:

«Проживание в сельских населенных пунктах должно обеспечивать гражданам получение стабильно высокого уровня доходов и доступность социальных благ» [21]. Соответственно, в качестве целевых были выбраны концепты «Доход на душу населения» и «Развитие социальной сферы». В рамках сценарного анализа моделировались сценарий саморазвития системы (т.е. при отсутствии вмешательства в протекающие в ней процессы), а также ряд сценариев ее управляемого развития. На рис. 5 показана динамика изменения состояния целевого концепта «Доход на душу населения» в результате реализации различных сценариев.

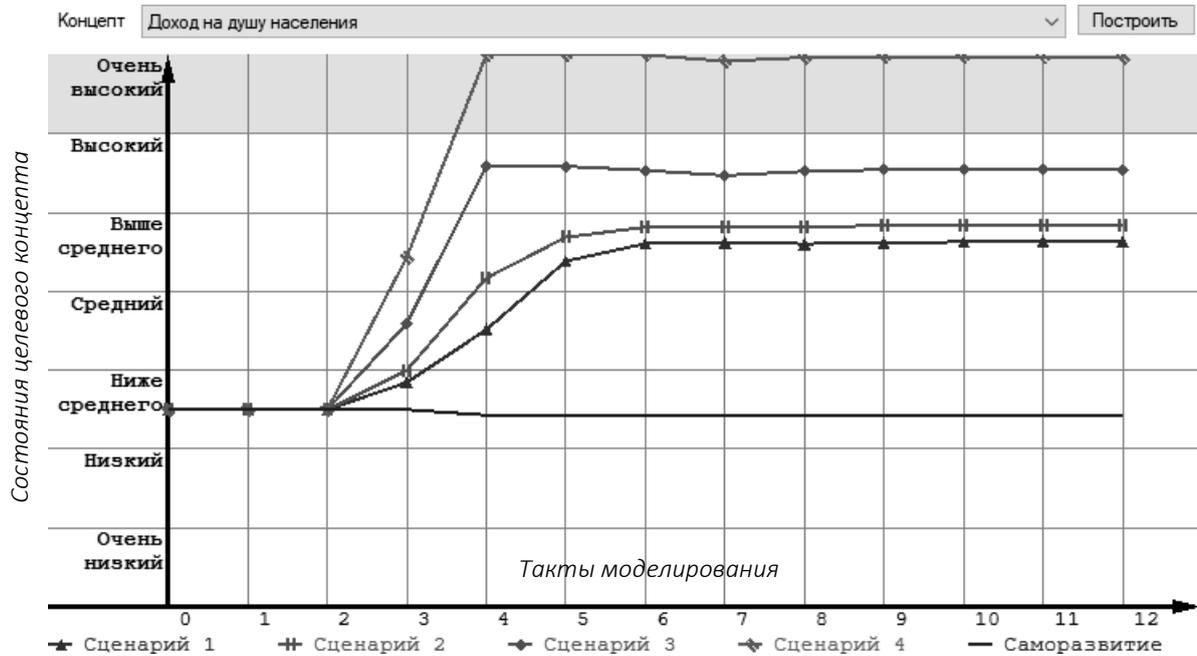


Рис. 5. Динамика состояний концепта «Доход на душу населения» при реализации различных сценариев (источник: [21])

Результаты сценарного анализа показали, что при отсутствии управляющих воздействий (сценарий саморазвития) основные тенденции, характерные для социально-экономических процессов сельских территорий в настоящее время, сохраняются и в последующие периоды времени. При реализации «умеренно-оптимистических» сценариев, когда воздействие оказывается только на один из трех управляемых концептов, динамика развития целевых факторов становится положительной, однако достичь целевых показателей не удастся даже в отдаленной перспективе. Наконец, «оптимистический» сценарий, когда управляющее воздействие оказывается одновременно на все управляемые концепты, не только обеспечивает положительную динамику развития целевых факторов, но и позволяет достичь их желаемого уровня. Вместе с тем, при реализации данного сценария угрозой развитию системы является ухудшение экологической обстановки, что потребует дополнительных решений по предотвращению негативного влияния развития производства на окружающую среду.

**3.3. Анализ и планирование программных проектов.** В качестве исходных данных для построения когнитивной модели программного проекта была взята системная модель «Пирамида программного проекта», описанная в работе [6]. Была рассмотрена одна из граней пирамиды, и на ее основе построена нечеткая когнитивная модель, учитывающая ресурсы проекта и точку зрения правообладателей [19]. На рис. 6 показано представление результатов структурно-целевого анализа данной модели в системе «ИГЛА». Степень насыщенности заливки концептов определяет влияние соответствующих факторов на систему. Так, наибольшее влияние на результаты реализации проекта оказывают такие факторы, как «Объем требований заказчика» и «Сложность технического задания». Результаты структурно-целевого анализа послужили исходными данными для поиска и исследования стратегий управления с помощью сценарного анализа. В качестве целевых

были выбраны концепты «Ценность ПП с точки зрения пользователей» и «Количество доработок на стадии внедрения».



Рис. 6. Представление когнитивной карты анализа и планирования программных проектов в системе «ИГЛА» (источник когнитивной карты: [19])

Для данного исследования была впервые использована авторская модифицированная модель сценарного анализа [23], что позволило повысить гибкость сценарного анализа и достоверность получаемых на его основе результатов. В частности, количество найденных эффективных сценариев развития моделируемой системы увеличилось в 3 раза, по сравнению с использованием стандартной модели импульсного процесса. Результаты динамического моделирования некоторых сценариев показаны на рис. 7.

По итогам сценарного анализа был сделан вывод, что если в момент описания предметной области программного проекта определены все источники требований (выделены заинтересованные лица, операционное окружение и организационная среда программного проекта), извлечены и проанализированы все полученные требования, обнаружены и разрешены конфликты между ними, все разделы технического задания полностью описаны и соответствуют требованиям стандартов, а также подготовлен и согласован технический проект, то можно ожидать следующие результаты [19]:

- на этапе внедрения программного продукта придется осуществлять лишь незначительные его доработки вследствие низкой вероятности появления изменений в требованиях;
- хорошо проработанные и полностью описанные требования позволят упростить процессы разработки и тестирования программного продукта, обеспечить его высокое качество, повысить ценность продукта с точки зрения пользователей (в том числе в соотношении с его стоимостью), а также избежать ситуаций, связанных как с невыполнением каких-либо требований, так и с реализацией избыточных функций.

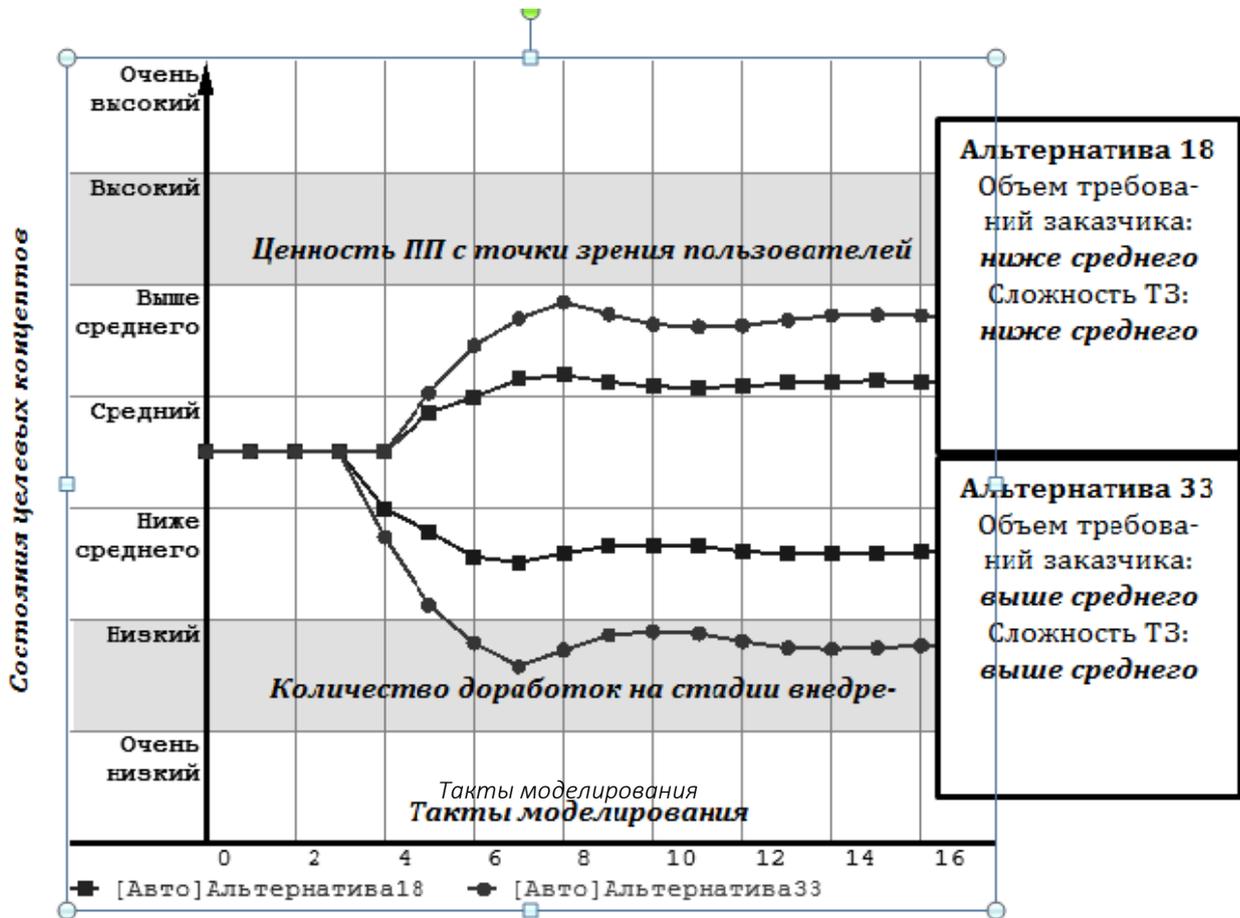


Рис. 7. Некоторые результаты сценарного анализа нечеткой когнитивной модели программного проекта (источник: [19])

Предложенная модель позволяет уже на начальном этапе разработки программного проекта оценить тенденции, связанные с выбором и реализацией того или иного решения, и выявить факторы, оказывающие наибольшее влияние на результаты. В частности, применение данной модели может быть актуальным при анализе портфелей проектов, поскольку позволяет снизить риски старта «неблагоприятных» проектов, а также на этапе планирования работ проекта, т.к. позволяет спрогнозировать их ход.

Помимо этого, разработанная информационная технология поддержки когнитивного моделирования и программная система «ИГЛА» в разное время применялись при решении следующих задач анализа и моделирования слабоструктурированных организационно-технических и социально-экономических систем:

- Моделирование стратегий управления научным потенциалом региона (на примере Брянской области) [2]. Построена когнитивная модель управления научным потенциалом, на основе которой выполнено прогнозирование региональной потребности в кадрах высшей научной квалификации, а также анализ и моделирование стратегий повышения их численности.
- Управление качеством в сварочном производстве [24]. Построена когнитивная модель процесса сварочного производства (при участии специалистов ФГУП ЦНИИ КМ «Прометей»), на основе которой выполнено моделирование стратегий выбора прочностных характеристик металла шва для совершенствования состава покрытия

сварочных электродов. При построении когнитивной модели принимался во внимание тот факт, что качество и прочность сварного соединения определяются не только параметрами сварочного материала, но и целым рядом других факторов, среди которых общетехнический уровень сварочного производства, средняя заработная плата сварщиков, качество проектной и нормативной документации, уровень аттестации, сертификации и контроля качества. Использование результатов моделирования позволило сократить объем экспериментов в 5-8 раз за счет успешного прогнозирования результатов испытаний.

- Моделирование трудноформализуемых задач организации производства [13]. Были построены и исследованы нечеткие когнитивные модели принятия концептуальных решений на ранних стадиях проектирования оснастки: оценка целесообразности использования модульных пресс-форм и выбор технологии изготовления штампованного изделия.

Известны и другие примеры применения системы «ИГЛА» для моделирования социально-экономических систем, в частности:

- управление инновационной деятельностью вуза (на примере Брянского государственного технического университета) [7];
- моделирование стратегий управления экономическим ростом сельскохозяйственной отрасли [3, 4, 12];
- моделирование тенденций и закономерностей трудовой деятельности в России [15].

**Заключение.** В статье представлена информационная технология поддержки когнитивного моделирования слабоструктурированных систем, основанная на использовании нечетких когнитивных карт Силова, хорошо зарекомендовавших себя в практических приложениях. Описан опыт применения указанной технологии и входящих в ее состав методов и алгоритмов построения и анализа нечетких когнитивных моделей при решении практических задач исследования стратегий управления социально-экономическими системами, среди которых:

- стратегическое управление инновационным механизмом предприятия;
- моделирование стратегий управления развитием сельских территорий;
- анализ и планирование программных проектов.

Применение разработанной информационной технологии поддержки когнитивного моделирования и программной системы «ИГЛА», реализующей данную технологию, в этих и ряде других задач исследования стратегий управления социально-экономическими системами позволило: сократить трудоемкость разработки и обоснования решений, повысить уровень их оперативности при одновременном снижении объема используемых вычислительных ресурсов, а также обеспечить возможность нахождения неочевидных для аналитиков стратегий и сценариев управления.

В целом применение когнитивного подхода в моделировании слабоструктурированных социально-экономических систем и управлении такими системами позволяет обнаруживать скрытые взаимосвязи между факторами, положительно и отрицательно влияющими на динамику функционирования и развития таких систем, выявлять факторы, не оказывающие какого-либо заметного влияния, а также находить точки

наиболее эффективного приложения управляющих воздействий. Это позволяет уже на ранних стадиях анализа систем оценивать последствия тех или иных управленческих решений и возможные риски реализации различных стратегий управления.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, проект № 19-07-00844.

#### СПИСОК ЛИТЕРАТУРЫ

1. Авдеева З.К., Коврига С.В., Макаренко Д.И. Когнитивное моделирование для решения задач управления слабоструктурированными системами (ситуациями) // Управление большими системами. 2007. Вып. 16. С. 26-39.
2. Аверченков В.И., Кожухар В.М., Подвесовский А.Г., Сазонова А.С. Мониторинг и прогнозирование региональной потребности в специалистах высшей научной квалификации. Брянск: БГТУ. 2010. 163 с.
3. Анохина М.Е. Моделирование стратегии управления экономическим ростом сельского хозяйства // Экономика сельскохозяйственных и перерабатывающих предприятий. 2019. Часть 1. № 2. С. 23-34.
4. Анохина М.Е. Моделирование стратегии управления экономическим ростом сельского хозяйства. Часть 2 // Экономика сельскохозяйственных и перерабатывающих предприятий. 2019. № 3. С. 50-53.
5. Борисов В.В., Круглов В.В., Федулов А.С. Нечеткие модели и сети. М.: Горячая линия -Телеком. 2012. 284 с.
6. Гвоздев В.Е., Ильясов Б.Г. Пирамида программного проекта // Программная инженерия. 2011. № 1. С. 16-24.
7. Горленко О.А., Лагерев Д.Г., Мирошников В.В., Чистоклетов Н.Ю. Методологические основы управления инновационной деятельностью вуза на основе когнитивных моделей. Брянск: БГТУ. 2009. 127 с.
8. Ерохин Д.В., Лагерев Д.Г., Ларичева Е.А., Подвесовский А.Г. Стратегическое управление инновационной деятельностью предприятия. Брянск: БГТУ. 2010. 196 с.
9. Захарова А.А., Подвесовский А.Г., Исаев Р.А. Математическое и программное обеспечение поддержки когнитивного моделирования слабоструктурированных организационно-технических систем // Междунар. конф. СРТ2019. Нижний Новгород: Изд-во ННГАСУ и НИЦФТИ. 2019. С. 131-141.
10. Захарова А.А., Подвесовский А.Г., Лагерев Д.Г. Визуальная аналитика и когнитивные методы для обработки и анализа гетерогенных данных мультисенсорных систем: проблемы и тенденции // Информационные и математические технологии в науке и управлении. 2019. № 4 (16). С. 60-74.
11. Колоколов В.А. Инновационные механизмы функционирования предпринимательских структур // Менеджмент в России и за рубежом. 2002. № 1. С. 95- 104.
12. Кондрашина О.Н., Анохина М.Е. Использование нечетких когнитивных карт в оценке качества экономического роста отдельной отрасли // Экономика и предпринимательство. 2017. № 5-1. С. 896-899.

13. Копелиович Д.И., Подвесовский А.Г., Сафонов А.Л., Вилюха А.В., Исаев Р.А. Применение нечетких когнитивных моделей в автоматизации проектирования технологической оснастки // Вестник компьютерных и информационных технологий. 2018. № 3. С.20-35.
14. Кулинич А.А. Ситуационный, когнитивный и семиотический подходы к принятию решений в организациях // Открытое образование. 2016. Т. 20. № 6. С. 9-17.
15. Мельник М.С., Орехов В.Д., Причина О.С. Моделирование тенденций и закономерностей трудовой деятельности в России: когнитивный подход // Проблемы экономики и юридической практики. 2018. № 3. С. 94-101.
16. Подвесовский А.Г., Исаев Р.А. Метафоры визуализации нечетких когнитивных карт // Научная визуализация. 2018. Т. 10. № 4. С. 13-29.
17. Подвесовский А.Г., Исаев Р.А. Идентификация структуры и параметров нечетких когнитивных моделей: экспертные и статистические методы // International Journal of Open Information Technologies. 2019. Т. 7. № 6. С. 35-61.
18. Подвесовский А.Г., Лагерев Д.Г., Коростелев Д.А. Применение нечетких когнитивных моделей для формирования множества альтернатив в задачах принятия решений // Вестник Брянского государственного технического университета. 2009. № 4 (24). С. 77-84.
19. Подвесовский А.Г., Титарев Д.В., Исаев Р.А. Нечеткие когнитивные модели в задачах анализа и планирования программных проектов // Вестник компьютерных и информационных технологий. 2019. № 8. С. 22-31.
20. Подгорская С.В., Подвесовский А.Г., Исаев Р.А., Антонова Н.И. Построение нечетких когнитивных моделей социально-экономических систем на примере модели управления комплексным развитием сельских территорий // Бизнес-информатика. 2019. Т. 13. № 3. С. 7–19.
21. Подгорская С.В., Подвесовский А.Г., Исаев Р.А., Тарасов А.С., Бахматова Г.А. Моделирование сценарного развития сельских территорий на основе нечеткой когнитивной модели // Проблемы управления. 2019. № 5. С. 49-59.
22. Силов В.Б. Принятие стратегических решений в нечеткой обстановке. М.: ИНПРО-РЕС. 1995. 228 с.
23. Isaev R.A., Podvesovskii A.G. Generalized Model of Pulse Process for Dynamic Analysis of Sylov's Fuzzy Cognitive Maps // CEUR Workshop Proceedings of the Mathematical Modeling Session at the International Conference Information Technology and Nanotechnology (MM-ITNT 2017). Vol. 1904. Pp. 57-63.
24. Podvesovskii A.G., Gulakov K.V. , Dergachyov K.V., Korostelyov D.A., Lagerev D.G. The choice of parameters of welding materials on the basis of fuzzy cognitive model with neural network identification of nonlinear dependence // Proceedings of 2015 International Conference on Mechanical Engineering, Automation and Control Systems (MEACS). IEEE Catalog Number: CFP1561Y-ART. Pp. 02-38-NSAP.

## FUZZY COGNITIVE MODELS IN MANAGEMENT OF SEMI-STRUCTURED SOCIO-ECONOMIC SYSTEMS

**Alena A. Zakharova**

Dr. Tech. Sc., Professor of Informatics and Software Engineering Department

e-mail: [zaa@tu-bryansk.ru](mailto:zaa@tu-bryansk.ru)

ORCID: 0000-0003-4221-7710,

**Aleksandr G. Podvesovskii**

Cand. Tech. Sc., Head of Informatics and Software Engineering Department

e-mail: [apodv@tu-bryansk.ru](mailto:apodv@tu-bryansk.ru)

ORCID: 0000-0002-1118-3266,

**Ruslan A. Isaev**

Assistant Lecturer of Informatics and Software Engineering Department,

e-mail: [ruslan-isaev-32@yandex.ru](mailto:ruslan-isaev-32@yandex.ru)

ORCID: 0000-0003-3263-4051,

Bryansk State Technical University

50 let Oktyabrya blvd., 241035, Bryansk, Russia

**Abstract.** The article deals with the problems of modeling semi-structured socio-economic systems based on the use of the cognitive approach. The original information technology for supporting the cognitive modeling of such systems based on the use of fuzzy cognitive maps is proposed. The experience of using the proposed technology and the software system that implements it in solving a number of practical problems of researching strategies for managing socio-economic systems is described.

**Keywords:** management of socio-economic systems, semi-structured system, cognitive model, fuzzy cognitive map, structure-and-target analysis, scenario analysis.

### References

1. Avdeeva Z.K., Kovriga S.V., Makarenko D.I. Kognitivnoe modelirovanie dlja reshenija zadach upravlenija slabostrukturirovannymi sistemami (situacijami) [Cognitive Modeling for Solving Problems of Control of Semi-structured Systems (Situations)] // Upravlenie bol'shimi sistemami = Large Systems Control. 2007. Issue 16. Pp. 26-39 (in Russian).
2. Averchenkov V.I., Kozhukhar V.M., Podvesovskii A.G., Sazonova A.S. Monitoring i prognozirovanie regional'noj potrebnosti v specialistah vysshej nauchnoj kvalifikacii [Monitoring and Forecasting of Regional Demand for Specialists of the Higher Scientific Qualification] Bryansk: Bryanskiy gosudarstvennyy universitet=Bryansk: Bryansk State Technical University. 2010. 163 p. (in Russian).
3. Anokhina M.E. Modelirovanie strategii upravlenija jekonomicheskim rostom sel'skogo hozjajstva [Modeling Strategy for Management of Economic Growth of Agriculture] // Jekonomika sel'skohozjajstvennyh i pererabatyvajushhih predpriyatij = Economy of agricultural and processing enterprises. 2019. Part 1. № 2. Pp. 23-34 (in Russian).

4. Anokhina M.E. Modelirovanie strategii upravlenija jekonomicheskim rostom sel'skogo hozjajstva [Modeling Strategy for Management of Economic Growth of Agriculture] // Jekonomika sel'skohozjajstvennyh i pererabatyvajushhijh predpriyatij = Economy of agricultural and processing enterprises. 2019. Part 2. № 3. Pp. 50-53 (in Russian).
5. Borisov V.V., Kruglov V.V., Fedulov A.S. Nechetkie modeli i seti [Fuzzy Models and Networks]. M.: Goryachaya Liniya–Telekom = Hot Line–Telecom. 2012. 284 p. (in Russian).
6. Gvozdev V.E., Ilyasov B.G. Piramida programmnoho projekta [Software project pyramid] // Programmaja inzhenerija = Software Engineering. 2011. № 1. Pp. 16-24 (in Russian).
7. Gorlenko O.A., Lagerev D.G., Miroshnikov V.V., Chistokletov N.Yu. Metodologicheskie osnovy upravlenija innovacionnoj dejatel'nost'ju vuza na osnove kognitivnyh modelej [Methodological foundations of university innovation management based on cognitive models]. Bryansk: Bryanskiy gosudarstvennyy universitet=Bryansk: Bryansk State Technical University. 2009. 127 p. (in Russian).
8. Erokhin D.V., Lagerev D.G., Laricheva E.A., Podvesovskii A.G. Strategicheskoe upravlenie innovacionnoj dejatel'nost'ju predpriyatija [Strategic Management of Innovation Activity of Enterprise]. Bryansk: Bryanskiy gosudarstvennyy universitet=Bryansk: Bryansk State Technical University. 2010. 196 p. (in Russian).
9. Zakharova A.A., Podvesovskii A.G., Isaev R.A. Matematicheskoe i programmnoe obespechenie podderzhki kognitivnogo modelirovanija slabostrukturirovannyh organizacionno-tehnicheskijh system [Mathematical and Software Support for Cognitive Modeling of Semi-structured Organizational and Technical Systems]. In: CPT2019 International Conference Proceedings. Pub. NNGASU and SRCIPT. Nizhniy Novgorod. 2019. Pp. 131-141 (in Russian).
10. Zakharova A.A., Podvesovskii A.G., Lagerev D.G. Vizual'naja analitika i kognitivnye metody dlja obrabotki i analiza geterogennyh dannyh mul'tisensornyh sistem: problemy i tendencii [Visual analytics and cognitive methods for processing and analysis of heterogeneous data in multi-sensor systems: issues and trends] // Informacionnye i matematicheskie tehnologii v nauke i upravlenii = Information and mathematical technologies in science and management. 2019. № 4 (16). Pp. 60-74 (in Russian).
11. Kolokolov V.A. Innovacionnye mehanizmy funkcionirovanija predprinimatel'skijh struktur [Innovative mechanisms for the functioning of entrepreneurial structures] // Menedzhment v Rossii i za rubezhom = Management in Rissia and Abroad. 2002. № 1. Pp. 95-104 (in Russian).
12. Kondrashina O.N., Anokhina M.E. Ispol'zovanie nechetkijh kognitivnyh kart v ocnke kachestva jekonomicheskogo rosta otdel'noj otrasli [Using of Fuzzy Cognitive Maps in Quality Assessment of Economic Growth of Separate Sector] // Jekonomika i predprinimatel'stvo = Journal of Economy and entrepreneurship. 2017. № 5-1. Pp. 896-899 (in Russian).
13. Kopeliovich D.I., Podvesovskii A.G., Safonov A.L., Vilyukha A.V., Isaev R.A. Primenenie nechetkijh kognitivnyh modelej v avtomatizacii proektirovaniya tekhnologicheskoih osnastki [Application of Fuzzy Cognitive Models in Computer-Aided Production Tooling Design] // Vestnik komp'yuternyh i informacionnyh tekhnologij = Herald of Computer and Information Technologies. 2018. № 3. Pp. 20-35 (in Russian).
14. Kulinich A.A. Situacionnyj, kognitivnyj i semioticheskij podhody k prinjatiju reshenij v organizacijah [Contingency, cognitive and semiotic approaches to decision-making in the

- organizations] // Otkrytoe obrazovanie = Open Education. 2016. Vol. 20. № 6. Pp. 9-17 (in Russian).
15. Mel'nik M.S., Orekhov V.D., Prichina O.S. Modelirovanie tendencij i zakonomernostej trudovoj dejatel'nosti v Rossii: kognitivnyj podhod [Simulation of Trends and Legislation of Labor Activities in Russia: Cognitive Approach] // Problemy jekonomiki i juridicheskoj praktiki = Economic problems and legal practice. 2018. № 3. Pp. 94-101 (in Russian).
16. Podvesovskii A.G., Isaev R.A. Visualization metaphors for fuzzy cognitive maps // Scientific Visualization. 2018. Vol. 10. № 4. Pp. 13-29.
17. Podvesovskii A.G., Isaev R.A. Identifikacija struktury i parametrov nechetkih kognitivnyh modelej: jekspert-nye i statisticheskie metody [Identification of structure and parameters of fuzzy cognitive models: expert and statistical methods] // International Journal of Open Information Technologies. 2019. Vol. 7. № 6. Pp. 35-61 (in Russian).
18. Podvesovskii A.G., Lagerev D.G., Korostelyov D.A. Primenenie nechetkih kognitivnyh modelej dlja formirovanija mnozhestva al'ternativ v zadachah prinjatija reshenij Применение нечетких когнитивных моделей для формирования множества альтернатив в задачах принятия решений [Application of fuzzy cognitive models to form a set of alternatives in decision-making problems] // Vestnik Brjanskogo gosudarstvennogo tehničeskogo universiteta = Bulletin of Bryansk State Technical University. 2009. № 4 (24). Pp. 77-84 (in Russian).
19. Podvesovskii A.G., Titarev D.V., Isaev R.A. Nechetkie kognitivnye modeli v zadachah analiza i planirovanija programmnyh proektov [Fuzzy cognitive models in software projects analysis and planning] // Vestnik komp'yuternykh i informacionnykh tekhnologii = Herald of computer and information technologies. 2019. № 8. Pp. 22-31 (in Russian).
20. Podgorskaya S.V., Podvesovskii A.G., Isaev R.A., Antonova N.I. Fuzzy cognitive models for socio-economic systems as applied to a management model for integrated development of rural areas // Business Informatics. 2019. Vol. 13. № 3. Pp. 7-19.
21. Podgorskaya S.V., Podvesovskii A.G., Isaev R.A., Tarasov A.S., Bakhmatova G.A. Modelirovanie scenarnogo razvitija sel'skih territorij na osnove nechetkoj kognitivnoj modeli [Modeling of Scenario Development of Rural Territories Based on Fuzzy Cognitive Model] // Problemy upravlenija = Control Sciences. 2019. № 5. Pp. 49-59.
22. Sylov V.B. Prinyatie strategicheskikh reshenij v nechetkoj obstanovke [Strategical Decision Making in Fuzzy Environment]. Moscow: INPRO-RES. 1995. 228 p. (in Russian).
23. Isaev R.A., Podvesovskii A.G. Generalized Model of Pulse Process for Dynamic Analysis of Sylov's Fuzzy Cognitive Maps // CEUR Workshop Proceedings of the Mathematical Modeling Session at the International Conference Information Technology and Nanotechnology (MM-ITNT 2017). Vol. 1904. Pp. 57-63.
24. Podvesovskii A.G., Gulakov K.V., Dergachyov K.V., Korostelyov D.A., Lagerev D.G. The choice of parameters of welding materials on the basis of fuzzy cognitive model with neural network identification of nonlinear dependence // Proceedings of 2015 International Conference on Mechanical Engineering, Automation and Control Systems (MEACS). IEEE Catalog Number: CFP1561Y-ART. Pp. 02-38-NSAP.

УДК 004.8:620.9

**ПРИМЕНЕНИЕ МЕТОДОВ КОГНИТИВНОГО МОДЕЛИРОВАНИЯ  
ДЛЯ АНАЛИЗА УГРОЗ ЭНЕРГЕТИЧЕСКОЙ БЕЗОПАСНОСТИ**

**Массель Алексей Геннадьевич**

к.т.н., с.н.с. отдела «Систем искусственного интеллекта в энергетике»

e-mail: [amassel@isem.irk.ru](mailto:amassel@isem.irk.ru),

**Пяткова Наталья Ивановна**

к.т.н., с.н.с. лаборатории «Надежности топливо- и энергоснабжения»

e-mail: [nata@isem.irk.ru](mailto:nata@isem.irk.ru),

Институт систем энергетики им. Л.А. Мелентьева СО РАН,  
664033 г. Иркутск, ул. Лермонтова 130

**Аннотация.** В статье рассматривается двухуровневая информационная технология исследований проблем энергетической безопасности, где сочетаются уровни качественного и количественного анализа. Более подробно рассматривается использование когнитивных моделей разного уровня. Описан уровень качественного анализа угрозы «Недостаток инвестиций». Описывается структура традиционных оптимизационных моделей топливно-энергетического комплекса. Показана возможность интеграции когнитивных и математических моделей.

**Ключевые слова:** Семантическое моделирование, когнитивное моделирование, энергетическая безопасность.

**Цитирование:** Массель А. Г., Пяткова Н. И. Применение методов когнитивного моделирования для анализа угроз энергетической безопасности // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 24-33. DOI: 10.38028/ESI.2020.20.4.002

**Введение.** В ИСЭМ СО РАН традиционно ведутся исследования проблем энергетической безопасности [8]; в исследованиях применяются линейные экономико-математические модели и традиционные программные комплексы, что доставляет определенные неудобства с точки зрения экспертов: при комбинаторном подходе к многовариантным расчетам на эксперта возлагалась задача выбрать из множества решений только определенные. В связи с этим была предложена двухуровневая технология, интегрирующая этапы качественного анализа (с использованием инструментальных средств семантического моделирования) и количественного анализа (с использованием линейных экономико-математических моделей и традиционных программных комплексов) [4]. Предполагается, что при использовании уровня качественного анализа может быть значительно снижена нагрузка на эксперта. В статье более подробно рассматривается применение методов когнитивного моделирования на уровне качественного анализа.

Понятие и возможности использования когнитивных моделей рассматриваются в [10,12]. Когнитивные карты представляют собой графы, а когнитивное моделирование осуществляется путем представления проблемной ситуации системы в виде наглядного визуального образа, причем использование графических представлений этих моделей можно отнести к области когнитивной графики [2]. Так, в процессе работы с экспертами-энергетиками особенно хорошо зарекомендовали себя когнитивные карты, которые

позволяют графически отобразить причинно-следственные отношения между различными объектами и факторами в определенных ситуациях.

Предпосылкой развития когнитивного подхода является сложность в использовании точных моделей для анализа и моделирования проблемных ситуаций, возникающих в процессе развития социальных и технико-экономических систем, из-за необходимости учета большого числа факторов, многие из которых оказываются трудно формализуемыми.

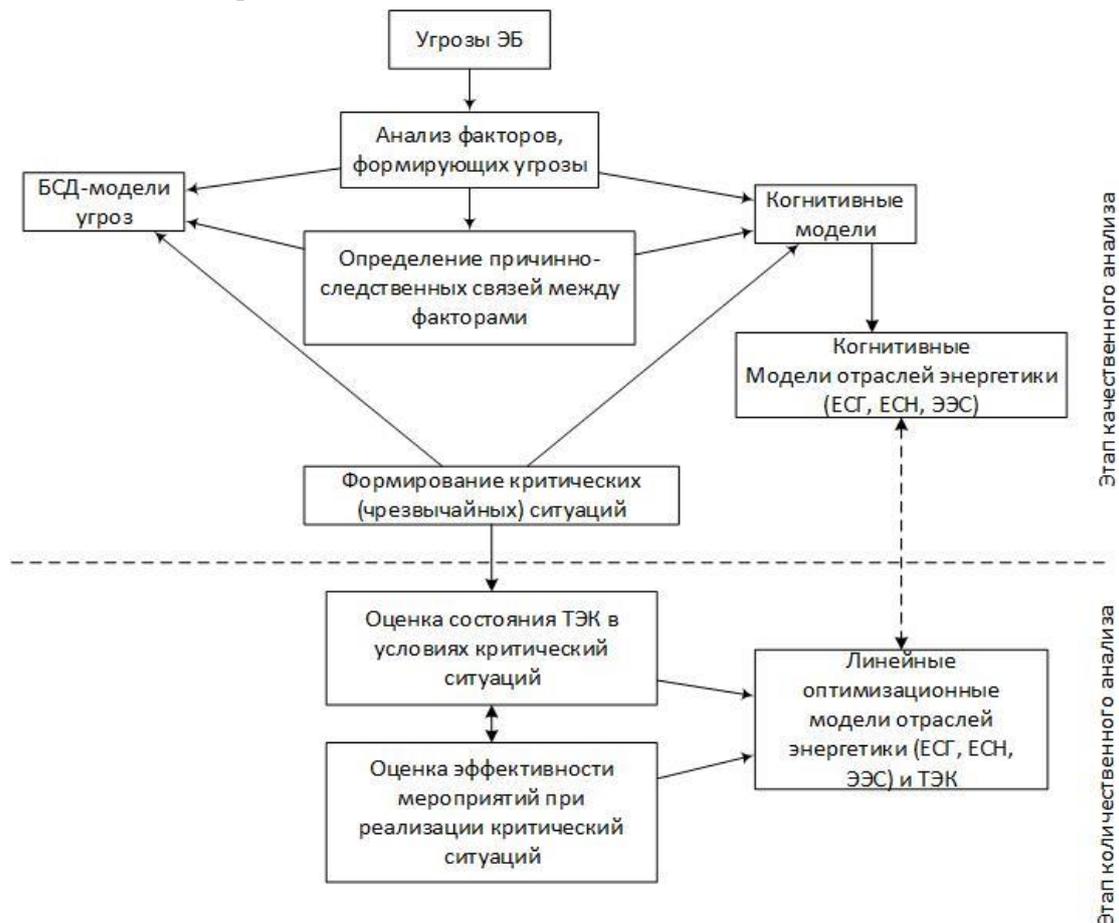
Помимо когнитивных моделей, в исследованиях проблем энергетической безопасности активно используется вероятностное моделирование (на основе Байесовских сетей доверия) [5], но в данной статье это направление не рассматривается

**Двухуровневая технология.** В ходе исследований было предложено перейти к двухуровневой технологии, которая включает этапы качественного и количественного анализа. Более подробно двухуровневая технология рассмотрена в [1, 4]. Иллюстрация интеграции математического и семантического (когнитивного) моделирования приведена на рис. 1.

К задачам качественного анализа относятся следующие этапы:

- Анализ факторов, формирующих угрозы.
- Определение причинно-следственных связей между факторами.
- Формирование критических и чрезвычайных ситуаций (КС и ЧС).

Для отображения и работы с этими этапами используются инструменты вероятностного и когнитивного моделирования.



**Рис. 1.** Решаемые задачи и используемые методы моделирования при исследованиях оценки влияния угроз на состояние энергетической безопасности

В настоящее время ведутся исследования по построению когнитивных моделей отраслей энергетики, что позволит связать их с конкретными линейными оптимизационными моделями отраслей энергетики. Представляется, что на этом этапе необходимо, в дальнейшем, переходить к динамическим когнитивным картам [11].

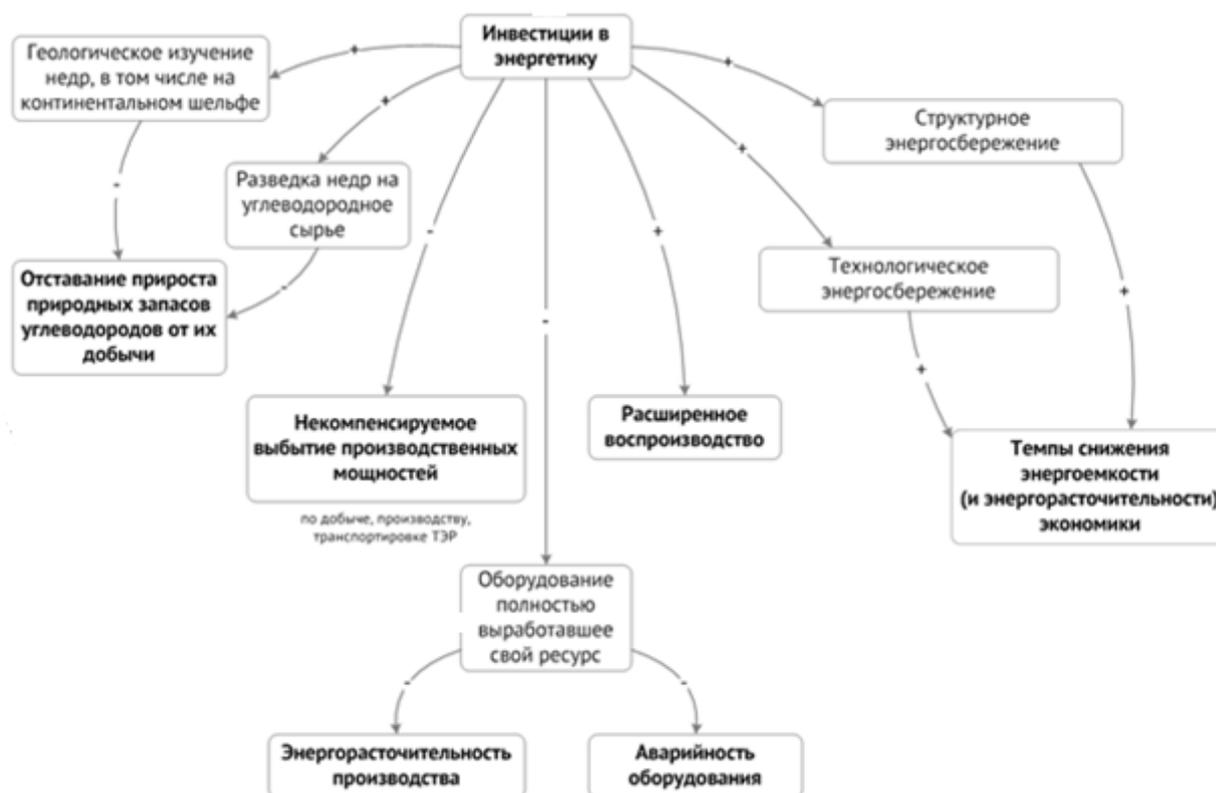
**Применение методов когнитивного моделирования.** В современных условиях наибольшую значимость приобретает проблема изучения угроз энергетической безопасности и факторов, формирующих эти угрозы. При этом возможно обосновать возникновение тех или иных критических или чрезвычайных ситуаций и дать вероятностную оценку их возникновения.

На основе анализа состояния энергетического сектора страны и условий его развития были выделены основные стратегические угрозы энергетической безопасности [9]:

- недостаточный уровень инвестиций в отраслях ТЭК;
- недостаточные приросты разведанных запасов углеводородов;
- снижение возможностей по увеличению объемов добычи газа;
- высокая изношенность и низкие темпы обновления оборудования в отраслях ТЭК.

Анализ этих угроз и формирование КС и ЧС на их основе выполняется на уровне качественного анализа, на котором возможно использование методов семантического моделирования (рис. 1), рассматриваемых в качестве методов анализа и обоснования угроз энергетической безопасности, формирования КС и ЧС и методов ситуационного анализа и управления.

Для одной из важных стратегических угроз "Недостаток инвестиций" построена система моделей ухудшения/улучшения ситуации с использованием методов когнитивного моделирования (рис. 2).



**Рис. 2.** Общая когнитивная модель угрозы «Недостаток инвестиций» (причинно-следственная связь факторов, образующих эту угрозу).

Итогом рассмотренного недофинансирования инвестиций в ТЭК (вместе с отвлечением значительных средств на сооружение экспортных трубопроводов, не имеющих прямого отношения к обеспечению энергетической безопасности России) были, остаются, и могут сохраниться в обозримой перспективе следующие *негативные явления (последствия)* [8, 9].

1. Некомпенсируемое выбытие производственных мощностей по добыче, производству, транспортировке и распределению ТЭР, либо, в других случаях, сохранение в работе оборудования (включая трубопроводы), полностью выработавшего свой ресурс (срок службы) или приближающегося к такому состоянию – низко-экономичного (в том числе энергорасточительного) и физически изношенного, ветхого, с повышенной аварийностью.

2. Использование ограниченных инвестиций преимущественно для обеспечения простого воспроизводства в отраслях ТЭК, без обеспечения расширенного воспроизводства и без необходимой модернизации производственного аппарата энергетики, повышения его экономической и энергетической эффективности.

3. Прогрессирующее снижение, или, по крайней мере, стагнация технического уровня предприятий ТЭК и соответственно ухудшение их технико-экономических показателей, отставание российской энергетики от мирового уровня, как негативный фактор, обусловленный в значительной мере дефицитом инвестиций и его рассмотренными последствиями (пп. 1, 2).

Следующая группа последствий недофинансирования состоит в обусловленности им других угроз ЭБ, в том числе:

4. Угроза энергорасточительности экономики (неприемлемо низких темпов снижения ее энергоемкости) обусловлена, во-первых, в значительной мере недостаточным финансированием мероприятий во всех отраслях народного хозяйства, прямо или косвенно направленных на повышение энергоэффективности используемых в них технологий, оборудования, зданий (финансирование технологического энергосбережения); и, во-вторых, опережающего развития малоэнергоемких (но наукоемких и трудоемких) отраслей материального производства и сферы услуг (структурное энергосбережение).

5. Недостаточными инвестициями в разведку недр на углеводородное сырье, в целом геологическое изучение недр, в том числе на континентальном шельфе, определяется, прежде всего, угрозой отставания прироста разведанных запасов углеводородов от объемов их добычи.

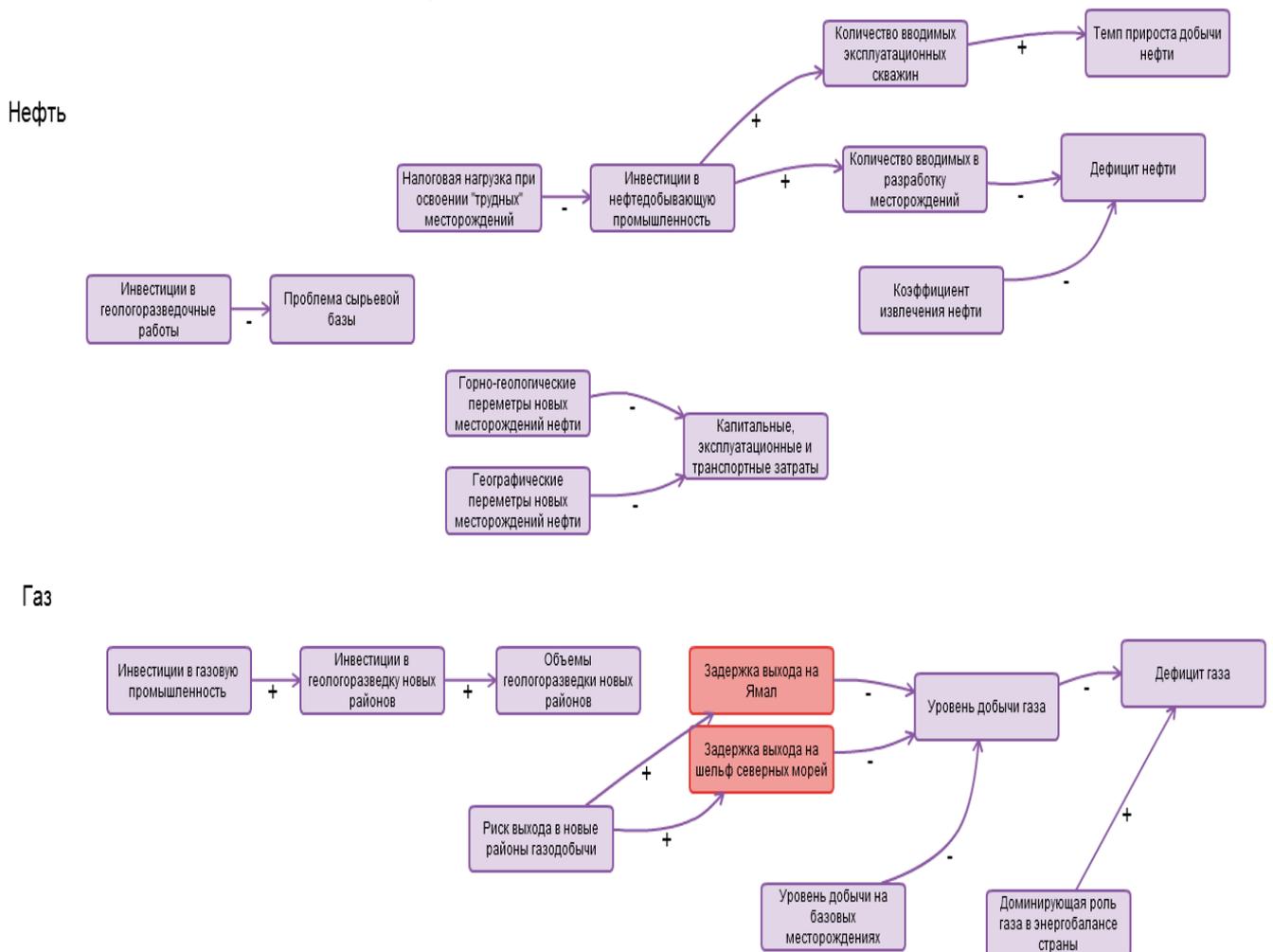
6. Угроза, сформулированная как доминирующая роль природного газа в ТЭБ европейских районов России, в значительной мере обусловлена недостаточными инвестициями в развитие атомной и возобновляемой энергетики, в техническое перевооружение угольной электро- и теплогенерации, в создание эффективной индустрии по производству качественных углепродуктов (включая обогащение энергетических углей) (рис. 3). Для этой угрозы построена когнитивная модель, представленная на рис. 3.

7. Угроза прогрессирующего снижения объемов добычи газа в России определяется не только экономическим риском освоения газовых ресурсов Ямала и шельфа северных морей (это, прежде всего, связано с привлекаемыми иностранными инвестициями), но и недостаточными собственными инвестициями в освоение указанных газовых ресурсов. В нефтяной отрасли состояние с инвестициями достаточно благополучное, но ухудшение состояния горно-геологических характеристик новых месторождений нефти, их качества, увеличения затрат на освоение этих месторождений, снижение коэффициента извлечения нефти на разработанных месторождениях влияет на величину требуемых инвестиций. По этим

причинам на основе рассуждений экспертов построена система когнитивных моделей ухудшения/улучшения ситуации по стратегической угрозе ЭБ «Недостаток инвестиций в в нефтяной и газовой отрасли» (рис. 4).



**Рис. 3.** Когнитивная модель по угрозе «Доминирующая роль природного газа в ТЭБ европейских районов России»



**Рис. 4.** Система моделей ухудшения/улучшения ситуации по стратегической угрозе ЭБ «Недостаток инвестиций в в нефтяной и газовой отрасли».

Помимо общих когнитивных моделей, которые, несомненно, помогают эксперту формировать нештатные ситуации, предлагается строить дополнительно когнитивные модели отраслей энергетики (рис. 5). Основное отличие этих моделей в их приближенности к технико-экономическим моделям, которые используются на этапе количественного анализа. У каждого концепта есть параметры, которые описывают его основные характеристики, и могут учитываться в модели. Таким образом, используя методы семантического моделирования угроз, мы можем обоснованно формировать различные нештатные ситуации и оценивать их на этапе качественного анализа. На рис. 5 представлена когнитивная карта угрозы «Недостаток инвестиций» для газовой отрасли, причем эксперт в своем исследовании может детализировать данную карту для задания более конкретного сценария.

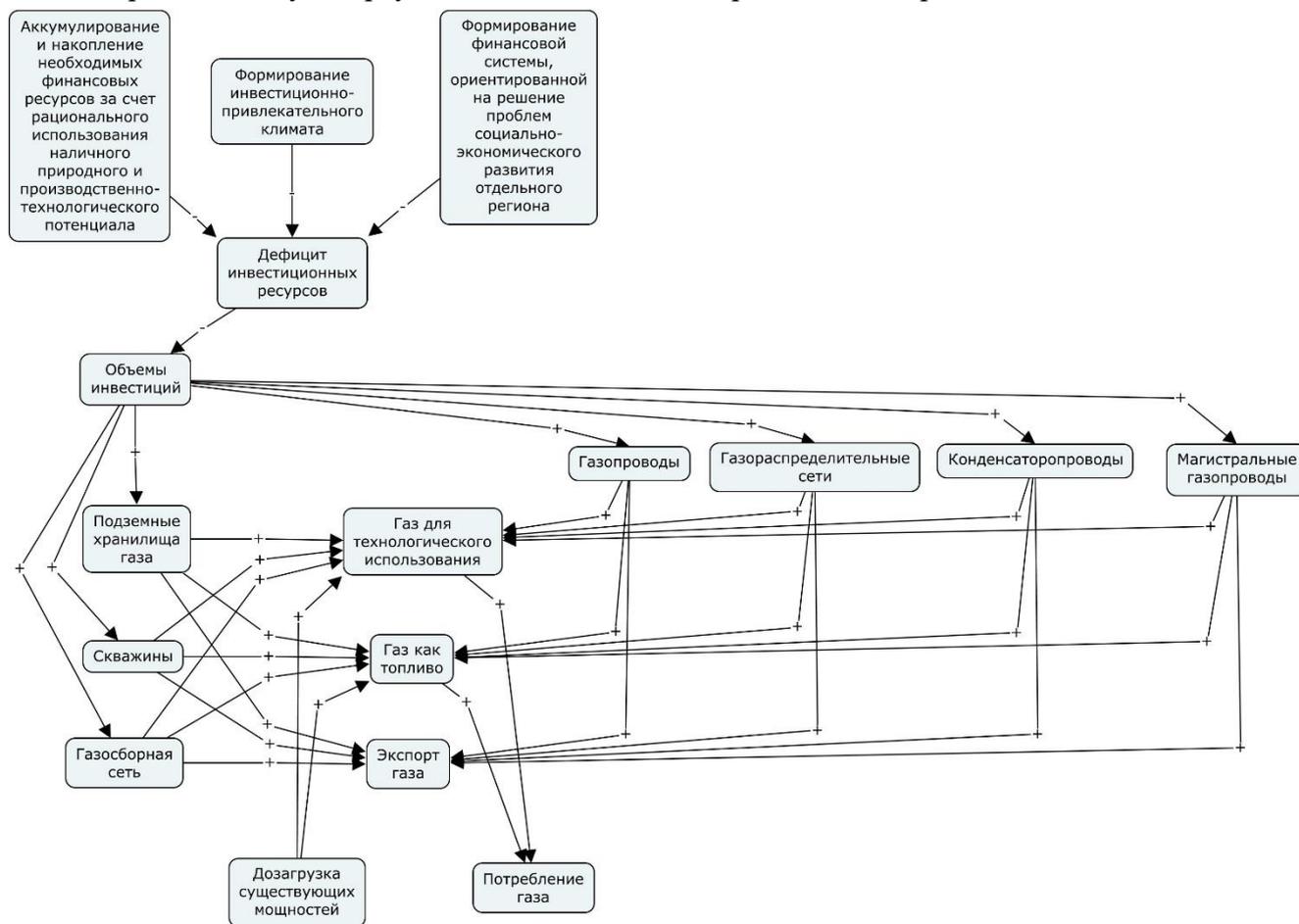


Рис. 5. Когнитивная карта угрозы «Недостаток инвестиций» для газовой отрасли.

Этап количественного анализа предполагает проведения исследований с помощью традиционных оптимизационных моделей топливно-энергетического комплекса. Для проведения экспериментальных расчетов формируется блок изменений в технико-экономические показатели ограничений модели, а именно: ограничения по производственным мощностям моделируемых объектов отраслевых систем ТЭК, пропускным способностям транспортных связей. Основой для получения количественных оценок служит модель оценки территориально-производственной структуры ТЭК с учетом требований ЭБ, подробно рассмотренная в [8]. Эта модель может использоваться в двух режимах:

- в режиме определения оптимального развития энергетических технологий (с учетом структурной избыточности в виде резервов мощностей, запасов топлива,

взаимозаменяемости энергоресурсов) и оптимального распределения потребляемых энергоресурсов,

– в режиме определения недопоставок энергоресурсов (дефицитов ТЭР) в целом по стране и по отдельным регионам.

Эта модель характеризуется иерархией уровней, представленной на рис. 6. Технологически она состоит из отраслевых подсистем энергетического комплекса (газовой, угольной, нефтеперерабатывающей (в части мазутоснабжения) отраслями, электро- и теплоэнергетикой).

В конечном варианте она дополнена финансовым блоком, описывающим инвестиционные затраты на реконструкцию, модернизацию действующих мощностей, вывод устаревшего оборудования, ввод новых мощностей на объектах энергетических отраслей. В ней же реализован учет динамики развития ситуации [3], позволивший отследить такие особенности многошагового процесса развития ТЭК, как:

- ввод новых производственных мощностей;
- демонтаж и консервацию старых объектов;
- реконструкцию объектов с изменением технологической схемы.

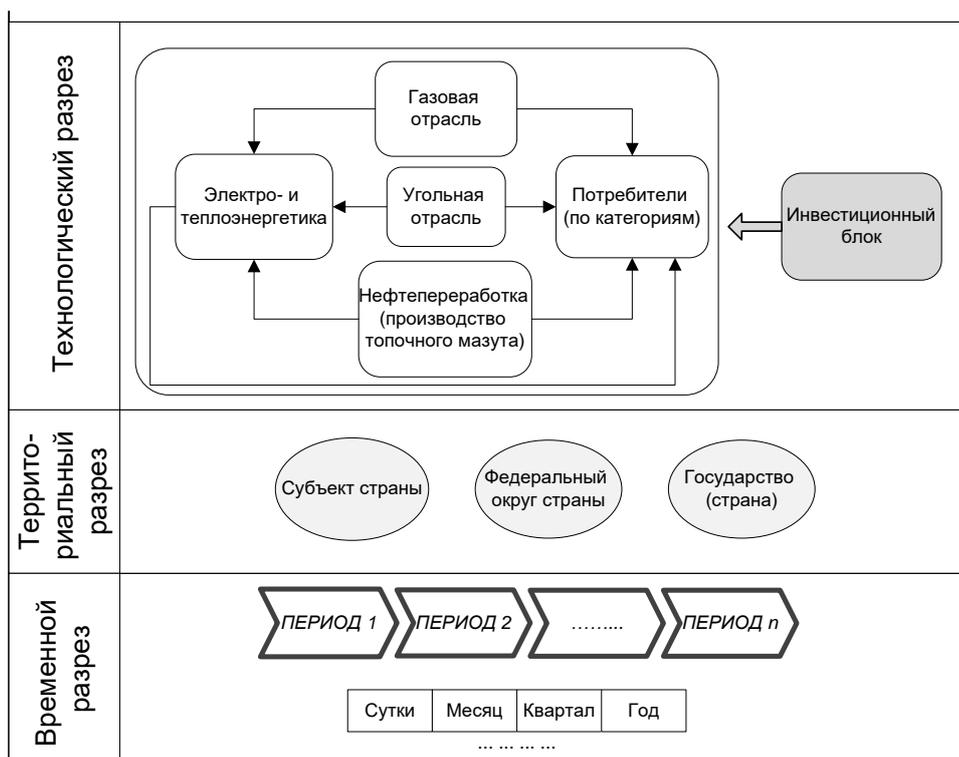


Рис. 6. Территориальная, временная и технологическая структура модели ТЭК

В целом модель используется для определения следующих характеристик (показателей):

– размеров недопоставок (дефицита) в отдельных видах энергоресурсов по рассматриваемым категориям потребителей, выделенным территориальным объединениям и в целом по стране, как величина невязки между заданной потребностью и возможностью производства данного вида энергоресурса (с учетом запасов, возможностей замещения этого вида энергоресурса у других потребителей и др.);

- изменений пропускных способностей межрайонных транспортных связей, определяемых путем сравнения соответствующих показателей рассматриваемого варианта с исходным;
- рекомендуемого рационального использования производственных мощностей энергетических объектов, а также распределение отдельных видов энергоресурсов по категориям потребителей.

**Заключение.** В статье рассмотрен подход к решению задачи оценки влияния угроз на состояние энергетической безопасности с применением семантических моделей. Приведены иллюстративные примеры использования методов когнитивного моделирования для задач анализа влияния угроз на состояние энергетической безопасности, в частности, применение когнитивных карт различных уровней для задания сценариев развития критических и чрезвычайных ситуаций. Предлагается интеграция математического и когнитивного моделирования для реализации этого подхода. Конкретизация когнитивных моделей позволяет строить когнитивные карты одного уровня с традиционными оптимизационными моделями топливно-энергетического комплекса, что в свою очередь, позволит в дальнейшем использовать когнитивные модели в качестве интерфейса корректировки экономико-математических моделей и их визуализации.

**Благодарности.** Работа была выполнена при частичной финансовой поддержке грантов РФФИ №18-07-00714, № 19-07-00351, 19-57-04003, №20-08-00367.

#### СПИСОК ЛИТЕРАТУРЫ

1. Аршинский В. Л., Массель А. Г., Сендеров С. М. Информационная технология интеллектуальной поддержки исследований проблем энергетической безопасности // Вестник Иркутского государственного технического университета. 2010. № 7 (47). С. 8–11.
2. Зенкин А. А. Когнитивная компьютерная графика / Под ред. Д. А. Поспелова. М.: Наука, Гл. ред. физ.-мат. лит. 1991. 192 с.
3. Макаров А. А., Мелентьев Л. А. Методы исследования и оптимизации энергетического хозяйства / А. А. Макаров, Л. А. Мелентьев. Новосибирск: Наука. 1973. 274 с.
4. Массель А. Г. Методологический подход к организации интеллектуальной поддержки исследований проблемы энергетической безопасности // Информационные технологии. 2010. № 9. С. 32–36.
5. Массель А. Г., Пяткова Е. В. Интеллектуальная ИТ-среда для ситуационного анализа проблем энергетической безопасности // Методические вопросы исследования надежности больших систем в энергетике. Иркутск: Институт систем энергетики им. Л.А. Мелентьева СО РАН. 2014. С. 472–483.
6. Массель Л. В., Массель А. Г. Интеграция семиотики, когнитивной графики и семантического моделирования в интеллектуальных семиотических системах ситуационного управления // Открытые семантические технологии проектирования интеллектуальных систем. 2016. № 6. С. 71–76.
7. Массель Л. В., Массель А. Г. Семантические технологии на основе интеграции онтологического, когнитивного и событийного моделирования // Открытые

- семантические технологии проектирования интеллектуальных систем. 2013. № 3. С. 247–250.
8. Энергетическая безопасность России / В. В. Бушуев, Н. И. Воропай, А. М. Мастепано, Ю. К. Шафраники, Новосибирск: Наука. Сиб. издат. фирма РАН. 1998. 306 с.
9. Н. И. Пяткова, В.И. Рабчук, С. М. Сендеров, Г. Б. Славин, М. Б. Чельцов Энергетическая безопасность России: проблемы и пути решения // Новосибирск: СО РАН. 2011. 198 с.
10. Axelrod R. Structure of decision: The cognitive maps of political elites / R. Axelrod. Princeton University Press. 1976. 1–405 с.
11. Gaskova D. A., Massel A. G. Semantic modeling of cyber threats in the energy sector using Dynamic Cognitive Maps and Bayesian Belief Network // Proceedings of the 2nd international conference on modelling, simulation and applied mathematics (msam2017). 2019.
12. Kosko B. Fuzzy cognitive maps // International Journal of Man-Machine Studies. 1986. № 1 (24). С. 65–75.
- 

**UDK 004.8:620.9**

**APPLICATION OF COGNITIVE MODELING METHODS  
FOR ENERGY SECURITY THREAT ANALYSIS**

**Aleksei G. Massel**

Ph.D., Senior researcher, Department of Artificial Intelligence Systems  
in the Energy Sector e-mail: [amassel@isem.irk.ru](mailto:amassel@isem.irk.ru),

**Natalia I. Pyatkova**

Ph.D., Senior researcher, Laboratory Reliability of Fuel and Energy Supply  
e-mail: [nata@isem.irk.ru](mailto:nata@isem.irk.ru),

Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences  
130, Lermontov Str., 664033, Irkutsk, Russia

**Abstract.** The article discusses a two-level information technology for researching energy security problems, where the levels of qualitative and quantitative analysis are combined. The use of cognitive models of different levels is discussed in more detail. The level of qualitative analysis of the threat "Under-investment" is described. The structure of traditional optimization models of the fuel and energy complex is described. The possibility of integrating cognitive and mathematical models is shown.

**Keywords:** Semantic modeling, cognitive modeling, energy security.

**References**

1. Arshinskii V. L., Massel A. G., Senderov S. M. Informatsionnaia tekhnologiya intellektualnoi podderzhki issledovaniu problem energeticheskoi bezopasnosti [Information technology of intelligent support for the study of energy security problems] // Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta. 2010. №7(47) Pp. 8-11.

2. Zenkin A. A. Kognitivnaia kompiuternaia grafika. [Cognitive computer graphics] / Pod red. D. A. Pospelova. – M.: Nauka. Gl. Red. Fiz. –mat. Lit. 1991. 192 p.
3. Makarov A. A., Melentev L. A. Metody issledovaniia i optimizatsii energeticheskogo khoziaistva. [Research methods and optimization of the energy economy]. Novosibirsk: Nauka=Science. 1973. 274 p.
4. Massel A. G. Metodologicheskii podkhod k organizatsii intellektualnoi podderzhki issledovaniu problemy energeticheskoi bezopasnosti. [Methodological approach to organizing intellectual support for research on energy security] // Informatsionnye tekhnologii. 2010. №9. Pp 32-36.
5. Massel A. G., Piatkova E. V. Intellektualnaia IT-sreda dlia situatsionnogo analiza problem energeticheskoi bezopasnosti. [Intelligent IT environment for situational analysis of energy security problems] // Metodicheskie voprosy issledovaniya nadezhnosti bol'shikh sistem v energetike = Methodological issues of research on the reliability of large systems in the energy sector. Irkutsk: Institut sistem energetiki im. L. A. Melenteva SO RAN= Irkutsk: L. A. Melentyev Institute of Energy Systems SB RAN. 2014. Pp. 472-483.
6. Massel' L. V., Massel' A. G. Integratsiia semiotiki, kognitivnoj grafiki i semanti-cheskogo modelirovaniia v intellektual'nykh semioticheskikh sistemakh situatsionnogo upravleniia [Integration of semiotics, cognitive graphics and semantic modeling in intelligent semiotic systems of situational control] // Otkrytye semanticheskie tekhnologii proektirovaniia intellektual'nykh sistem. 2016. (6). Pp. 71–76.
7. Massel' L. V., Massel' A. G. Semanticheskie tekhnologii na osnove integratsii ontologicheskogo, kognitivnogo i sobytijnogo modelirovaniia [Semantic technologies based on the integration of ontological, cognitive and event modeling] // Otkrytye semanticheskie tekhnologii proektirovaniia intellektual'nykh sistem. 2013. № 3. Pp. 247–250.
8. Energeticheskaya bezopasnost' Rossii [Energy security of Russia] / V.V. Bushuev, N.I. Voropay, A.M. Mastepanov, Yu.K. Shafranik i dr. Novosibirsk: Nauka. Sibirskaya izdatel'skaya firma RAN = Novosibirsk: Science. Siberian Publishing Company RAS]. 1998. 302 p.
9. N. I. Pjatkova, V. I. Rabchuk, S. M. Senderov, G. B. Slavin, M. B. Chel'cov Energeticheskaja bezopasnost' Rossii: problemy i puti resheniia [Energy security of Russia: problems and solutions] // Novosibirsk: SO RAN. 2011. 198 p.
10. Axelrod R. Structure of decision: The cognitive maps of political elites / R. Axelrod, Princeton University Press, 1976. 405 p.
11. Gaskova D. A., Massel A. G. Semantic modeling of cyber threats in the energy sector using Dynamic Cognitive Maps and Bayesian Belief Network // Proceedings of the 2nd international conference on modelling, simulation and applied mathematics (msam2017). 2019.
12. Kosko B. Fuzzy cognitive maps // International Journal of Man-Machine Studies. 1986. № 1 (24). C. 65–75.

## КОМПЛЕКС СРЕДСТВ ПОДДЕРЖКИ ПРОЦЕССОВ РАЗРАБОТКИ И СОПРОВОЖДЕНИЯ РЕШАТЕЛЕЙ ДЛЯ СИСТЕМ С ОНТОЛОГИЧЕСКИМИ БАЗАМИ ЗНАНИЙ

**Грибова Валерия Викторовна**

д.т.н, с.н.с, зам.директора по научной работе

e-mail: [gribova@iacp](mailto:gribova@iacp),

**Шалфеева Елена Арефьевна,**

к.т.н, доцент, с.н.с.

e-mail: [shalf@dvo.ru](mailto:shalf@dvo.ru),

Институт автоматизации и процессов управления ДВО РАН, Владивосток,  
690041, Приморский край, г. Владивосток, ул. Радио, 5

**Аннотация.** Использование онтологического подхода является одним из современных подходов к созданию систем с базами знаний. Для построения жизнеспособных программных сервисов, работающих с такими базами знаний, и управления их коллективной разработкой предложена инструментальная среда. Метод инструментальной поддержки нацелен на создание и развитие библиотек онтолого-базируемых операций, на использование их при конструировании программных средств, на распределение полномочий по созданию компонентов систем с базами знаний, на контроль и интеграцию их в облачную сопровождаемую систему поддержки принятия решений на основе знаний.

**Ключевые слова:** онтологическая база знаний, алгоритм решения, решатель задачи, повторное использование, сопровождение, коллективная разработка.

**Цитирование:** Грибова В. В., Шалфеева Е.А. Комплекс средств поддержки процессов разработки и сопровождения решателей для систем с онтологическими базами знаний // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 34-43. DOI:10.38028/ESI.2020.20.4.003

**Введение.** Современным подходом к созданию *систем с базами знаний* является использование онтологического подхода. Все преимущества такого способа создания баз знаний активно обсуждаются в литературе и, в настоящее время, как отмечено в [1], являются стандартом «де факто».

Технология разработки баз знаний на основе онтологий достаточно широко обсуждается в литературе, для разработки онтологических баз знаний предложен ряд инструментальных средств и платформ, типичными представителями являются OSTIS, Protégé, OntoEdit, IACPaaS и ряд других [2, 7,].

Вместе с тем проблема разработки решателей задач, основанных на онтологических базах знаний, освещена в литературе гораздо меньше. Традиционно считается, что в качестве решателя системы с базой знаний (СБЗ) выступает машина вывода, обрабатывающая продукционные правила, алгоритмический подход не применим к созданию такого класса систем, именно поэтому часто онтологические базы знаний преобразуют в базу правил и для нее разрабатывают машину вывода [9]. Для создания решателей используются и

специализированные, нетрадиционные модели и технологии, например, технология OSTIS, в которой решатель, по определению авторов, рассматривается в «неклассическом варианте» и представляет собой графодинамическую SC-машину [3].

При разработке онтологических баз знаний, представленных семантическими сетями, решатель задач может быть представлен в виде алгоритма, выполняющего обход базы знаний на основе знания ее структуры (онтологии) для сопоставления информации из базы знаний входным данным [6]. Очевидно, что разработка решателя для онтологических СБЗ должна отвечать всем основным требованиям к разработке программных систем: обеспечение снижения трудоемкости разработки и сопровождения, прозрачность, поддержка коллективной разработки. Первые три требования реализуются через применение специализированных оболочек, декларативно-компонентный метод разработки решателей, упрощение интерфейсов и уменьшение сцепления программных компонентов, их повторную используемость.

Применение специализированных оболочек, ориентированных на конкретный класс систем с базами знаний, обеспечило заметное сокращение сроков разработки и усилий на сопровождение систем, но у специализированных оболочек "жесткий" решатель и встроенный в него интерфейс, которые, в случае изменения требований, трудно модифицировать. Альтернатива оболочкам - применение объектно-ориентированного программирования и архитектуры MVC. При их использовании «в коде создаются программные объекты, описывающие свойства моделируемых объектов реального мира и операции над ними», поэтому «в программном коде прикладных автоматизированных систем все еще содержится много логики, связанной непосредственно с предметной областью и конкретными решаемыми задачами». «Очевидным ограничением этого подхода является необходимость модифицировать код при внесении изменений в структуру модели» [4].

Для коллективной разработки необходимо обеспечивать каждый тип разработчиков собственным «рабочим местом», предоставляющим ему необходимую функциональность для исполнения своих задач.

Современные технологии Protege, OSTIS и IACPaaS [2, 6, 7] поддерживают работу по созданию знаний и работе с ними на дистанционно доступных собственных платформах. На базе инструментального комплекса АТ-Технология разработана веб-ориентированная версия для поддержки построения прикладных интегрированных экспертных систем на всех традиционных этапах их создания [5].

Целью данной работы является описание комплекса средств поддержки процессов разработки и сопровождения решателей для систем с онтологическими базами знаний, реализованных в виде алгоритма.

**1. Поддержка разработки и сопровождения решателей: онтологические решатели в системах с базами знаний.** Онтологический решатель в общем случае обрабатывает несколько информационных ресурсов (*ИнфРес*), сформированных на основе онтологии.

Онтология предметной области ПрОбл определяет правила представления и обработки информации (анализ, формирование, модификация) и включает

$$Onto = KnOnt \cup SitOnt \cup Agree \cup Dict,$$

где *KnOnt* определяет структуру профессиональных знаний о решении задачи,

*SitOnt* – структуру информации об объектах действительности,

*Agree* – онтологические соглашения о правилах обработки, т.е. сопоставления фактов и знаний в процессе рассуждения и принятия решений, а также о правилах выдвижения,

подтверждения, опровержения гипотез, содержащихся в знаниях или создаваемых с их помощью;

*Dict* – словарь названий и области значений понятий ПрОбл.

Онтологический решатель *Solv* ( $KnOnt \cup SitOnt \cup Agree$ ) в составе системы с базой знаний (СБЗ) должен быть способен выдвигать гипотезы по результату сопоставления входной информации (об объекте) – предложениям из базы знаний (Knowledge Base, KB) (*KnOnt*, *Dict*) на основе онтологических соглашений *Agree* о связи данных и знаний и с учетом знания структуры предложений (связей терминов) *KnOnt*, и ограничений на интерпретацию смысла терминов. Задача онтологического решателя – предложить (путем «рассуждения») одно или несколько обоснованных решений-гипотез, использующих входную информацию и согласованных с Базой знаний KB (*KnOnt*). Обоснование каждой гипотезы – явное указание ее связи с теми предложениями Базы знаний, которые ее подтверждают или не отвергают.

При применении соглашений о соответствии фактов знаниям о рассматриваемых явлениях, связываемые понятия используются для промежуточных заключений. Часто в роли посылок (причин для принятия промежуточных решений) выступают факты из ситуаций действительности и ожидаемые, согласно знаниям, проявления анализируемых процессов. Следствиями являются значения предикатов их соответствия. Обработка – это поиск (в документе, описывающем действительность) фактов для подтверждения условий течения процессов (и «вывод»/получение значений соответствующих предикатов); вывод значений предикатов (на множестве условий, на множестве проявлений) – предикатов существования гипотез из элементов модели знаний по полученным значениям предикатов-«посылок», выдвижение гипотез на основе соответствующих вычисленных предикатов (подтвержденных или (не-) отвергнутых посылок).

Онтологический решатель (в технологии IASPaas [6]) – программно реализованное «рассуждение», алгоритм, проводящий обход каждой Базы знаний KB (представленной семантической сетью) для сопоставления входным данным (фактам или условиям) знаний и соглашений, выдвижения и объяснения гипотез и построения семантической сети, фиксирующей этапы рассуждения – шаги принятия или отклонения возможных логических заключений.

**2. Проектирование онтологического решателя из программных единиц.** Решатель (умеющий выдвигать гипотезы и объяснять их) планируется к реализации из программных единиц (*ПрЕд*), среди которых единицы ( $Unit_m$ ) разных типов, с доступом и без доступа к *ИнфРес* (информационным компонентам СБЗ), для вывода (или промежуточного заключения), для поиска фактов, для вычислений, для связи с внешним окружением.

*ПрЕд* для вывода – запрашивает и получает информацию о подтверждении или отрицании элементов отношений или их цепочек, прописанных в *Kn* (часто как ответ с областью значений {данные соответствуют; противоречат; отсутствуют}). Например, *ПрЕд* проверяет истинность или подтвержденность варианта развития процесса для подтверждения существования процесса; подтвержденность последовательности периодов развития для подтверждения Варианта. Программная единица, обрабатывая такие виды связей между элементами *ИнфРес*, делает промежуточные заключения процесса логического вывода.

*ПрЕд* ( $Unit_m$ ), которые делают некоторое заключение о соответствии подмножества фактов знаниям некоторого типа (причинно-следственных или других структурных связей

между понятиями), «имеют дело» с *Agreem*. В совокупности такие *ПрЕд*, как правило «покрывают» все утверждения из  $KnOnt \cup Agreeem$ . Решатель, создаваемый из *ПрЕд*, программируемых по такому принципу – онтологический.

В алгоритме решателя происходит поочередный вызов *ПрЕд* – процедур вывода следствий из обработанных посылок, которые записывают результат проверки посылок в объяснение. Например, *ПрЕд* «проверить вариант развития процесса» активируется после активации *ПрЕд* «проверить гипотезу о процессе» и после активации *ПрЕд* «проверить необхУсловия процесса». Она реализует процедуру вывода заключения о подтверждении или опровержении наличия комплекса признаков из обработанных результатов подтверждения или опровержения всех обязательных признаков и всех дизъюнктивных наборов признаков.

*ПрЕд* для поиска фактов получает список условий на события и прочие возможные факторы, ищет в документе названия таких наблюдений, сравнивает с условиями и дает ответ: данные соответствуют условию или данные не соответствуют или данные отсутствуют; она также получает значения наблюдений, ищет в документе названия таких наблюдений, сравнивает с условиями и дает ответ: {данные соответствуют; противоречат/ не соответствуют; отсутствуют}.

В составе решателя также могут быть и другие *ПрЕд*, не нуждающиеся в доступе к *ИнфРес* – вычислительные, интерфейсные.

**3. Онтологические операции доступа к информации как вид программных единиц и метод инструментальной поддержки коллективного развития библиотек операций.** При разработке алгоритма для каждого шага вывода требуется либо обращение к элементам описания модели (знаний), либо к фактам действительности, либо к тому и другому.

*ПрЕд* для поиска фактов и *ПрЕд* для вывода обращаются с запросами к той КВ, где содержатся знания об известных связях, типы которых зафиксированы в онтологии знаний. Например, *ПрЕд* для поиска методов для воздействия на признак для изменения его значений в заданном направлении *getMethodsForTrendSign* реализует отношения «состояние – воздействие – результат» (между методом воздействия, признаком(-ами) до начала воздействия и признаком(-ами) желаемыми\целевыми).

При поиске фактов для подтверждения условий существования искомых явлений или признаков их проявления (т.е. «посылок» для вывода) требуется обращение к фактам действительности (к возрасту, к размеру, к температуре...). Каждое такое обращение за фактом или их комплексом – «кандидат» на самостоятельную операцию доступа  $Op_i$  к документу с данными  $Unit_m$ , которые делают некоторое заключение о соответствии подмножества фактов знаниям некоторого типа, сопоставляют результаты выполнения операций над информацией  $Sit$  и над Базами знаний КВ. Операции могут получать на вход список параметров, на выходе обеспечивать вычисленные значения либо модификацию *ИнфРес*, например, объяснения или создаваемого плана\проекта (породить экземпляр указанного понятия, записать значение указанного элемента).

Операции доступа к содержимому онтологических *ИнфРес* наиболее связаны со структурными и причинно-следственными связями используемых понятий (определяемыми  $KnOnt$  и  $SitOnt$ ), поэтому являются онтологическими. С конкретной онтологией связывается множество таких операций, чтобы применяться затем к любым *ИнфРес* (знаниям, данным), управляемым этой онтологией.  $Op_i$  осуществляют запросы к элементам структуры *ИнфРес* в соответствии с правилами интерпретации и соглашениями.

Набор операций отражает востребованные запросы, возникающие в переборных алгоритмах решения классов задач. Формирование и использование операций доступа  $Op_i$  для управляемого программного доступа к целевым базам знаний и другим хранимым *ИнфРес* требует библиотек онтолого-базированных операций доступа к *ИнфРес*.

В процессе разработки (и сопровождения) онтологического решателя IASPaas и его *ПрЕд* (IASPaas-агентов) предлагается использование готовой операции (которую можно вызывать из агента). Разработка новой операции – это декларативное описание операции, редактирование, сохранение и компиляция исходного кода операции с использованием средства «Редактор библиотеки операций».

При генерации заготовки агента, исходя из его декларативного описания с указанием онтологий (из  $KnOnt \cup SitOnt$ ) обрабатываемых *ИнфРес*, предложено оповещать о наличии операций над такими *ИнфРес*, предложить список их имен и описаний с возможностью сгенерировать в коде шаблон-заготовку вызова-обращения к выбранной операции для *ИнфРес*.

Таким образом, эффективное конструирование (снижение трудозатрат) решателей ряда практических задач основано на повторном использовании программных единиц и операций. Одни и те же виды структурных и причинно-следственных связей могут приниматься во внимание при решении разных классов задач (прогноз, диагностика, планирование и др.) по единой модели течения процесса. В частности, построение заключения о соответствии подмножества фактов знаниям некоторого типа причинно-следственных связей, характерного для задач разных классов (например, диагностики и прогноза) является часто выполняемой обработкой. Для Решателей таких задач (или интегрирующих их составных задач) шанс повторно использовать  $Unit_m$ , которые делают такое заключение, увеличивается. Даже при создании одного Решателя частые (повторяемые) варианты поиска, сравнения, выбора, добавления информации рекомендуется оформлять в самостоятельные *ПрЕд*, обрабатывающие *ИнфРес*, для их повторного использования.

**4. Интегрированная среда построения жизнеспособных программных компонентов.** Для обеспечения сборки Решателя из заменяемых компонентов и повышения прозрачности его архитектуры актуальны средства декларирования всех составных частей – *ПрЕд* (и порядка их запуска), контроля доступности используемых ими *ИнфРес*, стандартизация их взаимодействий (шаблоны обращений), что особенно важно в условиях коллективной разработки и развития. Следовательно, важна инструментальная поддержка декларирования решателя (произвольной задачи) и декларирования его компонентов.

Декларация онтологической *ПрЕд* содержит среди параметров имена обрабатываемых *ИнфРес* и дополнительно содержит список ссылок на онтологии этих *ИнфРес*. Декларирование включает и описание назначения *ПрЕд*, что способствует их повторному использованию. Декларация *ПрЕд* содержит формат обращения к ней со списком типов параметров. Для интегрируемости онтологических *ПрЕд* ( $SemUnit$  и  $SemOp$ ) их внешние связи (форматы знаний и данных, входных и выходных сообщений) тоже специфицируются (в соответствующей декларации этой агента).

Создаваемая IASPaas-декларация решателя – это документирующий *ИнфРес*, он содержит список ссылок на онтологии всех обрабатываемых *ИнфРес*, читаемых и формируемых. Такая спецификация Решателя может быть интерпретируема человеком и машиной. Кроме спецификации форматов знаний и данных, могут быть определены все  $Unit_m$

и даже порядок их выполнения и связей друг с другом. Декларация Решателя может использоваться (1) автоматически для контроля сборки, для динамической (runtime) загрузки частей по мере потребности в них, (2) человеком – для сборки СБЗ из Решателя и всех обрабатываемых *ИнфРес*, а также для его сборки из *ПрЕд*, если для них создаются свои специфицирующие документы. Проектирование архитектуры решателей задач из декларируемых программных компонентов разных типов – шаг в направлении создания жизнеспособных систем (и еще один ключевой принцип проектирования). В декларации решателя задач, как минимум, указывается агент, начинающий работу. Другие *Unit<sub>m</sub>* могут явно не перечисляться, а быть указаны внутри “первого” агента при передаче им управления.

Кроме средств (инструментальных сервисов) декларирования нужны средства создания заготовок исходных кодов по декларациям, средства кодирования новых *ПрЕд*, средства каталогизации *Unit<sub>m</sub>* как потенциально повторно используемых компонентов и средства интеграции ПИК и новых *ПрЕд* в новые решатели или их новые версии. Поэтому в *среде* декларативно-компонентного метода разработки и развития СБЗ инструментарий для формирования *программных решателей* включает:

- онтологию *ПрОбл* (включая базу терминов),
- инструменты декларирования решателя и *ПрЕд*,
- среда разработки программ (IDE) для кодирования *ПрЕд*.
- библиотеки подготовленных и протестированных *ПрЕд*, обрабатывающих *ИнфРес*.

Методы и инструменты декларирования решателя нацелены на обеспечение прозрачности его архитектуры, а декларирование *ПрЕд* – на поддержку создания их исходных кодов и динамическое подключение в процессе выполнения.

**5. Поддержка коллективной разработки программных компонентов.** В соответствии с традициями инструментальной поддержки разработки интеллектуального сервиса (Типовые Программные Процедуры поддержки ресурсно-календарного планирования [5]) в IASPaas среде на основе декларативного описания процесса разработки создаются инструменты планирования работ, ответственных исполнителей, сроков, передачи результатов производства компонентов СБЗ и фрагментов этих компонентов.

В *онтологии коллективной разработки интеллектуальных сервисов*, по которой может быть построено декларативное описание очередного Проекта, выделяются основные типы участников разработки, всевозможные действия и полномочия участников. Участники разработки реализуют свои полномочия в собственных кабинетах – «рабочих местах».

Для разработки СБЗ строится дерево задач. Данное дерево отображает, что необходимо сделать для создания интеллектуального сервиса, а каждая вершина является задачей для конкретного исполнителя. В вершинах обозначены роли ответственных за исполнение задачи. Корневой вершине соответствует всё дерево, а исполнение её задачи соответствует завершению всего проекта по разработке интеллектуальной системы.

Задача имеет ряд атрибутов, среди которых – стадия. С помощью стадий определяется жизненный цикл задачи. В отношении задачи всегда существуют участники разработки в двух ролях: автор задачи, который внёс задачу в систему и выбрал её исполнителя, и, непосредственно, исполнитель задачи. В зависимости от роли, пользователю доступны разные переходы по стадиям.

Для управления коллективной разработкой предоставляется функциональность для создания *дерева задач процесса разработки* интеллектуального сервиса на основе IASaaS-технологии разработки сервисов, основанных на знаниях.

Инициатор разработки или *руководитель Проекта* должен назначить исполнителей на все задачи в дереве. В свою очередь, исполнители задачи могут декомпозировать (на несколько подзадач) или делегировать (всю) свою задачу. В этом случае к вершине дерева задач добавляются новые вершины. Исполнитель имеет возможность исполнить свою задачу (перемещая её по стадиям в рамках жизненного цикла) и возможность управлять задачей – например, создавать дочерние задачи, предоставлять результат исполнения. По мере исполнения задач проводится управляемое объединение результатов задач.

Комплекс инструментальных сервисов поддержки процесса коллективной разработки создается на основе онтологического подхода. В этом случае технология разработки рассматривается как *ПрОбл*, а онтология процесса разработки является основой онтологических *программных средств разработки* (инструментальных сервисов). В частности, инструментальные сервисы декларирования решателя и *ПрЕд* – это Редакторы (соответствующей информации), генерируемые по составной части онтологии этой технологии – по документу, описывающему содержание и формат декларации.

Декларативное определение процесса коллективной разработки позволяет наглядно представлять процессы и состояние разработки.

**Заключение.** На платформе IASaaS [6], используемой для дистанционного создания и использования онтологических баз знаний, разработан комплекс средств поддержки процессов разработки и сопровождения онтологических решателей.

Онтологический решатель обрабатывает несколько информационных ресурсов, сформированных на основе онтологии. В процессе рассуждения и принятия решений алгоритм опирается на формализованные связи и зафиксированные онтологические соглашения. Особенности, принципиально отличающие онтологический решатель от других решателей:

- алгоритм формируется на основе онтологии и не зависит от самих БЗ, что соответствует современному подходу к разработке интеллектуальных систем;
- в нем «естественно» реализуется подробное и понятное специалисту объяснение предлагаемого решения.
- Для снижения трудоемкости разработки и сопровождения онтологического решателя предложены следующие решения:
  - декларативное представление всех программных единиц;
  - декларативное представление решателя из составных частей, включая порядок их запуска;
  - контроль доступности используемых *ИнфРес*,
  - генерация фрагментов кода по этому представлению,
  - инструменты поддержки коллективного использования и развития библиотеки онтологических операций доступа к *ИнфРес*.

Инструментальной средой для создания таких решателей задач является облачная платформа IASaaS – среда для построения проблемно-ориентированных онтологий и решателей соответствующих классов задач. Под управлением проблемно-ориентированных онтологий (с общим доступом) участниками процесса создаются фрагменты баз знаний и

*ПрЕд* для решателей в рамках коллективной разработки СБЗ, обладающих свойством жизнеспособности.

Работа выполнена при частичной финансовой поддержке РФФИ (проекты №№ 19-07-00244, 18-07-01079).

#### СПИСОК ЛИТЕРАТУРЫ

1. Гаврилова Т.А., Страхович Э.В., Визуально-аналитическое мышление и интеллектуальные карты в онтологическом инжиниринге // *Онтология проектирования*. 2020. Т. 10. № 1. С. 87-99.
2. Голенков В.В., Гулякина Н.А. Проект открытой семантической технологии компонентного проектирования интеллектуальных систем. Часть 2: унифицированные модели проектирования // *Онтология проектирования*. 2014. №4(14). С 34-53;
3. Голенков В.В., Д.В. Шункевич, Давыденко И.Т. Семантическая технология проектирования интеллектуальных решателей задач на основе агентно-ориентированного подхода // *Программные системы и вычислительные методы*. 2013. №1(2). С.82-94.
4. Горшков С. В. Онтологическое моделирование предприятий: методы и технологии: монография. Екатеринбург: Изд-во Урал. ун-та. 2019. 236 с.
5. Рыбина Г. В., Блохин Ю. М., Иващенко М.Г. Интеллектуальная технология построения интегрированных экспертных систем // *Искусственный интеллект и принятие решений*. 2011. № 3. С. 48-57.
6. Gribova, V.V., Kleschev A.S., Moskalenko Ph.M., Timchenko, V.A., Fedorischev L.A., Shalfeeva E.A. 2017a. The IACPaaS Cloud Platform: Features and Perspectives. In Proc. of 2017 Second Russia and Pacific Conference on Computer Technology and Applications (RPC). Vladivostok. Pp. 80-84.
7. Musen M. The Protégé Project: A Look Back and a Look Forward // *AI Matters*. 2015 Jun; № 1(4). Pp. 4-12.
8. Sure Y., M. Erdmann, J. Angele, S. Staab, R. Studer, Wenke D. OntoEdit: Collaborative ontology development for the Semantic Web // *Proc. of the Inter. Semantic Web Conference (ISWC 2002)*. Sardinia. Italia.
9. Verhodubs O., Grundspençkis, J. Evolution of Ontology Potential for Generation of Rules,” *Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics, Craiova*. 2012. Article № 58. Pp. 1–4. DOI:10.1145/2254129.2254201.

УДК 004.4`2+004.89

**A SET OF TOOLS TO SUPPORT DEVELOPMENT AND MAINTENANCE  
OF SOLVERS FOR SYSTEMS WITH ONTOLOGICAL KNOWLEDGE BASES**

**Valeria.V. Gribova**

Doctor of Technical Sciences, Senior Scientist, Deputy Director for Scientific Work

e-mail: [gribova@iacp](mailto:gribova@iacp),

**Elena.A. Shalfeeva**

Candidate of Technical Sciences, Associate Professor, Senior Scientist

e-mail: [shalf@dvo.ru](mailto:shalf@dvo.ru),

Institute of Automation and Control Processes

Far Eastern Branch of the Academy of Sciences,

690041, Vladivostok, Primorsky Territory, Vladivostok, st. Radio, 5

The modern approach to creating systems with knowledge bases is based on an ontological approach. A tool environment is proposed for building viable software services that work with such knowledge bases and managing their collective development. The tool support method is aimed at creating and developing libraries of ontology-based operations, using them in the design of software tools, allocating authority to create components of systems with knowledge bases, and controlling and integrating them into a cloud-based, supported knowledge-based decision support system.

**Keyword:** ontological knowledge base, solution algorithm, problem solver, reuse, maintainability, collaborative development.

**References**

1. Gavrilova T.A., Strakhovich E.V., Vizual'no-analiticheskoye myshleniye i intellekt-karty v ontologicheskoy inzhenerii [Visual-analytical thinking and mind maps in ontological engineering] // Ontologiya dizayna = Design Ontology. 2020. Vol. 10. № 1. Pp. 87-99.
2. Golenkov V.V., Gulyakina N.A. Proyeckt otkrytoy semanticheskoy tekhnologii komponentnogo proyecktirovaniya intellektual'nykh sistem. Chast' 2: unifitsirovannyye modeli proyecktirovaniya [Project of open semantic technology for component design of intelligent systems. Part 2: unified design models] // Ontologiya dizayna = Design Ontology. 2014. №. 4 (14). Pp .34-53.
3. Golenkov V.V., Shunkevich D.V., Davydenko I.T. Semanticheskaya tekhnologiya proyecktirovaniya intellektual'nykh reshateley zadach na osnove agentno-oriyentirovannogo podkhoda [Semantic technology for designing intelligent problem solvers based on an agent-based approach] // Programmnyye sistemy i vychislitel'nyye metody = Software systems and computational methods. 2013. №1 (2). Pp. 82-94.
4. Gorshkov S. V. Ontologicheskoye modelirovaniye predpriyatiy: metody i tekhnologii [Ontological modeling of enterprises: methods and technologies]. Yekaterinburg: Izd-vo Ural'skogo universiteta = Yekaterinburg: Publishing house of the Ural University. 2019. 236p.
5. Rybina G.V., Blokhin Yu.M., Ivaschenko M.G. Intellektual'naya tekhnologiya postroyeniya integrirovannykh ekspertnykh system [Intelligent technology for building integrated expert

- systems] // *Iskusstvennyy intellekt i prinyatiye resheniy = Artificial Intelligence and Decision Making*. 2011. № 3. Pp . 48-57.
6. Gribova, V.V., Kleshev A.S., Moskalenko Ph.M., Timchenko, V.A., Fedorischev L.A., Shalfeeva E.A. 2017a. The IACPaaS Cloud Platform: Features and Perspectives. In Proc. of 2017 Second Russia and Pacific Conference on Computer Technology and Applications (RPC). Vladivostok. Pp.80-84.
7. Musen M. The Protégé Project: A Look Back and a Look Forward // *AI Matters*. 2015 Jun; № 1(4). Pp. 4-12.
8. Sure Y., M. Erdmann, J. Angele, S. Staab, R. Studer, Wenke D. *OntoEdit: Collaborative ontology development for the Semantic Web* // Proc. of the Inter. Semantic Web Conference (ISWC 2002). Sardinia. Italia.
9. Verhodubs O., Grundspençkis, J. Evolution of Ontology Potential for Generation of Rules,” Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics, Craiova. 2012. Article № 58. Pp. 1–4. DOI:10.1145/2254129.2254201.

## ЛИНГВИСТИЧЕСКИЕ АСПЕКТЫ ИССЛЕДОВАНИЯ АРГУМЕНТАЦИИ НА ОСНОВЕ ОНТОЛОГИИ

**Кононенко Ирина Семеновна**

Н.с., e-mail: [irina\\_k@cn.ru](mailto:irina_k@cn.ru),

**Ахмадеева Ирина Равильевна**

М.н.с., e-mail: [i.r.akhmadeeva@iis.nsk.su](mailto:i.r.akhmadeeva@iis.nsk.su),

**Сидорова Елена Анатольевна**

К.ф.-м.н., с.н.с., e-mail: [lsidorova@iis.nsk.su](mailto:lsidorova@iis.nsk.su),

Институт систем информатики им. А.П. Ершова СО РАН,

630060 г. Новосибирск, пр. Лаврентьева 6.

**Аннотация.** Обсуждаются вопросы аннотирования корпуса русскоязычных научно-популярных текстов. Особое внимание уделяется схеме разметки, в которой предметом аннотирования являются не только прагматические и семантические аспекты аргументации, но и лингвистические индикаторы аргументации. Предлагается представление индикаторов в виде лексических и лексико-грамматических шаблонов, которые автоматически формируются по размеченному фрагменту текста, а затем при необходимости уточняются экспертом вручную по результатам поиска в корпусе текстов и анализа всех соответствующих фрагментов. Полученный словарь индикаторов используется для поиска и автоматического выделения индикаторов аргументации в неразмеченном тексте.

**Ключевые слова:** онтология аргументации, аргументативная разметка, схема аргументации, индикатор аргументации, онтологическая модель разметки текста, словарь индикаторов.

**Цитирование:** Кононенко И.С., Ахмадеева И.Р., Сидорова Е. А. Лингвистические аспекты исследования аргументации на основе онтологии // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 44-55. DOI: 10.38028/ESI.2020.20.4.004

**Введение.** В последнее десятилетие появилась потребность в решении ряда прикладных задач автоматической обработки текста, требующих обращения к глубинным уровням представления дискурса для анализа приведенных в тексте аргументов - мнений и их обоснований. Основой развития теории аргументации и предпосылкой успеха в области автоматического распознавания аргументов является создание корпусов текстов значительного объема, снабженных аргументативной разметкой. Такого рода ресурсы развиваются для разных языков на базе текстов различных жанров и функциональных стилей, однако для русского языка аргументативно размеченные корпуса текстов до сих пор не созданы. Ниже представлены некоторые результаты работы в рамках исследовательского проекта, направленного на создание корпуса русскоязычных научно-популярных текстов с разметкой аргументации.

В процессе предварительного анализа научно-популярных текстов замечено, что нередко аргументы вводятся в текст с помощью явных текстовых индикаторов. О присутствии

в тексте аргумента могут сигнализировать дискурсивные коннекторы (*потому что, именно, следовательно, и т.п.*) или свободные индикаторные конструкции. В качестве примера таких конструкций для рассуждений, ссылающихся на мнение экспертов (авторитетных в данной предметной области), можно привести конструкцию с ментальным предикатом и авторитетом в позиции субъекта:

*Недавно **ученые** из Университета Йорка **показали, что шимпанзе все-таки способны выучивать новые звуки.***<sup>1</sup>

Одним из важнейших аспектов аннотирования аргументации является регистрация индикаторов аргументации, которые играют ключевую роль в процессе понимания дискурса, облегчая идентификацию и реконструкцию шагов аргументации, осуществляемых в аргументативном диалоге или тексте [12].

Основы индикаторного подхода к извлечению информации из текста были заложены еще в 70-е годы прошлого столетия (см. достаточно детальное описание в [2]). Извлечение онтологической информации из научных текстов в [4] основано на индикаторном методе, в основе которого лежит обнаружение в тексте подсказок в виде словесных клише (образцов, маркеров), которые являются показателями аспектов содержания текста. В последнее время появились исследования, направленные на выявление шаблонных индикаторов структуры дискурса: так, в [11] описываются основные модели образования свободных конструкций, сигнализирующих о наличии причинно-следственного риторического отношения в дискурсе.

В работе [8] предлагается пилотный вариант инструкции по аннотированию схем аргументации в корпусе аргументированных эссе, уже снабженных разметкой аргументативной структуры. Для решения новой, существенно более сложной, задачи предлагается использование лингвистических индикаторов, которые приписаны отдельным схемам в неформальной инструкции для аннотаторов и могут служить им подсказками при разметке в текстах схем рассуждения (в особенности речь идет о сильных, т.е. однозначных индикаторах). Однако нам неизвестны проекты создания корпусов текстов, в которых предметом аргументативного аннотирования являлись бы не только структуры и схемы аргументации, но и лингвистические индикаторы и ставилась бы задача автоматического выделения индикаторов в неразмеченных текстах.

Целью данного исследования является обеспечение возможности использования индикаторов аргументации при решении задачи автоматического распознавания аргументов. Для этого при разметке текстов корпуса производится первичное ручное аннотирование индикаторов аргументации и создание словаря индикаторов, который далее используется для автоматического обнаружения индикаторов в еще не размеченных текстах по лингвистическим свойствам текстовых фрагментов (наличие лексического маркера, грамматического или пунктуационного признака) и/или свойствам, связанным с онтологией аргументации (т.е. по связи с типовой схемой рассуждения), что, в свою очередь, позволяет привлечь внимание аннотатора к потенциально аргументативным фрагментам размечаемого текста.

**1. Модель аргументативной разметки текста.** Любая разметка текста опирается на схему, которая задает перечень сущностей, используемых при аннотировании, их типизацию,

<sup>1</sup> В примере индикаторы выделены жирным шрифтом, а заключительное утверждение подчеркнуто.

формат и интерпретацию. В нашей работе в качестве схемы для разметки аргументации используется формат AIF (Argument Interchange Format) [6] и онтология, построенная на его основе [3, 9].

Отличительной особенностью этой онтологии является ее ориентированность на графовое представление аргументации, а также наличие разветвленной системы классов для метаописания сущностей предметной области. Таким образом, в онтологии AIF можно выделить три базовые концепции:

$$O = \langle O^G, O^C, O^F \rangle, \text{ где}$$

$O^G$  – графовое представление аргументации, содержащее вершины Node и бинарные отношения между ними. В классе Node выделяются два типа вершин, S-node, предназначенные для визуализации аргументов, и I-node – для визуализации утверждений, входящих в состав аргументов;

$O^C$  – предметная область аргументации, содержащая классы типовых схем рассуждения (аргументов), классы утверждений и отношения между ними, а также экземпляры этих классов. Эта онтология реализует Уолтоновский подход [13] к представлению аргументации на основе аргументативных схем, которые формально описывают внутреннюю структуру аргументов;

$O^F$  – онтология метаописаний, содержащая классы дескрипторов (форм) для схем аргументации и их атрибутов. Эта онтология позволяет задавать языковое описание схем аргументов.

На рис. 1 приведен пример описания типовой схемы рассуждения, построенного по онтологическому представлению класса *CouseToEffect\_Inference* и его формы.

Аргумент «CouseToEffect_Inference»		
Роль	Тип утверждения	Описание утверждения
Causal_Premise	Causal_Statement	A causes B
Conclusion	OccurrenceB_Statement	In this case, B will occur
OccurrenceA_Premise	OccurrenceA_Statement	In this case, A occurs

Рис. 1. Описание типовой схемы аргумента *CouseToEffect*.

Модель аргументативной разметки корпуса текста, в соответствии с заданной онтологией, можно представить как следующую систему:

$$M^C = \langle T, O^{CA}, A \rangle, \text{ где}$$

T – класс, предназначенный для представления текстов, включая информацию о его источнике, жанровой принадлежности и набор мета-атрибутов (*автор, дата\_публикации* и т.п.);

$O^{CA} = CS \cup CA \cup C^R \cup P, O^{CA} \in O^C$  – часть онтологии аргументации, содержащая набор классов размечаемых сущностей:

$$CS = UCS_i \text{ – конечное непустое множество классов утверждений,}$$

$CA = UCA_i$  – конечное непустое множество схем (классов) аргументов,

$C^R = UC^R_i$  – конечное непустое множество классов конфликтных отношений;

$P = DT \cup R \cup RC$  – множество свойств классов онтологии, где

$DT$  – конечное множество атрибутов простых типов, данных  $V$ , и

$F^{DT}: CSUCAUC^R \rightarrow 2^{DT \times V}$  – функция, которая определяет имена и типы атрибутов для классов онтологии,

$R = Ua_i$  – конечное множество ролей (атрибутов) в структуре схем аргументов, и

$F^R: CA \rightarrow 2^{R \times (CSUCA)}$  – функция, которая определяет имена и типы ролей для классов аргументов (отметим, что значением атрибута (роли) в аргументе могут выступать как утверждения, так и другие аргументы);

$RC = Urc_i$  – конечное множество связей конфликтных отношений, и

$F^{RC}: C^R \rightarrow 2^{R \times (CSUCA)}$  – функция, которая определяет имена и типы ролей для конфликтных отношений (конфликт может быть, как между утверждениями, так и между утверждением и аргументом);

$A = UA_t$  – множество аннотаций и  $F^A: T \rightarrow 2^A$  – функция, сопоставляющая тексту множество его аннотаций  $F^A(t) \subseteq A$ .

Отличительной особенностью рассматриваемой модели аннотирования текста в сравнении с другими схемами представления аргументации (например, AIF) является то, что в предлагаемую схему явным образом вводятся компоненты для разметки индикаторов. Таким образом, аннотация текста представляется системой вида:

$A_t = \langle Fr_t, S_t, Arg_t, Ind_t \rangle$ , где

$Fr_t = Ufr_i$  – множество размеченных текстовых фрагментов (в общем случае фрагмент может быть разрывным и представляться мультиинтервалом);

$S_t = Us_i$ ,  $s_i \in CS_i$  – множество размеченных утверждений:

$s_i = \langle desc, ws \rangle$ , где

$desc$  – выраженная на естественном языке пропозиция, или пропозициональное содержание текстового фрагмента, размеченного в тексте, и

$F^S: S \rightarrow 2^{Fr}$  – функция, сопоставляющая утверждению множество фрагментов текста (в общем случае утверждение может несколько раз встречаться в тексте),

$ws$  – числовая характеристика, сопоставленная утверждению и характеризующая степень уверенности автора в его истинности или достоверности.

$Arg_t = Ua_i$  – множество аргументов, найденных в тексте, где  $a_i$  класса  $CA_i \in CA$  имеет набор ролей  $r_i$ , связывающих аргумент с утверждениями и другими аргументами:

$a_i = U_j(r_j, s_j)U_k(r_k, a_k)$ , так, что  $(r_j, s_j) \in F^{SA}(CA_i)$ ,  $(r_k, a_k) \in F^{SA}(CA_i)$ .

$Ind_t = U(ind_i, tm_i)$ ,  $ind_i \in M$ ,  $tm_i \in T_m$  – множество индикаторов  $ind_i$  различных аспектов аргументации  $tm_i$ , и

$R^{AM}: ArgUS \rightarrow 2^{M \times T_m}$  – функция, сопоставляющая аргументу или утверждению, входящему в состав аргумента, набор индикаторов и аспектов аргументации.

Множество  $T_m$  определяет различные типы соотношения индикатора и сигнализируемых им аспектов аргументации:

- 1) тип аргументативного отношения (поддержка vs. конфликт);
- 2) структура аргументации (множественная vs. последовательная аргументация);

- 3) сила или убедительность аргумента;
- 4) отношение вывода между двумя утверждениями (наличие аргументации);
- 5) роль утверждения в отношении вывода (посылка vs. заключение);
- 6) семантико-онтологическое отношение, на котором основана применяемая в данном случае типовая схема рассуждения;
- 7) степень уверенности автора в утверждении.

Аспекты 1-4 сигнализируют о наличии аргументации, что в случае ручной разметки фокусирует внимание эксперта на соответствующих фрагментах текста, а при автоматической – требует наличия других показателей для более точного выявления класса аргумента. Аспекты 4-6 соотносят индикатор не столько с аргументом, сколько с утверждением (или несколькими утверждениями) в структуре аргумента. Аспекты 3, 7 влияют на оценку убедительности аргументации.

Таким образом, аннотация состоит из множества фрагментов текста, каждому из которых поставлено в соответствие утверждение или индикатор. Аргументы определяются как n-местные отношения над размеченными утверждениями. Роль индикаторов заключается в фиксации свойств и границ аргументов и/или утверждений, входящих в их состав.

**2. Индикаторы.** Индикаторы аргументации – это используемые в дискурсе языковые средства (слова и конструкции), которые служат подсказками при определении структуры аргументации: они помогают установить наличие аргументов и их составляющих в данном сегменте текста, реконструировать связи между утверждениями, соотносят аргумент с определенной схемой рассуждения (формой вывода, выражающей взаимосвязь посылок и заключений).

Индикаторы различаются не только по типу семантической и прагматической информации (что отмечено выше), но и по степени грамматикализации (первичные и вторичные индикаторы, см. [11]), по семантике опорного знаменательного слова (например, в них используются такие лексико-семантические классы, как речевые и ментальные предикаты, предикаты вывода и ментального воздействия, модальные слова, выражающие различные степени убежденности в высказываемом мнении) и по типу конструкции (см. классификацию и примеры в [7]).

**2.1. Язык описания индикаторов.** Множество индикаторов, образующих словарь индикаторов, можно создать с помощью специального синтаксиса, формальное описание которого в расширенной форме Бэкуса-Наура представлено ниже:

```
<словарь индикаторов> ::= {<шаблон>}
<шаблон> ::= <имя индикатора> '=' (<альтернатива> ['->' <атрибуты>])+ ['=>' <атрибуты>]
<альтернатива> ::= <цепочка> | <разрыв>
<цепочка> ::= '[' <индикатор> (';' <индикатор>) ']'
<разрыв> ::= '[' <ограничение> ':' <индикатор> {';' <ограничение> ':' <индикатор>} ']'
<индикатор> ::= (<имя индикатора> | <простой индикатор>) [<ограничения на атрибуты>]
<простой индикатор> ::= <тип простого индикатора> '/' <описание простого индикатора>
<тип простого индикатора> ::= 'w' | 'ph' | 't' | 's'
<ограничение> ::= 'begin' | 'end' | 'include' | 'not_include' | 'begin!' | 'end!' | 'main'
<ограничения на атрибуты> ::= <атрибуты>
<атрибуты> ::= '<' <атрибут> {';' | '><'} <атрибут> '>'
```

**<атрибут>** ::= <имя атрибута> [=] <значение> {, <значение>}

Можно выделить два основных способа задания индикаторов: индикаторы с простой структурой *<простой индикатор>* и составные конструкции, описываемые с помощью шаблонов *<шаблон>*. Словарь индикаторов состоит из множества шаблонов, которые в своем описании используют индикаторы с простой структурой, а также ссылаются на другие составные индикаторы.

К индикаторам с простой структурой относятся однословные и многословные дискурсивные коннекторы, и полнозначная лексика (предикаты вывода и причины, речевые и ментальные предикаты и т.п.).

Описание составных индикаторов включает несколько альтернатив, каждая из которых может быть цепочкой индикаторов *<цепочка>* или разрывной конструкцией *<разрыв>*. Любой из альтернатив (или сразу всем) при желании можно приписать грамматические или семантические значения *<атрибуты>*. Для простых индикаторов значения атрибутов извлекаются автоматически при анализе текста. Полученные значения атрибутов можно в дальнейшем использовать для фильтрации вхождений *<ограничения на атрибуты>*.

**3. Индикаторы как инструмент исследования аргументации.** Для поддержки исследования аргументации была разработана платформа, которая включает набор веб-инструментов, обеспечивающих создание корпусов текстов, визуализацию используемой онтологии аргументации (т.е. схем аргументации и их метаописаний), построение графа аргументации, поиск различных сущностей в корпусах текстов в терминах онтологии и т.п.

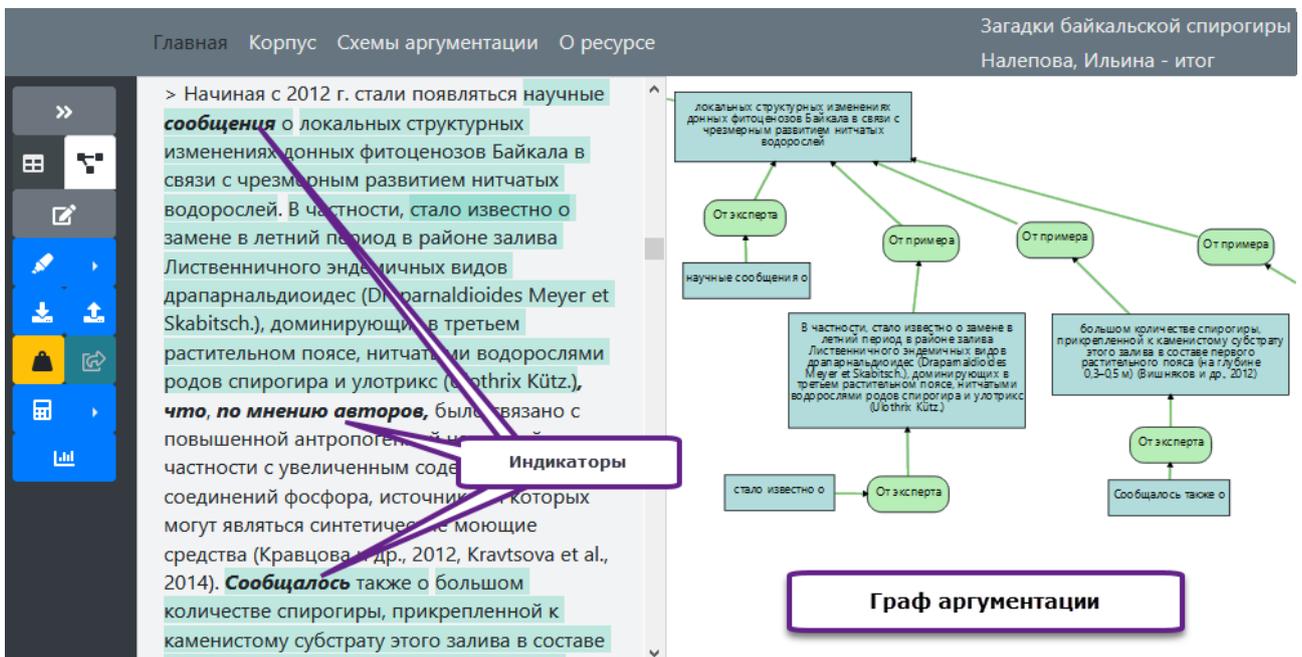
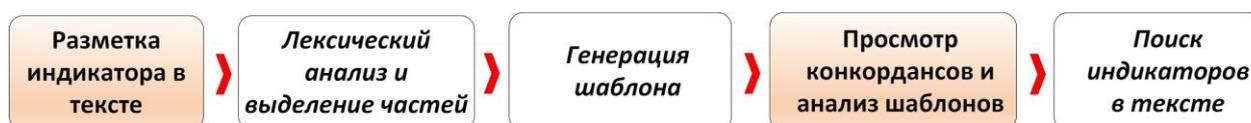


Рис. 2. Выделение индикаторов в аргументативной разметке текста.

В рамках данного исследования были созданы дополнительные инструменты для поиска и исследования аргументации в текстах на основе индикаторов. Созданные инструменты обеспечивают следующие функциональные возможности:

- подключение словаря индикаторов аргументации, созданного на основе риторических маркеров, и автоматический поиск и выделение в текстах индикаторов аргументации в режиме аргументативной разметки текста и построения графа аргументации (рис. 2),
- исследование адекватности выявления особенностей аргументации с помощью индикаторов на основе корпусного исследования и построения конкордансов.

**3.1. Формирование словаря индикаторов.** Формирование словаря индикаторов происходит на базе ручной разметки индикаторов. На рис.3 приведены основные этапы процесса создания и исследования индикаторов (блоки со светлым фоном сопоставлены с полностью автоматическими процедурами, блоки с темным фоном - с процедурами, осуществляемыми экспертом).



**Рис. 3.** Основные этапы формирования словаря индикаторов.

Автоматически можно сформировать шаблоны для первичных индикаторов (функциональных слов) и задать базовое структурное определение шаблонов для вторичных индикаторов, включающих опорное слово. Далее полученное формальное описание индикаторов исследуется экспертом. С этой целью осуществляется поиск индикаторов в тексте, построение конкорданса.

Для исследования индикаторов аргументации были созданы веб-инструменты, обеспечивающие:

- визуализацию словаря индикаторов, просмотр состава индикатора, поиск индикаторов и/или их составных частей в тексте,
- создание новых индикаторов и пополнение словаря индикаторов пользователем,
- поиск индикаторов и/или их частей в выбранном(ых) корпусе, построение конкордансов и визуализация расширяемых контекстов.

**3.2. Пополнение словаря.** Список знаменательных слов, которые могут служить индикаторами или опорными словами индикаторных конструкций, разнороден и принципиально неполон, как далек от полноты и словарь шаблонов индикаторных конструкций. Поэтому при разметке аннотатор сталкивается с индикаторными фрагментами, не покрытыми текущим словарем индикаторов. В этом случае у аннотатора есть возможность создать пользовательский шаблон, выделив соответствующий фрагмент текста (возможно, разрывный).

На основе размеченного аннотатором фрагмента автоматически генерируется шаблон, в котором зафиксированы состав и нормализованная форма лексических единиц, знаки препинания и разрывы (если пользователь выделил разрывный фрагмент).

Так, на рис.4 в числе прочих представлен пользовательский шаблон *президент ...отметил, что*, который не был автоматически выделен идентичным по структуре шаблоном *expert\_opinion*, поскольку слово *президент* не было заранее отнесено к лексико-

семантическому классу **expert**. В данном случае, не являясь авторитетным представителем лингвистической науки, президент, как представитель власти, является экспертом в области языковой политики, о которой идет речь в аннотируемом тексте, так что ввод данного пользовательского шаблона вполне правомерен и шаблон удален не будет.

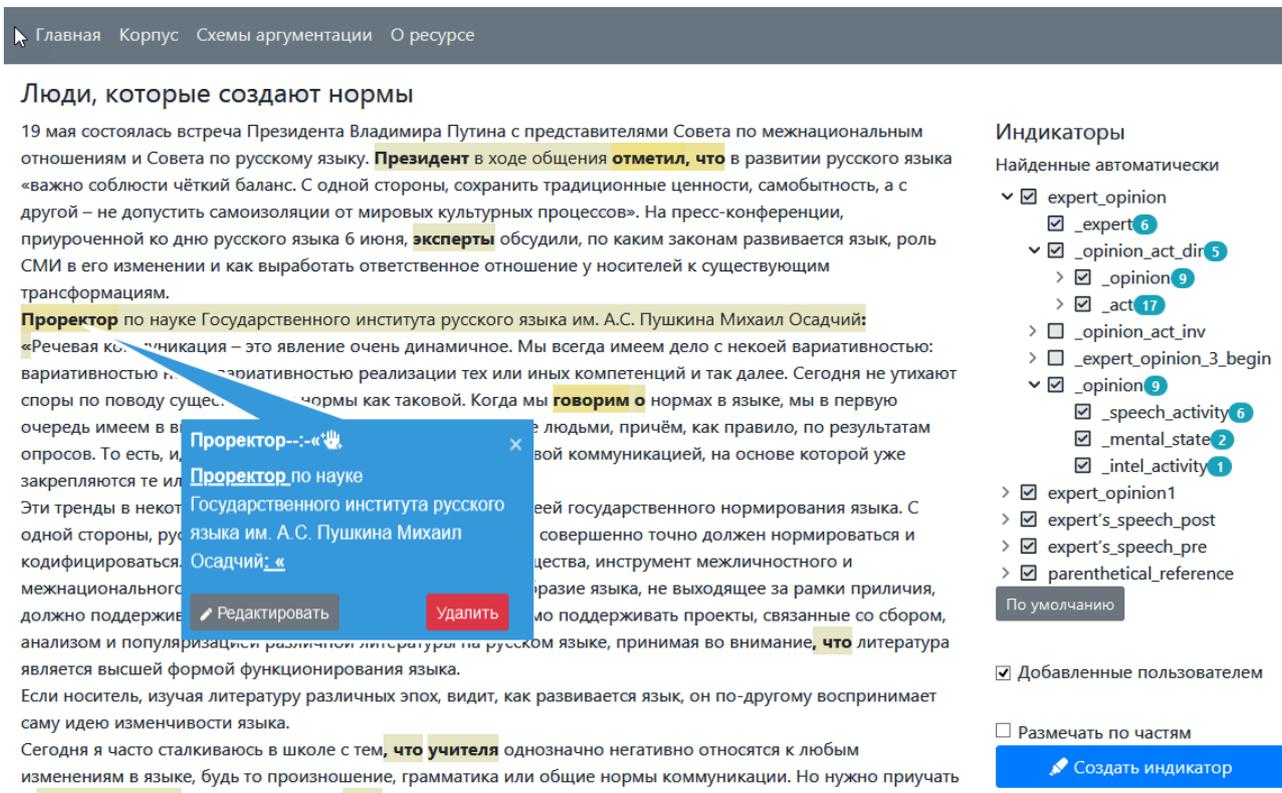


Рис. 4. Исследование индикаторов в тексте и пополнение словаря.

Кроме того, рис.4 демонстрирует еще один пользовательский шаблон *проректор по науке*, который также не был автоматически выделен в качестве индикатора типа экспертного мнения, но по другой причине – ввиду отсутствия речевого глагола. Такая структура (использующая пунктуацию прямой речи без речевого глагола) характерна для жанра интервью и в таком жанре имеет полное право на существование: соответствующий шаблон является кандидатом на пополнение словаря индикаторов.

**3.3. Корпусное исследование индикаторов.** Для расширения и обобщения лексического состава шаблона проводится анализ конкорданса (множества всех контекстов) шаблона и его частей.

При проведении исследований могут выполняться следующие шаги:

- формирование гипотез об обобщении отдельных опорных слов и создание лексико-семантических классов (так, в результате анализа аргументов, представляющих экспертное мнение, все глаголы, лежащие в основе конструкций прямой и косвенной речи, объединены в лексико-семантический класс речевых предикатов);
- проверка по конкордансу гипотезы об обобщении опорных слов разных классов в рамках одного шаблона (так, конкорданс показывает функциональное тождество и общность конструкций, формируемых речевыми и ментальными предикатами, что означает целесообразность объединения этих классов в рамках одного шаблона **\_opinion**);

- создание служебных шаблонов с альтернативами (так, рассмотрение контекстов уже введенных лексических индикаторов **\_opinion**, представленных речевыми и ментальными предикатами, позволяет увидеть альтернативные способы заполнения актантных позиций, т.е. средства ввода актанта, представляющего экспертное мнение, и сформировать варианты служебных шаблонов для прямой и инверсной актантных позиций, **\_opinion\_act\_dir**, см. рис. 5, и **\_opinion\_act\_inv**, которые далее выступают в качестве подшаблонов альтернативных вариантов полной индикаторной конструкции **expert\_opinion**);
- корректировка и проверка всех вхождений созданных индикаторов в корпус.

The screenshot shows a search interface with the following components:

- Parameters of search:** A tree view on the left with checkboxes for various indicators like `expert_opinion`, `_expert`, `_opinion_act_dir`, `_opinion_act_inv`, `_expert_opinion_3_begin`, `_opinion`, `expert_opinion1`, `expert's_speech_post`, `expert's_speech_pre`, and `parenthetical_reference`. A search button labeled "Искать" is visible.
- Search Results:** A table with columns: "Левый контекст", "Вхождение", "Правый ко...", "Текст", "Шаблон", and "Кор...". It lists several search results with their respective contexts and indicators.
- Corpus:** A list of sources on the left, including "Naked Science \*", "National Geographic", "STRF 76" (selected), "Антинаучные", "Дилетант", "Земля. Хроники Жизни", and "Лингвистический".
- Navigation:** Buttons for "Назад", "Страница 2 из 8", "10 строк", and "Вперед".

Рис. 5. Поиск индикаторов в корпусе текстов.

**Заключение.** Данное исследование проводилось в рамках проекта по созданию программного комплекса, предназначенного для поддержки исследования аргументации в русскоязычных научно-популярных текстах.

Предложено представление индикаторов аргументации в виде лексических единиц и лексико-грамматических шаблонов. Процедура автоматизированного создания словаря индикаторов аргументации позволяет на основе размеченного текстового фрагмента автоматически сформировать предварительную формальную спецификацию индикатора и уточнить ее вручную по результатам поиска в корпусе и анализа всех соответствующих ей фрагментов. Полученные индикаторы используются для поиска и автоматического выделения индикаторов аргументации в неразмеченном тексте. Это облегчает процесс аннотирования и исследования текстов, поскольку привлекает внимание эксперта-аннотатора к определенным фрагментам текста, потенциально содержащим аргументы.

Кроме того, проведены эксперименты в области автоматического распознавания аргументации, в которых индикаторы, извлекаемые с помощью достаточно надежных лексико-грамматических шаблонов (демонстрирующих точность более 70 %, см. [1]), используются в качестве признаков для машинного обучения при решении задачи автоматического извлечения аргументов [5]. В работе [10] предложен подход к частичному восстановлению аргументативной структуры текста, который может быть реализован при

недостатке размеченных коллекций. Утверждения, содержащие аргументы, выявляются автоматически на базе индикаторов аргументации. В условиях дефицита больших размеченных коллекций, с пополнением арсенала индикаторных шаблонов и расширением словарей индикаторов данный метод показывает себя вполне работоспособным.

Статья подготовлена по итогам исследования, проведённого в рамках проекта Российского фонда фундаментальных исследований № 18-00-01376 (18-00-00889).

#### СПИСОК ЛИТЕРАТУРЫ

1. Ахмадеева И.Р., Кононенко И.С., Саломатина Н.В., Сидорова Е.А. Подход к построению шаблонов индикаторов для извлечения аргументов из научно-популярных текстов // Тр. Международной конференции “Знания - Онтологии - Теории” (ЗОНТ-2019). Институт математики им. С. Л. Соболева СО РАН. Новосибирский государственный университет. 2019. С. 24-32.
2. Блюменау Д.И., Гендина Н.И., Добронравов И.С., Лахути Д.Г., Леонов В.П., Федоров Е.Б. Формализованное реферирование с использованием словесных клише (маркеров) // Научно-техническая информация. Сер.2. 1981. №2. С. 16–20.
3. Загорюлько, Ю.А., Гаранина Н.О., Боровикова О.И., Доманов О.А. Моделирование аргументации в научно-популярном дискурсе с использованием онтологий // Онтология проектирования. 2019. Т. 9. №4 (34). С.496-509.
4. Саломатина Н.В., Гусев В.Д. Автоматизация формирования индикаторных словарей и возможности их использования // Труды межд. конференции Диалог-2006 «Компьютерная лингвистика и интеллектуальные технологии». Бекасово. 2006. Москва: "Наука". С. 121–125.
5. Akhmadeeva I., Kononenko I., Salomatina N., Sidorova E. Indicator Patterns as Features for Argument Mining // 2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). Novosibirsk. Russia. 2019. Pp. 0886-0891.
6. Chesñevar C.I., McGinnis J., Modgil S., Rahwan I., Reed C., Simari G., South M., Vreeswijk G., Willmott S. Towards an argument interchange format. The knowledge engineering review 2006. № 21(4). Pp. 293-316.
7. Kononenko I., Sidorova E. Development of the Lexicon of Argumentation Indicators. In: Kuznetsov S., Panov A. (eds) Artificial Intelligence. RCAI 2019. Communications in Computer and Information Science. vol 1093. Springer. Cham. 2019. Pp. 154-168.
8. Musi E., Ghosh D., Muresan S. Towards Feasible Guidelines for the Annotation of Argument Schemes// Proceedings of the third workshop on argument mining (ArgMining 2016). Pp. 82-93.
9. Rahwan I, Banihashemi B, Reed C, Walton D, Abdallah S. Representing and classifying arguments on the semantic web. The Knowledge Engineering Review 2011. №26(4). Pp.487-511.
10. Salomatina N.V., Kononenko I.S., Sidorova E.A., Pimenov I.S. Identification of connected arguments based on reasoning schemes “from expert opinion” // Journal of Physics: Conference Series (в печати).
11. Toldova S., Pisarevskaya D., Vasilyeva M., Kobozeva M. The cues for rhetorical relations in Russian: "Cause-Effect" relation in Russian Rhetorical Structure Treebank // Computational

Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue”. 2018. Issue 17. Pp.747-761.

12. Van Eemeren F.H., Houtlosser P., and F. Snoeck Henkemans. Argumentative Indicators in Discourse: A Pragma-Dialectical Study // Dordrecht: Springer. 2007.

13. Walton D, Reed C, Macagno F. Argumentation schemes. Cambridge: Cambridge University Press. 2008.

---

**UDK 004.82:004.89:519.816**

## **LINGUISTIC ASPECTS OF ONTOLOGY-BASED ARGUMENTATION STUDY**

**Irina S. Kononenko**

Researcher, e-mail: [irina\\_k@cn.ru](mailto:irina_k@cn.ru),

**Irina R. Akhmadeeva**

Junior researcher, e-mail: [i.r.akhmadeeva@iis.nsk.su](mailto:i.r.akhmadeeva@iis.nsk.su),

**Elena A. Sidorova**

PhD, Senior researcher, e-mail: [lsidorova@iis.nsk.su](mailto:lsidorova@iis.nsk.su),

A.P. Ershov Institute of Informatics Systems

630090, Russia, Novosibirsk Acad. Lavrentjev pr., 6.

**Abstract.** Annotation of the corpus of Russian-language popular science texts is discussed. Particular attention is paid to the annotation scheme, in which the objects of annotation are not only pragmatic and semantic aspects of argumentation, but also their linguistic indicators. The presentation of argumentation indicators in the form of lexical and lexical-grammatical patterns is proposed, which are automatically generated from a marked-up fragment of text, and then, if necessary, are manually specified by an expert based on the results of a search in the corpus and analysis of all relevant fragments. The resulting dictionary of indicators is used to search and automatically highlight indicators of argumentation in unannotated text.

**Keywords:** ontology of argumentation, argumentation annotation, argumentation scheme, argumentation indicator, ontological model of text markup, lexicon of indicators.

### **References**

1. Akhmadeeva I.R., Kononenko I.S., Salomatina N.V., Sidorova E.A. Podhod k postroeniju shablonov indikatorov dlja izvlechenija argumentov iz nauchno-populjarnyh tekstov [Approach to building indicator patterns for extracting arguments from popular science texts] // Tr. Mezhdunarodnoj konferencii “Znanija - Ontologii - Teorii” (ZONT-2019) = International Symposium “Knowledge - Ontology - Theory” (KONT-19). Sobolev Institute of Mathematics. Novosibirsk State University. 2019. C. 24-32. (in Russian).
2. Bljumenau D.I., Gendina N.I., Dobronravov I.S., Lahuti D.G., Leonov V.P., Fedorov E.B. Formalizovannoe referirovanie s ispol'zovaniem slovesnyh klishe (markerov) [Formalized abstracting using verbal cliches (markers)] // Nauchno-tehnicheskaja informacija = Scientific and technical information. Vol.2. 1981. №2. Pp. 16–20.1. (in Russian).

3. Zagorulko, Yu.A., Garanina N.O., Borovikova O.I., Domanov O.A. Modelirovanie argumentacii v nauchno-populjarnom diskurse s ispol'zovaniem ontologij [Argumentation modeling in popular science dis-course using ontologies] // *Ontologija proektirovanija = Ontology of designing*. 2019. T. 9. №4 (34). Pp. 496-509 (in Russian).
4. Salomatina N.V., Gusev V.D. Avtomatizacija formirovanija indikatornyh slovarej i vozmozhnosti ih ispol'zovanija [Automation of cue dictionaries formation and their applications] // *Trudy mezhd. konferencii Dialog-2006 «Komp'juternaja lingvistika i intellektual'nye tehnologii»*. Bekasovo = Dialogue-2006 “Computational Linguistics and Intellectual Technologies”. Bekasovo. Moscow. "Nauka". 2006. Pp. 121–125. (in Russian).
5. Akhmadeeva I., Kononenko I., Salomatina N., Sidorova E. Indicator Patterns as Features for Argument Mining // 2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). Novosibirsk, Russia. 2019. Pp. 0886-0891.
6. Chesñevar C.I., McGinnis J., Modgil S., Rahwan I., Reed C., Simari G., South M., Vreeswijk G., Willmott S. Towards an argument interchange format. *The knowledge engineering review* 2006. № 21(4). Pp. 293-316.
7. Kononenko I., Sidorova E. Development of the Lexicon of Argumentation Indicators. In: Kuznetsov S., Panov A. (eds) *Artificial Intelligence. RCAI 2019. Communications in Computer and Information Science*, vol 1093. Springer. Cham. 2019. Pp. 154-168.
8. Musi E., Ghosh D., Muresan S. Towards Feasible Guidelines for the Annotation of Argument Schemes// *Proceedings of the third workshop on argument mining (ArgMining 2016)*. Pp. 82-93.
9. Rahwan I, Banihashemi B, Reed C, Walton D, Abdallah S. Representing and classifying arguments on the seman-tic web. *The Knowledge Engineering Review* 2011. №26(4). Pp.487-511.
10. Salomatina N.V., Kononenko I.S., Sidorova E.A., Pimenov I.S. Identification of connected arguments based on reasoning schemes “from expert opinion” // *Journal of Physics: Conference Series*. (to be published).
11. Toldova S., Pisarevskaya D., Vasilyeva M., Kobozeva M. The cues for rhetorical relations in Russian: "Cause-Effect" relation in Russian Rhetorical Structure Treebank // *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue”*. 2018. Issue 17. Pp.747-761.
12. Van Eemeren F.H., Houtlosser P., and F. Snoeck Henkemans. *Argumentative Indicators in Discourse: A Pragma-Dialectical Study* // Dordrecht: Springer. 2007.
13. Walton D, Reed C, Macagno F. *Argumentation schemes*. Cambridge: Cambridge University Press. 2008.

## ОБРАБОТКА СИНХРОНИЗИРОВАННЫХ ВЕКТОРНЫХ ИЗМЕРЕНИЙ МЕТОДАМИ НЕЧЕТКОЙ ЛОГИКИ ПРИ КИБЕРАТАКАХ НА ЭЛЕКТРОЭНЕРГЕТИЧЕСКУЮ СИСТЕМУ

Колосок Ирина Николаевна

Д.т.н., ведущий научный сотрудник лаборатории «Управление функционированием электроэнергетических систем», e-mail: [kolosok@isem.irk.ru](mailto:kolosok@isem.irk.ru),

Гурина Людмила Александровна

К.т.н., доцент, старший научный сотрудник лаборатории «Управление функционированием электроэнергетических систем», e-mail: [gurina@isem.irk.ru](mailto:gurina@isem.irk.ru),

Институт систем энергетики им. Л.А. Мелентьева СО РАН,  
664130 г. Иркутск, ул. Лермонтова 130

**Аннотация.** Применение интеллектуальных и цифровых технологий в системах измерения, передачи и обработки информации, являющихся частью информационно-коммуникационной инфраструктуры электроэнергетической системы (ЭЭС), направлено на повышение эффективности управления ЭЭС. Вместе с тем, при цифровизации ЭЭС отмечен рост уязвимостей к кибератакам. В связи с этим возрастает актуальность обеспечения задач управления в условиях кибератак своевременной, полной и достоверной информацией. Предложен алгоритм обработки синхронизированных векторных измерений на основе вейвлет-анализа и нечеткой логики, повышающий качество информации, используемой в дальнейшем при оценивании состояния ЭЭС.

**Ключевые слова:** качество данных, синхронизированные векторные измерения, кибератаки, вейвлет-анализ, нечеткая логика.

**Цитирование:** Колосок И.Н., Гурина Л.А. Обработка синхронизированных векторных измерений методами нечеткой логики при кибератаках на электро-энергетическую систему // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 56-63. DOI: 10.38028/ESI.2020.20.4.005

**Введение.** Внедрение систем мониторинга переходных режимов (за рубежом такие системы называют WAMS - Wide Area Measurement Systems) позволило при управлении ЭЭС наряду с традиционными измерениями SCADA<sup>1</sup> использовать синхронизированные векторные измерения (СВИ), поступающие от устройств PMU<sup>2</sup>. Благодаря более точным измерениям модулей и фаз напряжений в узлах, модулей и фаз токов в ветвях, уровень наблюдаемости и управляемости ЭЭС существенно возрастает. При оценивании состояния ЭЭС использование СВИ повышает точность получаемых оценок. Наряду с этим следует отметить уязвимость не только WAMS, но и всей информационно-коммуникационной инфраструктуры ЭЭС, частью которой является WAMS, к новым угрозам, возникающим в

<sup>1</sup> SCADA (аббр. от англ. Supervisory Control And Data Acquisition — диспетчерское управление и сбор данных) — программный пакет, предназначенный для разработки или обеспечения работы в реальном времени систем сбора, обработки, отображения и архивирования информации об объекте мониторинга или управления.

<sup>2</sup> PMU (аббр. от англ. Phasor Measurement Unit) – прибор, измеряющий комплексные величины тока и напряжения. В отличие от традиционных телеизмерений измерения от PMU синхронизированы по времени через GPS, точность их выше и поступают они в пункты сбора информации тысячами срезов в секунду, тогда как SCADA принимает один срез в несколько секунд

современном киберпространстве [9]. Приобретение ЭЭС киберфизического характера при цифровой трансформации ее свойств расширяет возможности реализации кибератак на информационно-коммуникационную инфраструктуру из-за наличия большого числа уязвимостей на всех уровнях иерархии управления ЭЭС [1]. В этих условиях обеспечение задач управления своевременной, полной и достоверной информацией, для сохранения устойчивого и надежного функционирования ЭЭС, особенно актуально.

Целью данной работы является разработка мер по повышению качества синхронизированных векторных измерений, нарушенного успешно проведенными кибератаками на информационно-коммуникационную инфраструктуру ЭЭС [6-8]. Под качеством информации понимается степень ее полноты и достоверности [3].

Отмечен целый класс кибератак, нарушающих такие свойства кибербезопасности информационно-коммуникационной инфраструктуры ЭЭС, как целостность и доступность [4]. Постоянное усовершенствование кибератак, чтобы быть не обнаруженными, требует пересмотра существующих методов обработки информации, необходимой для формирований управленческих решений.

Оценивание состояния (ОС) является ключевой задачей при управлении ЭЭС. При ее решении выполняется анализ наблюдаемости, обнаружение плохих данных, определяются оптимальные оценки параметров режима [2]. Разработка методов обработки данных, создающих барьер на пути распространения влияния кибератаках, не обнаруживаемые традиционными методами обнаружения плохих данных (ОПД), позволит повысить точность оценок измерений.

Использование только вероятностных методов обработки информации не всегда успешно из-за возникновения неопределенности данных, вызванной кибератаками. В этом случае предлагается совместное использование вейвлет-анализа и нечеткой логики.

**Алгоритм обработки измерений.** Последствием успешно проведенной кибератаки, например, атаки внедрения ложных данных, при решении оценивания состояния ЭЭС может быть выдача ложных оценок состояния, которые приведут к неправильным управленческим решениям, которые могут вызвать крупномасштабные аварии при функционировании ЭЭС. При этом возникает неопределенность данных измерений, ее характерными признаками являются неполнота и недостоверность. Важно обнаруживать и предотвращать влияние кибератак на качество информации.

Применение вейвлет-анализа и теории нечетких множеств при обработке данных измерений, как предварительный этап ОС ЭЭС, позволит существенно повысить как уровень достоверности, так и уровень полноты информации.

Так как традиционные методы оценивания состояния основаны на вероятностных предположениях об ошибках в измерениях и для надежной оценки состояния зачастую требуется избыточность измерений, применение нечеткой логики позволит использовать интервальный анализ данных и, тем самым, устранить проблему неполных данных.

Общая модель оценивания состояния представляется как

$$y = h(X) + \varepsilon, \quad (1)$$

где  $y$  - вектор измерений с  $m$  элементами;  $X$  - вектор переменных состояния с  $n$  элементами;  $h(.)$  – вектор-функция, который связывает переменные состояния и измерения ( $m$  функций);  $\varepsilon$  - вектор шума измерения.

Нечеткая постановка задачи оценивания состояния [11] предполагает, что, по меньшей мере, одно измерение представляется как нечеткое число.

Цель задачи состоит в минимизации взвешенной суммы квадратов ошибок, согласно выражению:

$$\min \varepsilon^T R^{-1} \varepsilon. \quad (2)$$

Уравнение (2) представляет задачу взвешенных наименьших квадратов, решение которой получается путем замены  $\varepsilon$  на выражение, полученное из (1). Эта задачи минимизации затем решается путем формирования системы уравнений

$$H(X)^T R^{-1} [y - h(X)] = 0, \quad (3)$$

где  $H(X)^T$  – транспонированная матрица Якоби,  $R$  – ковариационная матрица ошибок измерений.

Нечеткое число описывается, как

$$\tilde{Z} = \langle z_1, z_2, z_3, z_4 \rangle, \quad (4)$$

где  $\langle z_1, z_2, z_3, z_4 \rangle$  – кортеж для трапецидального нечеткого числа [10].

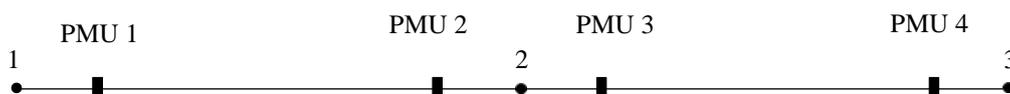
Параметры  $z_1, z_2, z_3, z_4$  определяются, исходя из следующих выражений

$$\begin{aligned} z_1 &= \min\{x_i\}, \\ z_2 &= M[x_i] - [M[x_i] - \min\{x_i\}]\sigma, \\ z_3 &= M[x_i] + [\max\{x_i\} - M[x_i]]\sigma, \\ z_4 &= \max\{x_i\}, \end{aligned} \quad (5)$$

где  $M[x_i]$  – математическое ожидание последовательности измерений,  $\sigma$  – среднее квадратическое отклонение последовательности измерений.

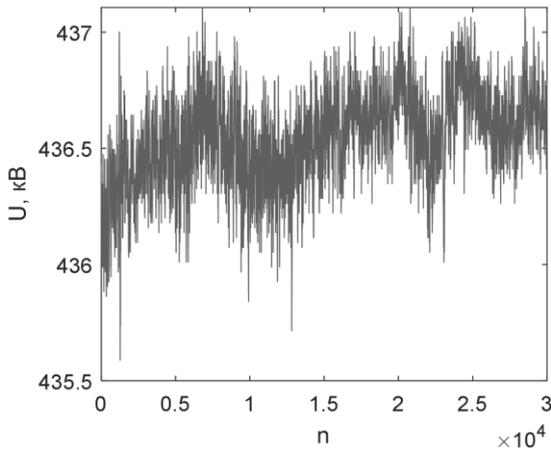
Если измерения содержат ошибки, вызванные кибератаками, возможно неточное определение параметров нечеткого числа. Для фильтрации ошибок и достоверизации данных предлагается проведение вейвлет-анализа потоков измерений [5].

**Пример.** Для демонстрации предложенного подхода обработки данных в качестве примера смоделирована кибератака внедрения ложных данных на устройства СВИ трехузловой схемы участка электрической сети (рис. 1).

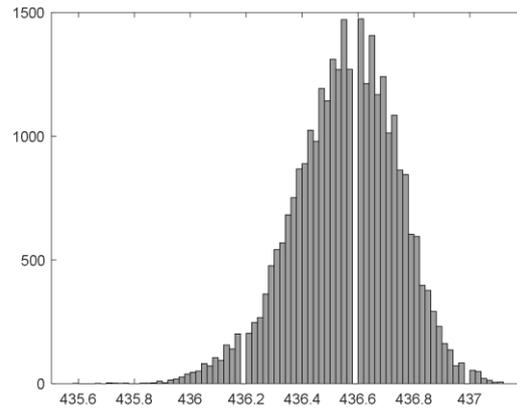


**Рис. 1.** Фрагмент схемы электрической сети.

На рис. 2, 3 представлены поток измерений напряжения и его гистограмма распределения по нормальному закону. Число измерений  $n = 30000$  с интервалом дискретизации  $\Delta t = 20$  мс. Результаты вейвлет-анализа показали, что измерения напряжения не содержат грубых ошибок.

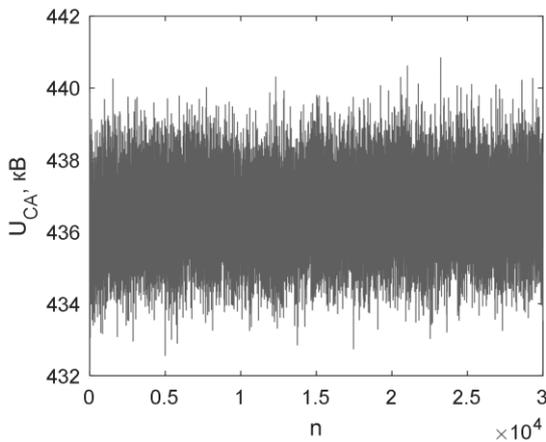


**Рис. 2.** Измерения напряжения, не содержащие грубых ошибок.

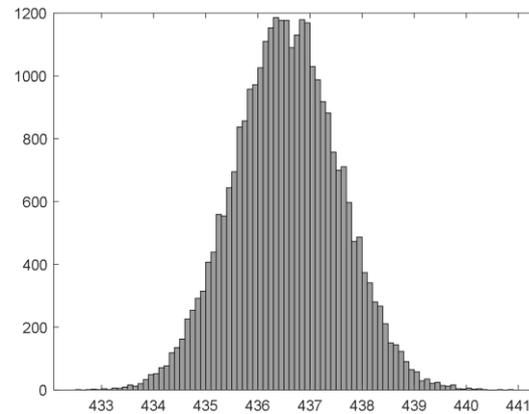


**Рис. 3.** Гистограмма распределения.

На поток измерения напряжения была сгенерирована атака в виде наложения шума  $a(t) = \xi_{КАИ}(t) \rightarrow N(0, \sigma_a^2)$  (рис. 4, 5).

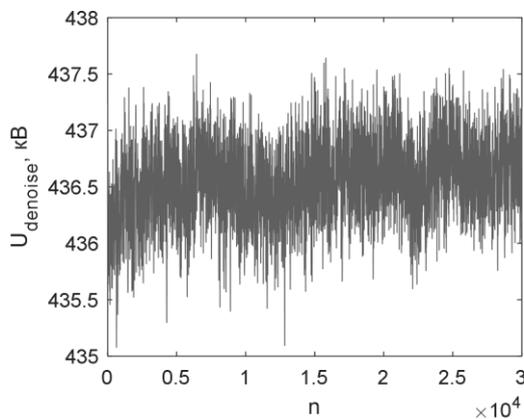


**Рис. 4.** Поток измерений напряжения при успешно реализованной кибератаке в виде наложения шума.

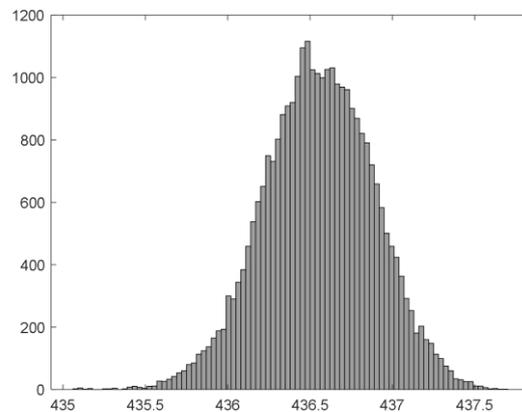


**Рис. 5.** Гистограмма распределения.

Достоверизация данных проведена путем применения вейвлет-преобразований (фильтрация шумов, восстановление) потока измерений напряжения  $U_{CA}$  (рис.6, 7).



**Рис. 6.** Восстановленный поток измерений напряжения.



**Рис. 7.** Гистограмма распределения.

В результате проведенного вейвлет-анализа получены характеристики (математическое ожидание  $m_U$ , минимальное  $min_U$  и максимальное  $max_U$  значения, среднее квадратическое отклонение  $\sigma_U$ ) изменения напряжения (табл. 1), требуемые для нахождения параметров нечетких чисел.

**Таблица 1.** Характеристики изменения напряжения

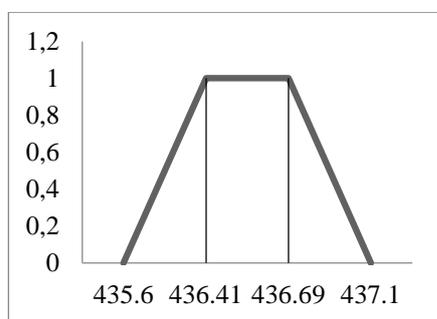
	$U$	$U_{CA}$	$U_{denoise}$
$m_U$	436.6	436.6	436.6
$min_U$	435.6	432.6	435.1
$max_U$	437.1	440.8	437.7
$\sigma_U$	0.19	1	0.33

Параметры нечетких чисел для потоков измерений напряжения  $\tilde{U}$ ,  $\tilde{U}_{CA}$ ,  $\tilde{U}_{denoise}$  (табл. 2) определялись согласно выражениям (1)-(4).

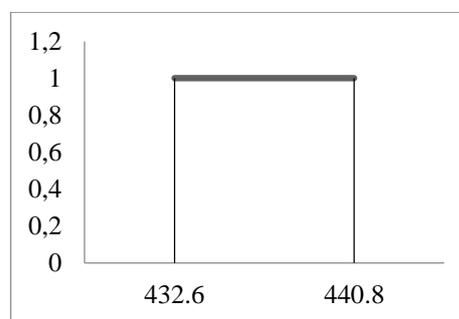
**Таблица 2.** Параметры нечетких чисел потоков измерения напряжения

	$U$	$U_{CA}$	$U_{denoise}$
$z_1$	435.6	432.6	435.1
$z_2$	436.4124	432.6	436.105
$z_3$	436.6938	440.8	436.96
$z_4$	437.1	440.8	437.7

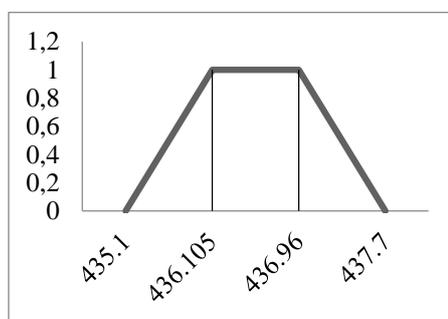
Согласно найденным параметрам нечетких чисел сформированы их трапецидальные функции принадлежности (рис. 8-10).



**Рис. 8.** Функция принадлежности нечеткого числа  $\tilde{U}$ .



**Рис. 9.** Функция принадлежности нечеткого числа  $\tilde{U}_{CA}$ .



**Рис. 10.** Функция принадлежности нечеткого числа  $\tilde{U}_{denoise}$ .

Сравнительный анализ полученных результатов (табл. 3) показал целесообразность применения вейвлет-анализа процессов изменения переменных режима при кибератаках на информационно-коммуникационную инфраструктуру, снижающих степень достоверности информации. Его применение позволяет устранить влияние искажающих факторов на параметры нечетких чисел, тем самым снижая неопределенность данных.

**Таблица 3.** Сравнительный анализ полученных результатов

	$\Delta = \tilde{U} - \tilde{U}_{CA}$	$\Delta = \tilde{U} - \tilde{U}_{denoise}$
$v_1$	3	0,5
$v_2$	3,8124	0,3074
$v_3$	-4,1062	-0,2662
$v_4$	-3,7	-0,6

Совокупное применение вейвлет-анализа и нечеткой логики при обработке данных увеличивает степень полноты и достоверности информации.

**Заключение.** Предложенный алгоритм обработки информации на основе вейвлет-анализа и нечеткой логики позволяет повысить точность измерений, необходимых для решения оценивания состояния ЭЭС и, тем самым, обеспечивая надежное управление ЭЭС. Использование нечеткой логики позволяет перейти к интервальным вычислениям при оценивании состояния ЭЭС в условиях неопределенности, вызванной потерей информации при кибератаках на информационно-коммуникационную инфраструктуру ЭЭС, что является предметом дальнейших исследований.

Работа выполнена в рамках научного проекта III.17.4.2. программы фундаментальных исследований СО РАН, рег. № АААА-А17-117030310438-1 и при частичной поддержке гранта РФФИ (№19-07-00351 А)

#### СПИСОК ЛИТЕРАТУРЫ

1. Воропай Н.И. Направления и проблемы трансформации электроэнергетических систем // *Электричество*. 2020. № 7. С. 12-21.
2. Гамм А.З., Колосок И.Н. Обнаружение грубых ошибок телеизмерений в электроэнергетических системах. – Новосибирск: Наука. 2000. 152 с.
3. Колосок И.Н., Гурина Л.А. Оценка качества данных SCADA и WAMS при кибератаках на информационно-коммуникационную инфраструктуру ЭЭС // *Информационные и математические технологии в науке и управлении*. 2020. № 1 (17). С. 68-78.
4. Колосок И.Н., Гурина Л.А. Оценка рисков управления киберфизической ЭЭС на основе теории нечетких множеств // *Методические вопросы исследования надежности больших систем энергетики: Вып. 70. Методические и практические проблемы надежности систем энергетики. В 2-х книгах. / Книга 1 /* отв. ред. Н.И. Воропай. Иркутск: ИСЭМ СО РАН. 2019. С. 238-247.
5. Kolosok I., Gurina L. Wavelet Analysis of PMU Measurements for Identification of Cyber Attacks on TCMS // *2018 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Moscow, Russia. 2018. Pp. 1-4. doi: 10.1109/ICIEAM.2018.8728768.*
6. Kyung Choi, Xinyi Chen, Shi Li, Mihui Kim, Kijoon Chae and JungChan Na. Intrusion Detection of NSM Based DoS Attacks Using Data Mining in Smart Grid // *Energies*. 2012. No 2. Pp. 4091-4109. doi:10.3390/en5104091.

7. L. Hu, Z. Wang, X. Liu, A. V. Vasilakos and F. E. Alsaadi. Recent advances on state estimation for power grids with unconventional measurements // in *IET Control Theory & Applications*. 2017. Vol. 11. № 18. Pp. 3221-3232. doi: 10.1049/iet-cta.2017.0629
  8. M. Yasinzadeh and M. Akhbari. Detection of PMU spoofing in power grid based on phasor measurement analysis // in *IET Generation, Transmission & Distribution*. 2018. Vol. 12. № 9. Pp. 1980-1987. doi: 10.1049/iet-gtd.2017.1445.
  9. S. Sridhar, A. Hahn and M. Govindarasu. Cyber-Physical System Security for the Electric Power Grid // in *Proceedings of the IEEE*. 2012. Vol. 100. № 1. Pp. 210-224. doi: 10.1109/JPROC.2011.2165269.
  10. Uluçay V., Deli I., Şahin M. Trapezoidal fuzzy multi-number and its application to multi-criteria decision-making problems // *Neural Comput & Applic*. 2018. № 30. Pp 1469–1478. <https://doi.org/10.1007/s00521-016-2760-3>
  11. V. Miranda, J. Pereira, and J. T. Saraiva. Load allocation in DMS with a fuzzy state estimator // *IEEE Trans. Power Syst.* 2000. vol. 15. № 2. Pp. 529–534.
- 

UDK 621.311 : 004.056

**PROCESSING OF SYNCHRONOUS PHASOR MEASUREMENTS BY FUZZY LOGIC METHODS IN THE CASE OF CYBERATTACKS ON INFORMATION-COMMUNICATION INFRASTRUCTURE OF A CYBER-PHYSICAL ELECTRIC POWER SYSTEM<sup>3</sup>**

**Irina N. Kolosok**

Professor, Leading Researcher of Laboratory of Electric Power Systems Operation and Control,  
e-mail: [kolosok@isem.irk.ru](mailto:kolosok@isem.irk.ru)

**Liudmila A. Gurina**

Doctor, Senior Researcher of Laboratory of Electric Power Systems Operation and Control,  
e-mail: [gurina@isem.irk.ru](mailto:gurina@isem.irk.ru)

Melentiev Energy Systems Institute

Siberian Branch of the Russian Academy of Sciences

130, Lermontov Str., 664033, Irkutsk, Russia

**Abstract.** The use of intelligent and digital technologies in systems for measuring, transmitting, and processing information, which are part of the information-communication infrastructure, is aimed at improving the efficiency of EPS management. At the same time, with the EPS digitalization, there has been a rise in vulnerabilities to cyberattacks. In this context, the urgency of providing timely, complete, and reliable data to perform control in the case of cyberattacks is increasing. We propose an algorithm to process synchronized vector measurements based on wavelet analysis and fuzzy logic, which improves the quality of information used in the state estimation of electric power systems.

**Keywords:** quality of data, PMU measurements, cyberattacks, wavelet analysis, fuzzy logic.

---

<sup>3</sup> This study is supported by the Siberian Branch of the Russian Academy of Sciences (Project III.17.4.2) of the Federal Program of Scientific Research (No. AAAAA-A17-117030310438-1)

## References

1. Voropaj N.I. Napravleniya i problemy transformacii elektroenergeticheskikh system [Prospects and problems of electric power system transformations] // *Elektrichestvo*. 2020. № 7. Pp. 12-21. (in Russian)
2. Gamm A.Z., Kolosok I.N. Obnaruzhenie grubyh oshibok teleizmerenij v elektroenergeticheskikh sistemah [Bad data detection in measurements in electric power systems]. Novosibirsk: Nauka = Science. 2000. 152 p.
3. Kolosok I.N., Gurina L.A. Ocenka kachestva dannyh SCADA i WAMS pri kiberatakah na informacionno-kommunikacionnyu infrastrukturu EES. Informacionnye i matematicheskie tekhnologii v nauke i upravlenii [Quality assessment of SCADA and WAMS data in the case of cyberattacks on information and communication infrastructure of EPS] // 2020. № 1 (17). Pp. 68-78. (in Russian)
4. Kolosok I.N., Gurina L.A. Ocenka riskov upravleniya kiberfizicheskoj EES na osnove teorii nechetkih mnozhestv [Risk control assessment of cyber-physical power system based on the theory of fuzzy sets] // *Metodicheskie voprosy issledovaniya nadezhnosti bol'shih sistem energetiki: Vyp. 70. Metodicheskie i prakticheskie problemy nadezhnosti sistem energetiki. V 2-h knigah. / Kniga 1 / otv. red. N.I. Voropaj. Irkutsk: ISEM SO RAN. 2019. Pp. 238-247.*
5. Kolosok I., Gurina L. Wavelet Analysis of PMU Measurements for Identification of Cyber Attacks on TCMS // 2018 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Moscow, Russia. 2018. Pp. 1-4. doi: 10.1109/ICIEAM.2018.8728768. (in Russian).
6. Kyung Choi, Xinyi Chen, Shi Li, Mihui Kim, Kijoon Chae and JungChan Na. Intrusion Detection of NSM Based DoS Attacks Using Data Mining in Smart Grid // *Energies*. 2012. № 2. Pp. 4091-4109. doi:10.3390/en5104091.
7. L. Hu, Z. Wang, X. Liu, A. V. Vasilakos and F. E. Alsaadi. Recent advances on state estimation for power grids with unconventional measurements // in *IET Control Theory & Applications*. 2017. Vol. 11. № 18. Pp. 3221-3232. doi: 10.1049/iet-cta.2017.0629
8. M. Yasinzadeh and M. Akhbari. Detection of PMU spoofing in power grid based on phasor measurement analysis // in *IET Generation, Transmission & Distribution*. 2018. Vol. 12. № 9. Pp. 1980-1987. doi: 10.1049/iet-gtd.2017.1445
9. S. Sridhar, A. Hahn and M. Govindarasu. Cyber-Physical System Security for the Electric Power Grid // in *Proceedings of the IEEE*. 2012. Vol. 100. № 1. Pp. 210-224. doi: 10.1109/JPROC.2011.2165269.
10. Uluçay V., Deli I., Şahin M. Trapezoidal fuzzy multi-number and its application to multi-criteria decision-making problems // *Neural Comput & Applic*. 2018. № 30. Pp 1469–1478. <https://doi.org/10.1007/s00521-016-2760-3>
11. V. Miranda, J. Pereira, and J. T. Saraiva. Load allocation in DMS with a fuzzy state estimator // *IEEE Trans. Power Syst*. 2000. vol. 15. № 2. Pp. 529–534.

## МЕТОД ОПРЕДЕЛЕНИЯ УРОВНЯ КИБЕРСИТУАЦИОННОЙ ОСВЕДОМЛЕННОСТИ ЭНЕРГЕТИЧЕСКИХ ОБЪЕКТОВ

Гаськова Дарья Александровна

м.н.с., e-mail: [gaskovada@gmail.com](mailto:gaskovada@gmail.com),

Институт систем энергетики им. Л.А. Мелентьева СО РАН,

664033, Россия, г. Иркутск, ул. Лермонтова, 130.

**Аннотация.** В статье рассматриваются результаты исследования ситуационной осведомленности в киберсреде энергетических объектов. Для повышения информированности о состоянии киберсреды таких объектов предлагаются модель сценариев экстремальных ситуаций в энергетике на основе байесовской сети доверия и численный метод определения киберситуационной осведомленности. Модель основана на причинно-следственных связях между уязвимостями локальной вычислительной сети (ЛВС) и возможных киберугрозах, представленных в виде векторов проникновения в ЛВС, векторов развития кибернетических атак и векторов атак на целевой актив, объединённых в сценарии. В работе ставится акцент на сценарии, последствия которых могут расцениваться как экстремальные ситуации в энергетике, вызванные киберугрозами.

**Ключевые слова:** байесовские сети доверия, экстремальные ситуации в энергетике, киберугрозы.

**Цитирование:** Гаськова Д.А. Метод определения уровня киберситуационной осведомлённости энергетических объектов // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 64-74. DOI: 10.38028/ESI.2020.20.4.006

**Введение.** Исследования кибербезопасности и киберситуационной осведомленности (КСО) в энергетике связаны с исследованиями критических инфраструктур (КИ). Это обусловлено рассмотрением энергетического сектора как одной из основных КИ и защитой ключевых информационных систем объектов КИ как направления кибербезопасности [16]. Нарушение функционирования или разрушение объектов энергетики приводит к экстремальным ситуациям (ЭКС), губительным последствиям и нарушению энергетической безопасности (ЭБ) как составляющей национальной безопасности (НБ) страны, а киберугрозы выделены как группа стратегических угроз в энергетике [4]. В Доктрине энергетической безопасности Российской Федерации [7], принятой в 2019 году, отражены официальные взгляды на обеспечение ЭБ в рамках стратегического планирования в сфере обеспечения НБ. Наряду с такими трансграничными угрозами ЭБ, как террористическая и диверсионная деятельность, неблагоприятные и опасные природные явления, выделены и угрозы противоправного использования информационно-телекоммуникационных технологий, в том числе осуществление компьютерных атак на объекты информационной инфраструктуры и сети связи, способные привести к нарушениям функционирования инфраструктуры объектов топливно-энергетического комплекса (ТЭК). С практической точки зрения наблюдается долгосрочная тенденция реализации кибератак на объектах энергетического сектора. В исследованиях [3] в 2020 году отмечаются достаточно выраженные различия в ландшафте угроз для всех рассматриваемых компьютеров автоматизированной системы управления

(АСУ) и компьютеров АСУ в энергетике в пользу вторых по основным типам угроз, заблокированных на компьютерах, в том числе: интернет-угрозам, почтовым угрозам, угрозам на съемных носителях. В исследованиях атак на ТЭК России [1] за 2018-2019 годы показано, что основными целями киберпреступников, атакующих ТЭК, являются разрушительное воздействие на инфраструктуру и промышленный шпионаж.

**Киберситуационная осведомленность.** Активно вопросами ситуационной осведомленности начали заниматься ещё в 80-х годах прошлого столетия. Термин «ситуационная осведомленность», или Situational Awareness, тесно связано в первую очередь с пионерными работами Мики Эндсли (Mica R. Endsley) [5, 11], в которых ситуационная осведомленность определена как «чувственное восприятие элементов обстановки в (едином) пространственно-временном континууме, осознанное восприятие их значения, а также проецирование их в ближайшее будущее». В своей работе [12] М. Эндсли отмечала, что проблемы нового класса технологий, заключающиеся в том числе в росте сложности систем и генерации в них огромного объема данных, являются одним из факторов роста интереса к ситуационной осведомленности. Развитие новых информационно-коммуникационных технологий добавило измерение «Cyber» к традиционным военным и деловым операциям ситуационной осведомленности, поскольку системы и сети, работающие в киберпространстве, имеют уязвимости, которые представляют значительные риски как для отдельных организаций, так и для национальной безопасности [17]. Такое измерение и называют киберситуационной осведомленностью.

Определение термина «ситуационная осведомленность» в контексте киберсреды приведено в глоссарии Национального института стандартов и технологий США (NIST) [18], в соответствии с которым ситуационная осведомленность – восприятие мер (стратегий, средств) обеспечения безопасности предприятия и его ландшафта угроз в единой пространственно-временной системе координат; понимание двух этих аспектов вместе (риск); прогноз их состояния на ближайшее будущее (*перевод автора*).

На основе определений, приведенных в [9, 13, 20,], в контексте этих исследований под киберситуационной осведомленностью понимается область исследований, связанная с применением методов искусственного интеллекта в области кибербезопасности, направленная на повышение осведомленности о возможных ситуациях нарушений кибербезопасности и автоматическое обнаружение киберугроз.

Под термином «Киберситуационная осведомленность энергетических объектов» будем понимать осведомленность о состоянии киберсреды энергетических объектов, включающую информацию о: критических уязвимостях энергетических объектов с точки зрения кибербезопасности; киберугрозах, инициирующих эти критические уязвимости, а также о техногенных угрозах энергетической безопасности, вызванных киберугрозами.

Эта сравнительно новая концепция для нашей страны уже получила достаточно широкое распространение за рубежом [9, 10, 13, 19, 20]. Отмечается, что киберситуационная осведомленность не может рассматриваться изолированно, она взаимозависима и является частью ситуационной осведомленности, которая ограничена рамками киберсреды. Киберсреда включает подключенные компьютерные устройства, персонал, инфраструктуру, приложения, сервисы, телекоммуникационные системы, а также совокупность передаваемой и/или хранящейся информации [21]. Самостоятельное понимание ситуационной осведомленности, полученное на основе анализа и оценки действий в сети, рассматривается

как дополнительная информация и знания для получения общей ситуационной осведомленности путем их объединения с информацией и знаниями о параметрах, состояниях, характеристиках системы вне киберсреды [13].

Современные решения для автоматизации технологического процесса на энергетических объектах становятся более сложными и используют передовые цифровые технологии [15], что приводит к увеличению рисков нарушения безопасности этих объектов, вплоть до возникновения экстремальных ситуаций. Ведущим направлением решения проблем обеспечения безопасности являются теории, основанные на концепции риска. Технические системы почти всегда проектируются, конструируются и эксплуатируются в неизбежных условиях риска и неопределенности. Концепция риска основана на определении текущих состояний элементов системы и условий возникновения и развития угроз при чрезвычайных, аварийных и катастрофических ситуациях, качественном и количественном описании сценариев и последствий достижения предельных состояний с возникновением аварий и катастроф [22].

**Модель сценариев ЭКС в энергетике, вызванных реализацией киберугроз.** В рамках исследований киберситуационной осведомленности автором предлагается модель сценариев ЭКС в энергетике, вызванных реализацией киберугроз, на основе байесовских сетей доверия (БСД). Модель строится в соответствии с типовой схемой атаки на технологический сегмент сети [8], включающей три основных этапа:

1. Проникновение в корпоративную информационную систему (КИС).
2. Проникновение из КИС в технологический сегмент (ТС) локальной вычислительной сети (ЛВС).
3. Атака на целевой актив.

На основе модели осуществляется вероятностный вывод и определяются вероятности киберугроз на каждом этапе. Вероятностный вывод требует совместного распределения вероятностей, вследствие чего возникает проблема экспоненциальной сложности с ростом количества переменных. Для решения подобной проблемы создаются модели байесовской сети доверия, разбивающие сложное распределение вероятностей на ряд простых узлов, что снижает проблематичность получения знаний и сложность вероятностного вывода [23].

В статье рассматривается уровень КСО как способ измерения ситуационной осведомленности в киберсреде. На основе структурирования знаний с использованием фрактального стратифицированного подхода [14] и онтологического инжиниринга [2] разработана модель сценариев ЭКС в энергетике, вызванных киберугрозами:

$$M = \{A, V, T, W, C, F\}, \quad (1)$$

где  $A$  – множество активов информационно-технологической системы энергетического объекта (аппаратно-программные, программные, протоколы и пр.),  $V$  – множество всех обнаруженных критических уязвимостей активов рассматриваемой ЛВС,  $T$  – множество киберугроз активов ЛВС,  $W$  – множество техногенных угроз ЭБ, вызванных киберугрозами  $T$ ,  $C$  – множество последствий реализации угроз,  $F$  – множество взаимосвязей между критическими активами, уязвимостями активов ЛВС, угрозами и последствиями реализации угроз, представленных в виде продукционных правил «ЕСЛИ-ТО».

Далее рассмотрим каждый компонент модели подробнее. Множество активов информационно-технологической системы энергетического объекта  $A$  включает подмножество активов ЛВС и технологической инфраструктуры:

$$A = \{A^E, A^I\}, \quad (2)$$

$$A^E = \{a_1^E, a_2^E, \dots, a_k^E\}, \quad (3)$$

$$A^I = \{a_1^I, a_2^I, \dots, a_l^I\}, \quad (4)$$

где  $A^E$  – множество активов ЛВС,  $A^I$  – множество активов технологической инфраструктуры (ТехИ) (энергетические объекты),  $a_1^E, a_2^E, \dots, a_k^E$  – активы ЛВС,  $a_1^I, a_2^I, \dots, a_l^I$  – активы ТехИ.

Каждый актив  $a_i^E$  содержит набор критических уязвимостей или, иными словами, каждой критической уязвимости из набора уязвимостей ставится в соответствие конкретный актив ЛВС, так что:

$$F_{a_i^E}^{V_i}: V_i \rightarrow a_i^E, \quad (5)$$

$$V_i = \{v_1, v_2, \dots, v_m\}, V_i \subset V, \quad (6)$$

где  $F_{a_i^E}^{V_i}$  – отображение взаимосвязи между активом ЛВС  $a_i^E$  и набором обнаруженных критических уязвимостей  $V_i$  актива.

Множество киберугроз  $T$  представлено двумя подмножествами:

$$T = \{T^V, T^W\}, \quad (7)$$

где  $T^V$  – множество киберугроз, инициирующих критические уязвимости,  $T^W$  – множество киберугроз, спровоцированных инициирующими киберугрозами  $T^V$ .

Каждая уязвимость  $v_i$  может быть использована одной или более инициирующей киберугрозой, т.е.:

$$T_i^V = \{t_1^V, t_2^V, \dots, t_h^V\}, T_i^V \subset T^V, \quad (8)$$

$$F_{v_i}^{T_i^V}: T_i^V \rightarrow v_i, \quad (9)$$

где  $t_1^V, t_2^V, \dots, t_h^V$  – набор киберугроз, которые могут инициировать уязвимость  $v_i$ ,  $F_{v_i}^{T_i^V}$  – отображение взаимосвязи между уязвимостью  $v_i$  и набором инициирующих её киберугроз.

В свою очередь, инициирующие киберугрозы влекут реализацию других киберугроз т.е. цепочки киберугроз (вектора атак), представленных, как:

$$T_i^W = \{t_1^W, t_2^W, \dots, t_z^W\}, T_i^W \subset T^W, \quad (10)$$

$$F_{T_i^V}^{T_i^W}: T_i^V \rightarrow T_i^W, \quad (11)$$

где  $t_1^W, t_2^W, \dots, t_z^W$  – набор киберугроз, спровоцированных киберугрозой  $t_i^V$ ,  $F_{T_i^V}^{T_i^W}$  – отображение взаимосвязи между инициирующей угрозой и набором спровоцированных киберугроз.

Кибернетические угрозы способны вызывать техногенные угрозы на объекте, которые также являются угрозами энергетической безопасности:

$$W = \{w_1, w_2, \dots, w_u\}, \quad (12)$$

$$F_W^{T_i}: T_i \rightarrow W, \quad (13)$$

где  $W$  – множество техногенных угроз ЭБ, вызванных киберугрозами  $T$ ,  $F_W^{T_i}$  – отображение взаимосвязи между вектором атаки и множеством техногенных угроз ЭБ  $W$ , рассматриваемым в модели как вектор киберугрозы.

Техногенные угрозы ЭБ на объекте могут привести к неблагоприятным последствиям, нарушающим функционирование объекта:

$$C = \{c_1, c_2, \dots, c_q\}, \quad (14)$$

$$F_W^{C_k}: C_k \rightarrow W, \tag{15}$$

где  $C$  – множество последствий реализации угроз,  $F_T^{C_k}$  – отображение взаимосвязи между множеством техногенных угроз ЭБ и последствиями реализации угроз.

Множество взаимосвязей модели описывается следующим образом:

$$F = \left\{ F_{a_i^E}^{V_i}, F_{v_i}^{T_i^V}, F_{T_i^V}^{t_i^W}, F_W^{T_i}, F_W^{C_k} \right\}, \tag{16}$$

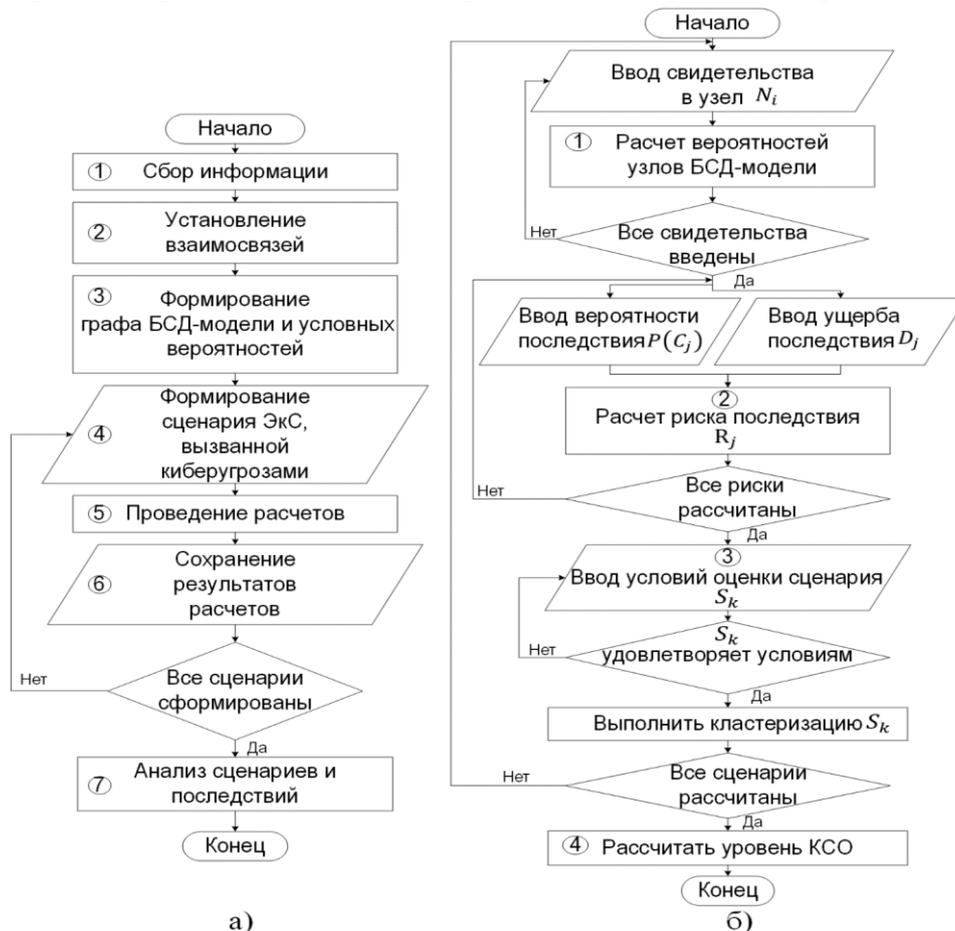
где  $F_{a_i^E}^{V_i}, F_{v_i}^{T_i^V}, F_{T_i^V}^{t_i^W}, F_W^{T_i}, F_W^{C_k}$  определены в формулах (9), (11), (13), (15).

**Численный метод определения уровня КСО энергетического объекта.** Определять уровень КСО предлагается с использованием описанной выше модели, по формуле:

$$L = \frac{Z}{2^{n-1}}, \tag{17}$$

где  $L$  – уровень КСО энергетического объекта,  $Z$  – количество рассчитанных сценариев ЭкС в энергетике, вызванных киберугрозами,  $2^n - 1$  – общее количество таких сценариев, где  $n = |V| + |T| + |W|$ . Значения  $V, T, W$  введены ранее в формулах (6)-(7), (12).

Численный метод реализуется алгоритмом, представленным на рис 1 а. Блоки 4-7 алгоритма на рис. 1 а представлены детализированным алгоритмом на рис. 1 б.



**Рис. 1.** а) алгоритм определения уровня КСО энергетического объекта, б) детализированный алгоритм блоков 4-7 алгоритма определения уровня КСО

**Комментарий к рис 1 а.** Блоки 1-3 соответствуют сбору информации, который заключается в определении множеств активов  $A$  формула (2), критических уязвимостей  $V$  (6), киберугроз  $T$  (7-11),  $W$  – техногенных угроз ЭБ (12),  $C$  – последствий реализации угроз (14), установлении взаимосвязей  $F$  (16).

Блок 4 представляет формирование графа БСД-модели и таблиц условных вероятностей (ТУВ) его узлов (для всех типов узлов в соответствии с  $M$  (1) – моделью сценариев ЭкС в энергетике, вызванной киберугрозами для расчета сценария). Начальные вероятностные характеристики (ТУВ) определяются на основе экспертных или статистических данных.

Для проведения расчетов (блоки 5-6) требуется вычислить вероятность последствий для каждого сценария, кластеризовать сценарии в соответствии с критериями кластеризации и определить уровень КСО. Сценарий описывается как:

$$S_k^C = \{S_1^C, S_2^C, \dots, S_j^C\}, \quad (18)$$

где  $S_k^C$  – множество кластеризованных сценариев ЭкС в энергетике, вызванных киберугрозами (все сценарии разбиваются на три кластера: норма, предкризис, кризис [6]),  $k = \overline{1,3}$ . Принимается, что существует взаимно-однозначное соответствие между сценарием  $S_j^C$  и вероятностью его последствия  $P(c_q^S)$ .

Анализ сценариев и последствий (блок 7) осуществляется путем анализа распределения рисков отдельных сценариев по кластерам, сопровождается описанием рисков  $R$ :

$$R = \{T, V, W, C, D\}, \quad (19)$$

где значения  $T, V, W, C$  введены ранее в (7-15),  $D$  – ущерб последствия сценария  $S^C$ .

**Комментарий к рис. 1 б.** Расчет вероятностей (блок 1) выполняется по формулам:

$$\begin{aligned} P(N_j|E_i = True) &= \sum_{j=1}^n P(E_i = True) \times P(N_j|E_i = True), \\ P(N_j|E_i = False) &= \sum_{j=1}^n P(E_i = False) \times P(N_j|E_i = False), \end{aligned} \quad (20)$$

$$N_j \in N,$$

где  $N$  – множество всех узлов графа БСД-модели,  $E_i$  – узел графа БСД-модели, в который введено свидетельство,  $N_j$  – множество дочерних узлов узла  $E_i$ ,  $i = \overline{1, |N| - |C|}$ ,  $n = \overline{1, |N_j|}$ ,  $C$  введено ранее в (14).

Количественная оценка рисков последствий сценария  $S_j^C$  (блок 2) выражается в вычислении вероятностного ущерба по формуле:

$$R_j = P(C_j) \times D_j, \quad (21)$$

где вероятность  $P$  является функцией  $G$  от сценария  $S^C$ :  $P(C_j) = G(S^C)$ ,  $P(C_j)$  – вероятность последствия сценария  $S_j^C$ , рассчитываемая по формулам (20),  $D_j$  – ущерб от реализации киберугроз этого сценария.

Условия оценки сценариев ЭкС, вызванных киберугрозами (блок 3), имеют вид:

$$0 \leq P(C_j) \times D_{C_j} \leq l_1, \quad (22)$$

$$l_1 \leq P(C_j) \times D_{C_j} \leq l_2, \quad (23)$$

$$l_2 \leq P(C_j) \times D_{C_j}, \quad (24)$$

$$0 \leq P(C_j) \leq 1, \quad (25)$$

$$0 \leq D_{C_j} \leq D_m, \quad (26)$$

где  $l_1, l_2$  – критерии кластеризации сценариев ЭкС, вызванных киберугрозами,  $D_m$  – максимальный ущерб.

Сценарии, удовлетворяющие условиям (22), (25), (26) считаем незначительными (норма), удовлетворяющие условиям (23), (25), (26) считаем соответствующими среднему уровню опасности (предкризис), удовлетворяющие условиям (24), (25), (26) – опасными (кризис).

Определение уровня КСО энергетического объекта (блок 4, рис. 1б) осуществляется по формуле (17).

**Заключение.** В статье приведены определения киберситуационной осведомленности, представлены модель сценариев ЭкС в энергетике, вызванных реализацией киберугроз, предложенная на основе выполненного ранее структурирования знаний с использованием фрактального стратифицированного подхода и онтологического инжиниринга, и численный метод определения уровня КСО энергетического объекта. Реализация предложенных модели и численного метода определения уровня КСО энергетического объекта выполнена в рамках разработки интеллектуального программного комплекса ИПК «ОКО» для анализа киберситуационной осведомленности энергетического объекта.

**Благодарности.** Работа выполнена в рамках выполнения проекта по госзаданию ИСЭМ СО РАН АААА-А17-117030310444-2 (проект №349-2016-0005) и при частичной финансовой поддержке грантов РФФИ №19-57-04003, № 19-07-00351, №18-07-00714, 20-010-00204.

#### СПИСОК ЛИТЕРАТУРЫ

1. АРТ-атаки на топливно-энергетический комплекс России: обзор тактик и техник / Аналитические статьи Positive Technologies. Режим доступа: <https://www.ptsecurity.com/ru-ru/research/analytics/apt-attacks-energy-2019/> (дата обращения: 15.11.2020).
2. Гаськова Д.А., Массель А.Г. Онтологический инжиниринг для разработки интеллектуальной системы анализа угроз и оценки рисков кибербезопасности энергетических объектов // Онтология проектирования. 2019. Т. 9. №2 (32). С. 225-238.
3. Кибератаки на системы АСУ ТП в энергетике в Европе. Первый квартал 2020 года / Отчеты Kaspersky ICS CERT. Режим доступа: <https://ics-cert.kaspersky.ru/reports/2020/09/03/cyberthreats-for-ics-in-energy-in-europe-q1-2020/> (дата обращения: 15.11.2020).
4. Массель Л.В., Воропай Н.И., Сендеров С.М., Массель А.Г. Киберопасность как одна из стратегических угроз энергетической безопасности // Вопросы кибербезопасности. 2016. № 4 (17). С. 2-10.
5. Массель Л.В., Иванов Р.А., Массель А.Г. Моделирование этапов принятия решений на основе сетецентрического подхода // Вестник ИрГТУ. 2013. №10(81). С.16-22.
6. Массель Л.В., Массель А.Г. Технологии и инструментальные средства интеллектуальной поддержки принятия решений в экстремальных ситуациях в энергетике // Вычислительные технологии. 2013. Т. 18. № S1. С. 37-44.
7. Об утверждении Доктрины энергетической безопасности Российской Федерации / Указ Президента Российской Федерации от 13 мая 2019 года № 216. Режим доступа: <https://minenergo.gov.ru/system/download-pdf/14766/9694114766> (дата обращения 15.11.2020).
8. Промышленные компании: векторы атак/ Аналитические статьи Positive Technologies. Режим доступа: <https://www.ptsecurity.com/ru-ru/research/analytics/ics-attacks-2018/> (дата обращения 15.11.2020).
9. Cheng Y., Deng J., Li J., DeLoach S.A., Singhal A., Ou X. Metrics of Security. In: Kott A., Wang C., Erbacher R. (eds) Cyber Defense and Situational Awareness. Advances in Information Security. 2014. Vol 62. Springer. Cham.

10. Eckhart M., Ekelhart A., Weippl E. Enhancing Cyber Situational Awareness for Cyber-Physical Systems through Digital Twins // Proceedings of the 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). 2019. Pp. 1222-1225. DOI:10.1109/etfa.2019.8869197.
11. Endsley M.R. Situation awareness global assessment technique (SAGAT) // Proceedings of the IEEE 1988 National Aerospace and Electronics Conference. 1988. Pp. 789-795. DOI:10.1109/naecon.1988.195097.
12. Endsley M.R. Theoretical underpinnings of situation awareness: A critical review. In: Endsley M.R., Garland D.J. Situation awareness analysis and measurement. 2000. Pp. 3-32.
13. Frank U., Brynielsson J. Cyber Situational Awareness – A systematic review of literature // Computer Security. 2014. Vol. 46. Pp. 18–31. DOI: 10.1016/j.cose.2014.06.008.
14. Gaskova D. Fractal Stratified Model Development for Critical Infrastructure from the standpoint of Energy and Cyber Security // Proceedings of the VIth International Workshop “Critical Infrastructures: Contingency Management, Intelligent, Agent-Based, Cloud Computing and Cyber Security (IWCI 2019)”. Publisher: Atlantis Press. 2019. Irkutsk: MESI SB RAS. Pp. 179-183. DOI:10.2991/iwci-19.2019.31.
15. Irmak E., Erkek I. An overview of cyber-attack vectors on SCADA systems // Proceedings of 6th International Symposium on Digital Forensic and Security (ISDFS). 2018. DOI:10.1109/isdfs.2018.8355379.
16. ISO/IEC 27032:2012 ISO standard of Information technology. Security techniques. Guidelines for cybersecurity. Режим доступа: <https://www.iso.org/ru/standard/44375.html> (дата обращения 15.11.2020).
17. MITRE Capabilities overviews. Режим доступа: <https://www.mitre.org/capabilities/cybersecurity/situation-awareness> (дата обращения 15.11.2020).
18. National Institute of Standards and Technology (NIST) Glossary. Режим доступа: <https://csrc.nist.gov/glossary> (дата обращения 15.11.2020).
19. Okolica J.S., McDonald J.T., Peterson G.L., Mills R.F., Haas M.W. Developing Systems for Cyber Situational Awareness // Proceedings of the 2nd Cyberspace Research Workshop. Shreveport. Louisiana. 2009.
20. Pappaterra M. J., Flammini F. A Review of Intelligent Cybersecurity with Bayesian Networks // Proceedings of IEEE International Conference on Systems, Man and Cybernetics (SMC). Bari. Italy. 2019. Pp. 445-452. DOI:10.1109/smc.2019.8913864.
21. T-REC-X.1205 – ITU-T: Overview of cybersecurity. Режим доступа: <https://www.itu.int/rec/T-REC-X.1205-200804-I> (дата обращения 15.11.2020).
22. Yacov Y. Haimes. Systems-based risk analysis. In: Bostrom N., Cirkovic M.M. Global Catastrpphic Risks. Oxford. 2008. Pp. 146-163.
23. Zhang Q., Zhou C., Xiong N., Qin Y., Li X., Huang S. Multimodel-Based Incident Prediction and Risk Assessment in Dynamic Cybersecurity Protection for Industrial Control Systems // Proceedings of IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2016. Vol. 46. № 10. Pp. 1429-1444. DOI: 10.1109/TSMC.2015.2503399.

## METHOD FOR DETERMINING THE LEVEL OF CYBER SITUATIONAL AWARENES ON ENERGY FACILITIES

**Daria A. Gaskova**

Junior Researcher, e-mail: [gaskovada@gmail.com](mailto:gaskovada@gmail.com),  
Melentiev Energy Systems Institute of  
Siberian Branch of the Russian Academy of Sciences,  
130, Lermontov Str., 664033, Irkutsk, Russia.

**Annotation.** The article describes the research direction of situational awareness in the cyber environment on energy facilities. A model based on a Bayesian Belief Network and a numerical method for determining cyber situational awareness are proposed to increase awareness of the cyber environment state on such facilities. The model is based on causal relationships between the vulnerabilities of the local area network and possible cyber threats, presented in the form of vectors of penetration into the network, vectors of the development of cyber attacks and attack vectors on the target asset, combined in scenarios. The work focuses on scenarios, the consequences of which can be regarded as extreme situations in the energy sector caused by cyber threats.

**Keywords:** Bayesian Belief Network, emergency situation in the energy sector, cyber threats.

### References

1. APT-ataki na toplivno-energeticheskij kompleks Rossii: obzor taktik i tekhnik [APT attacks to Russian energy sector: review of tactics and techniques] / Analytics of Positive Technologies. Available at: <https://www.ptsecurity.com/ru-ru/research/analytics/apt-attacks-energy-2019> accessed 15.11.2020 (in Russian).
2. Gaskova D.A., Massel A.G. Ontologicheskij inzhiniring dlya razrabotki intellektual'noj sistemy analiza ugroz i ocenki riskov kiberbezopasnosti energeticheskikh obektov [Ontological engineering for the development of the intelligent system for threats analysis and risk assessment of cybersecurity in energy facilities] // Ontologiya proektirovaniya = Ontology of Designing. 2019. Vol. 9. Issue 2(32). Pp. 225-238 (in Russian).
3. Kiberataki na sistemy ASU TP v energetike v Evrope [Cyberattacks on ICS in energy in in Europe] / Reports of Kaspersky ICS CERT. Available at: <https://ics-cert.kaspersky.ru/reports/2020/09/03/cyberthreats-for-ics-in-energy-in-europe-q1-2020> accessed 15.11.2020 (in Russian).
4. Massel L.V., Voropay N.I., Senderov S.M, Massel A.G. Kiberopasnost' kak odna iz strategicheskikh ugroz energeticheskoy bezopasnosti [Cyber danger as one of the strategic threats to russia's energy security] // Voprosy kiberbezopasnosti = Cybersecurity issues. 2016. №. 4(17). Pp. 2-10 (in Russian).
5. Massel L.V., Ivanov R.A., Massel A.G. Modelirovanie etapov prinyatiya reshenij na osnove setecentricheskogo podhoda [Decision-making stages modeling based on network-centric approach] // Vestnik IrGTU = Proceedings of ISTU. 2013. № 10(81). Pp. 16-22 (in Russian).

6. Massel L.V., Massel A.G. Tekhnologii i instrumental'nye sredstva intellektual'noj podderzhki prinyatiya reshenij v ekstremal'nyh situacijah v energetike [Technologies and tools for intelligent decision-making support in extreme situations in the energy sector] // Vychislitel'nye tekhnologii = Computational Technologies. 2013. Vol. 18. № S1. Pp. 37-44 (in Russian).
7. Ob utverzhdenii Doktriny energeticheskoy bezopasnosti Rossijskoj Federacii [On approval of the Energy Security Doctrine of the Russian Federation] / Ukaz Prezidenta Rossijskoj Federacii ot 13 maya 2019 goda no. 216 = Decree of the President of the Russian Federation of May 13. 2019. № 216. Available at: <https://minenergo.gov.ru/system/download-pdf/14766/9694114766>, accessed 15.11.2020 (in Russian).
8. Promyshlennye kompanii: vektory atak [Industrial companies: attack vectors] / Analytics of Positive Technologies. Available at: <https://www.ptsecurity.com/ru-ru/research/analytics/ics-attacks-2018> accessed 15.11.2020 (in Russian).
9. Cheng Y., Deng J., Li J., DeLoach S.A., Singhal A., Ou X. Metrics of Security. In: Kott A., Wang C., Erbacher R. (eds) Cyber Defense and Situational Awareness. Advances in Information Security. 2014. Vol 62. Springer, Cham.
10. Eckhart M., Ekelhart A., Weippl E. Enhancing Cyber Situational Awareness for Cyber-Physical Systems through Digital Twins // Proceedings of the 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). 2019. Pp. 1222-1225. DOI:10.1109/etfa.2019.8869197.
11. Endsley M.R. Situation awareness global assessment technique (SAGAT) // Proceedings of the IEEE 1988 National Aerospace and Electronics Conference. 1988. Pp. 789-795. DOI:10.1109/naecon.1988.195097.
12. Endsley M.R. Theoretical underpinnings of situation awareness: A critical review. In: Endsley M.R., Garland D.J. Situation awareness analysis and measurement. 2000. Pp. 3-32.
13. Frank U., Brynielsson J. Cyber Situational Awareness – A systematic review of literature // Computer Security. 2014. Vol. 46. Pp. 18–31. DOI: 10.1016/j.cose.2014.06.008.
14. Gaskova D. Fractal Stratified Model Development for Critical Infrastructure from the standpoint of Energy and Cyber Security // Proceedings of the VIth International Workshop “Critical Infrastructures: Contingency Management, Intelligent, Agent-Based, Cloud Computing and Cyber Security (IWCI 2019)”. Publisher: Atlantis Press. 2019. Irkutsk: MESI SB RAS. Pp. 179-183. DOI: 10.2991/iwci-19.2019.31.
15. Irmak E., Erkek I. An overview of cyber-attack vectors on SCADA systems // Proceedings of 6th International Symposium on Digital Forensic and Security (ISDFS). 2018. DOI: 10.1109/isdfs.2018.8355379.
16. ISO/IEC 27032:2012 ISO standard of Information technology. Security techniques. Guidelines for cybersecurity. Режим доступа: <https://www.iso.org/ru/standard/44375.html> (дата обращения 15.11.2020).
17. MITRE Capabilities overviews. Режим доступа: <https://www.mitre.org/capabilities/cybersecurity/situation-awareness> (дата обращения 15.11.2020).
18. National Institute of Standards and Technology (NIST) Glossary. Режим доступа: <https://csrc.nist.gov/glossary> (дата обращения 15.11.2020).

19. Okolica J.S., McDonald J.T., Peterson G.L., Mills R.F., Haas M.W. Developing Systems for Cyber Situational Awareness // Proceedings of the 2nd Cyberspace Research Workshop. Shreveport. Louisiana. 2009.
20. Pappaterra M. J., Flammini F. A Review of Intelligent Cybersecurity with Bayesian Networks // Proceedings of IEEE International Conference on Systems, Man and Cybernetics (SMC). Bari. Italy. 2019. Pp. 445-452. DOI:10.1109/smc.2019.8913864.
21. T-REC-X.1205 – ITU-T: Overview of cybersecurity. Режим доступа: <https://www.itu.int/rec/T-REC-X.1205-200804-I> (дата обращения 15.11.2020).
22. Yacov Y. Haimes. Systems-based risk analysis. In: Bostrom N., Cirkovic M.M. Global Catastrpphic Risks. Oxford. 2008. Pp. 146-163.
23. Zhang Q., Zhou C., Xiong N., Qin Y., Li X., Huang S. Multimodel-Based Incident Prediction and Risk Assessment in Dynamic Cybersecurity Protection for Industrial Control Systems // Proceedings of IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2016. Vol. 46. №. 10. Pp. 1429-1444. DOI: 10.1109/TSMC.2015.2503399.

УДК 004.8 : 004.056.5

**МОДИФИКАЦИИ АЛГОРИТМА ПРЫГАЮЩИХ ЛЯГУШЕК ДЛЯ ОТБОРА ПРИЗНАКОВ В НЕЧЕТКОМ КЛАССИФИКАТОРЕ ПРИ АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЯ ПО РУКОПИСНОЙ ПОДПИСИ**

**Ходашинский Илья Александрович**

Д.т.н., профессор,

e-mail: [hodashn@rambler.ru](mailto:hodashn@rambler.ru),

**Бардамова Марина Борисовна**

Аспирантка,

e-mail: [722bmb@gmail.com](mailto:722bmb@gmail.com),

Томский государственный университет систем управления и радиоэлектроники  
(ТУСУР), 634050 г. Томск, пр. Ленина 40.

**Аннотация.** В работе проведено сравнение эффективности модификаций алгоритма прыгающих лягушек, позволяющих метаэвристике функционировать в бинарном пространстве поиска. Для задачи отбора признаков в нечетком классификаторе опробованы методы, основанные на модифицированных алгебраических операциях, функциях трансформации и операции слияния, а также их комбинации. В эксперименте использован набор данных SVC2004, содержащий большое количество признаков для аутентификации пользователя на основе динамических признаков рукописной подписи.

**Ключевые слова:** нечеткий классификатор, алгоритм прыгающих лягушек, метаэвристики, отбор признаков, бинаризация

**Цитирование:** Ходашинский И.А., Бардамова М. Б. Модификации алгоритма прыгающих лягушек для отбора признаков в нечетком классификаторе при аутентификации пользователя по рукописной подписи // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С.75-83. DOI: 10.38028/ESI.2020.20.4.007

**Введение.** Метаэвристические алгоритмы широко применяются для оптимизации в различных задачах, в том числе с целью отбора информативных подмножеств признаков при построении модели машинного обучения. Удаление лишних признаков помогает избежать переобучение модели и уменьшить её сложность. При использовании метаэвристик в качестве инструмента отбора признаков необходимым условием является способность алгоритма осуществлять поиск в бинарном пространстве. Однако в то время, как некоторые метаэвристики, например, генетический алгоритм, изначально были созданы для работы с бинарными входными векторами, другие были разработаны для работы в непрерывной области поиска. Примерами последних могут служить как алгоритм роящихся частиц, так и множество других метаэвристик. Очевидно, что один только генетический алгоритм не способен быть универсальным инструментом для любых данных, поскольку это противоречит теореме о бесплатных завтраках [13]. Поэтому важно иметь аналогичные альтернативы и разрабатывать модификации, способные трансформировать изначально непрерывный алгоритм в бинарный. Самыми распространенными методами бинаризации являются использование пороговых или трансформационных функций [2, 10, 12].

Алгоритм прыгающих лягушек (АПЛ) успешно применяется для оптимизации параметров классификатора [1, 3]. Существуют и примеры использования этой метаэвристики в бинарном пространстве [6, 12, 15]. В [12] АПЛ совмещается с алгоритмом роящихся частиц, откуда заимствует объемную процедуру вычисления скорости элементов входных векторов, и использует процедуры мутации и кроссовера из генетического алгоритма.

Авторы предлагают вычислительно простые способы бинаризации АПЛ, позволяющие обеспечить высокую скорость работы алгоритма. Все описанные методы будут применены к этой метаэвристике впервые. Проверка эффективности бинарного алгоритма при решении задачи отбора признаков в нечетком классификаторе будет проведена на наборе данных SVC2004. Этот набор содержит сигналы, описывающие динамические характеристики рукописной подписи с целью определения легитимности пользователя при прохождении процедуры аутентификации.

**1. Непрерывный алгоритм прыгающих лягушек.** Идея метаэвристики, разработанной в 2003 году, заключается в воспроизведении поведения группы лягушек в процессе поиска пищи [5]. В данной работе в качестве основы взята версия, описанная в [4]. На вход подаются вектора вида  $\mathbf{Z} = (z_1, z_2, \dots, z_{|Z|})$ ,  $z \in \mathbb{R}$ . В алгоритме комбинируются итерационно повторяющиеся внешний глобальный и внутренний локальный поиск. Глобальный поиск заключается в сортировке входных векторов по значению фитнес-функции и разбиением популяции на подгруппы. Внутри каждой подгруппы независимо осуществляется локальный поиск, в котором происходит обновление худших векторов.

Обновление векторов происходит согласно следующему принципу. Выбираются  $\mathbf{Z}_b$  и  $\mathbf{Z}_w$  – вектора с лучшим и худшим значением фитнес-функции в группе. Далее вычисляется промежуточный вектор  $\mathbf{Z}^*$ :

$$\mathbf{Z}^* = r \times c \times (\mathbf{Z}_b - \mathbf{Z}_w) + \mathbf{Z}_w, \quad (1)$$

где  $r$  – равномерно распределенное случайное число из промежутка от 0 до 1,  $c$  – коэффициент обновления вектора. Если значение фитнес-функции вектора  $\mathbf{Z}^*$  превосходит фитнес-функцию вектора  $\mathbf{Z}_w$ ,  $\mathbf{Z}^*$  замещает собой  $\mathbf{Z}_w$ . В противном случае вектор  $\mathbf{Z}^*$  вычисляется заново, но вместо  $\mathbf{Z}_b$  используется глобально лучший вектор  $\mathbf{Z}_0$ . Если и в данном случае не удастся улучшить вектор  $\mathbf{Z}_w$ , то он перезаписывается со случайно сгенерированным отклонением.

Алгоритм принимает на вход следующие параметры: число подгрупп  $G$ , число векторов в данной группе  $F$ , коэффициент для обновления входных векторов  $c$ , количество итераций для глобального и локального поиска  $T_{gl}$  и  $T_{lc}$  соответственно.

**2. Бинарный алгоритм прыгающих лягушек.** На вход бинарной версии метаэвристики подается популяция, состоящая из векторов вида  $\mathbf{S} = (s_1, s_2, \dots, s_n)$ . Каждый элемент вектора соответствует признаку из таблицы наблюдения,  $n$  – количество признаков. Значение элемента вектора  $s_i = 0$  показывает, что  $i$ -ый признак не участвует в классификации,  $s_i = 1$  означает, что  $i$ -ый признак используется классификатором. Алгоритм должен найти такой вектор  $\mathbf{S}$ , который позволил бы максимально улучшить фитнес-функцию при уменьшении количества признаков.

В фитнес-функции важно учитывать как качество построения модели классификации, так и долю отобранных признаков с целью уменьшения их количества:

$$Fit(\mathbf{S}) = \alpha \times Error(\mathbf{S}) + (1 - \alpha) \times \frac{n^*}{n}, \quad (2)$$

где  $Error(\mathbf{S})$  – ошибка классификации на векторе  $\mathbf{S}$ ,  $n^*$  – количество элементов в векторе  $\mathbf{S}$ , равных единице,  $\alpha$  – коэффициент приоритета одной части функции над другой.

**2.1. Метод модифицированных алгебраических операций** позволяет алгоритму оперировать бинарными входными векторами, так как все арифметические операторы заменяются на логические [8, 9]. В соответствии с этой идеей операции умножения, сложения и вычитания в (1) заменяются на конъюнкцию, дизъюнкцию и строую дизъюнкцию соответственно:

$$\mathbf{S}^* = r \wedge (\mathbf{S}_b \oplus \mathbf{S}_w) \vee \mathbf{S}_w, \quad (3)$$

где  $\mathbf{r}$  – случайный бинарный вектор.

**2.2. Метод слияния**, предполагающий использование оригинальной функции слияния для осуществления бинарного поиска, впервые был предложен в [7]. Функция поэлементно сравнивает два бинарных вектора; если значение элемента на одной и той же позиции совпадает, то в результирующем векторе на эту позицию запишется данное значение. В ином случае осуществляется генерация случайного числа из интервала от 0 до 1. Если оно меньше или равно 0,5, то в соответствующую позицию нового вектора записывается элемент из худшего вектора. В противном случае на этом месте будет выставлен элемент из лучшего вектора.

Таким образом, функция слияния может быть задана следующим образом:

$$\text{merge}(\mathbf{S}_w, \mathbf{S}_b) = \begin{cases} s_i^* = s_{wi} = s_{bi}, & \text{если } s_{wi} = s_{bi} \\ s_i^* = s_{wi}, & \text{если } s_{wi} \neq s_{bi} \text{ и } rand \leq 0,5, \\ s_i^* = s_{bi}, & \text{если } s_{wi} \neq s_{bi} \text{ и } rand > 0,5 \end{cases} \quad (4)$$

где  $rand$  – случайное равномерно распределенное число,  $rand \in [0;1]$ .

В алгоритме прыгающих лягушек описанную функцию удобно применять в локальном поиске вместо оператора (1). При первой генерации вектора  $\mathbf{S}^*$  в функцию передаются худший и локально лучший вектор, если  $\mathbf{S}^*$  не улучшает  $\mathbf{S}_w$ , производится повторная генерация  $\mathbf{S}^*$ . На повторной генерации вместо локального лидера используется глобальный.

### 2.3. Комбинация методов модифицированных арифметических операций и слияния

Данная модификация представляет собой совмещение оператора, использующего логические операции, и функции слияния двух векторов. Создание нового вектора на первой стадии локального поиска выглядит следующим образом:

$$\mathbf{S}^* = \text{merge}(r \wedge (\mathbf{S}_b \oplus \mathbf{S}_w), \mathbf{S}_b). \quad (5)$$

Как и во всех предыдущих случаях, на второй стадии локального поиска вместо  $\mathbf{S}_b$  применяется глобально лучший вектор  $\mathbf{S}_0$ .

**2.4. Функции трансформации.** Для использования функций трансформации необходимо наличие некоторого непрерывного вектора, каким-либо образом характеризующего признаки [11]. Чаще всего таким вектором служит вектор скорости. В оригинальном АПЛ такого вектора нет, но авторы предлагают рассчитывать в локальном поиске скорость худшего вектора  $\mathbf{S}_w$  следующим образом:

$$\mathbf{V} = (\mathbf{S}_0 - \mathbf{S}_w) \times r_1 + (\mathbf{S}_b - \mathbf{S}_w) \times r_2, \quad (6)$$

где  $r_1$  и  $r_2$  – вектора, заполненные случайными вещественными значениями в диапазоне от 0 до 1. Далее полученному значению необходимо поставить в соответствие бинарный эквивалент. Отображение происходит в два шага. На первом шаге функция трансформации, принимая на вход значение скорости элемента вектора  $\mathbf{S}_w$ , вычисляет некоторое число,

принадлежащее диапазону  $[0;1]$ . Второй шаг заключается в непосредственном обновлении элементов в соответствии с трансформационным правилом.

В этом случае локальный поиск сокращается до двух этапов. На первом происходит создание промежуточного вектора  $\mathbf{S}^*$  на базе  $\mathbf{S}_w$  с помощью функции трансформации и обновление  $\mathbf{S}_w$  в том случае, если  $\mathbf{S}^*$  обладает лучшей фитнес-функцией. В противном случае осуществляется второй этап, на котором на место  $\mathbf{S}_w$  записывается случайный вектор.

Далее описаны исследуемые авторами вариации функций и правил.

Выделяют несколько семейств функций трансформации. Два наиболее часто используемых семейства имеют S- и V-образные графики [10]. В качестве S-образной функции трансформации была использована базовая версия «сигмоиды»:

$$F_1(v_i) = 1/(1 + e^{-v_i}), \quad (7)$$

где  $v_i$  – значение скорости  $i$ -ого элемента.

V-образные функции обладают большим разнообразием вариаций, поэтому для исследования были выбраны две функции. Первая вычисляется с помощью гиперболического тангенса:

$$F_2(v_i) = |\tanh(v_i)|, \quad (8)$$

вторая задается следующим выражением:

$$F_3(v_i) = \left| v_i / \sqrt{1 + v_i^2} \right|. \quad (9)$$

Трансформационные правила отличаются принципом обновления элементов. В первом правиле  $R_1$  элементу строго присваивается бинарное значение:

$$\text{Если } rand < F(v_i), \text{ то } s_i = 1, \text{ иначе } s_i = 0, \quad (10)$$

где  $F(v_j)$  – одна из трех функций трансформации,  $rand$  – равномерно распределенное случайное число,  $rand \in [0;1]$ . Второе правило  $R_2$  либо заменяет элемент на противоположный, либо не меняет его:

$$\text{Если } rand < F(v_i), \text{ то } s_i = s_i \oplus 1. \quad (11)$$

**2.5. Комбинация функций трансформации и операции слияния.** В этом случае первый этап локального поиска заключается в применении функции трансформации на векторе  $\mathbf{S}_w$  и попытке замены худшего вектора на промежуточный  $\mathbf{S}^*$ . Если замена не происходит, то на втором этапе используется операция merge, осуществляющая слияние  $\mathbf{S}^*$  и вектора  $\mathbf{S}_0$ .

**3. Описание эксперимента.** Набор данных SVC2004 (Task 1) посвящен задаче аутентификации пользователя по динамическим характеристикам его рукописной подписи, таким, как скорость изменения координат, ускорение изменения координат, давление на поверхность планшета, длительность подписи и так далее [14]. Всего в SVC2004 содержатся 40 наборов различных подписей, соответствующих разным пользователям. Каждый набор включает в себя 20 вариаций подлинной подписи от её автора и 20 высококачественных подделок, как минимум от четырех злоумышленников. Каждая подпись – это набор сигналов, снятых с графического планшета. После проведения процедур предобработки (устранение разрывов, исправление различий в масштабе, позиции и наклоне подписи) и выделения признаков из сигналов, итоговая таблица наблюдений для каждой подписи состоит из 100

признаков и бинарной метки класса, показывающей принадлежность подписи легитимному пользователю или злоумышленнику.

Поскольку каждая подпись содержит всего 40 экземпляров данных, в эксперименте использована двукратная кросс-валидация. На каждой выборке алгоритм запускался 30 раз. Полученные показатели качества классификации были усреднены по количеству запусков и двум выборкам каждой подписи. Снимались следующие показатели: общая точность, ошибка I рода (False Rejection Rate, FRR), ошибка II рода (False Acceptance Rate, FAR) и количество признаков, оставшихся в наборе.

Структура классификатора генерировалась алгоритмом на основе экстремальных значений классов [7]. В качестве термов были выбраны треугольные функции принадлежности. Параметры алгоритма прыгающих лягушек следующие: количество групп равняется 4, в каждой из них по 10 векторов; локальных итераций 5, глобальных итераций 50. Коэффициент  $\alpha$  в фитнес-функции равнялся 0,5. В таблице 1 представлен перечень использованных в АПЛ способов бинаризации и их обозначения.

**Таблица 1.** Нумерация модификаций алгоритма прыгающих лягушек

Метод		Метод	
MAO	Модифицированные алгебраические операции (3)	F3R2	Функция трансформации $F_3$ (9) + правило $R_2$ (11)
merge	Операция слияния (4)	F1R1m	Функция трансформации $F_1$ (7) + правило $R_1$ (10) + операция слияния
MAOm	Модифицированные алгебраические операции + операция слияния (5)	F2R1m	Функция трансформации $F_2$ (8) + правило $R_1$ (10) + операция слияния
F1R1	Функция трансформации $F_1$ (7) + правило $R_1$ (10)	F3R1m	Функция трансформации $F_3$ (9) + правило $R_1$ (10) + операция слияния
F2R1	Функция трансформации $F_2$ (8) + правило $R_1$ (10)	F1R2m	Функция трансформации $F_1$ (7) + правило $R_2$ (11) + операция слияния
F3R1	Функция трансформации $F_3$ (9) + правило $R_1$ (10)	F2R2m	Функция трансформации $F_2$ (8) + правило $R_2$ (11) + операция слияния
F1R2	Функция трансформации $F_1$ (7) + правило $R_2$ (11)	F3R2m	Функция трансформации $F_3$ (9) + правило $R_2$ (11) + операция слияния
F2R2	Функция трансформации $F_2$ (8) + правило $R_2$ (11)		

**4. Результаты эксперимента.** В таблице 2 содержатся полученные показатели классификации на тестовых выборках, усредненные по всем 40 пользователям. Жирным шрифтом отмечены лучшие результаты по показателю.

**Таблица 2.** Результаты построения нечеткого классификатора на подмножествах признаков, отобранных бинарным алгоритмом прыгающих лягушек

Метод	Точность	FRR	FAR	Признаки
MAO	83,9 ± 7,5	25,1 ± 14,0	7,1 ± 8,1	43,9 ± 4,3
merge	86,7 ± 6,9	20,2 ± 13,0	6,3 ± 7,6	34,2 ± 3,6
MAOm	85,5 ± 7,7	21,1 ± 14,3	7,9 ± 9,2	20,5 ± 4,1
F1R1	85,2 ± 7,4	23,0 ± 13,3	6,5 ± 8,0	40,4 ± 4,1
F2R1	85,6 ± 7,3	21,9 ± 13,5	7,0 ± 8,2	28,0 ± 4,2
F3R1	86,1 ± 7,2	20,7 ± 13,7	7,0 ± 8,3	24,6 ± 4,3
F1R2	83,8 ± 7,6	25,4 ± 14,2	7,1 ± 8,2	43,5 ± 4,1
F2R2	87,0 ± 7,2	19,3 ± 13,3	6,8 ± 8,1	30,0 ± 4,0

F3R2	<b>87,1</b> ± 7,0	19,4 ± 13,1	6,4 ± 7,7	31,1 ± 4,0
F1R1m	85,3 ± 7,3	23,2 ± 13,5	<b>6,2</b> ± 7,8	39,1 ± 3,9
F2R1m	85,9 ± 7,5	20,1 ± 13,9	8,0 ± 9,3	19,7 ± 5,1
F3R1m	85,6 ± 7,7	20,1 ± 14,1	8,6 ± 9,6	<b>17,6</b> ± 4,6
F1R2m	84,2 ± 7,4	24,2 ± 13,6	7,5 ± 8,6	44,8 ± 4,1
F2R2m	<b>87,1</b> ± 7,0	<b>19,0</b> ± 12,9	6,9 ± 8,1	29,4 ± 4,0
F3R2m	86,7 ± 6,9	20,0 ± 13,0	6,7 ± 7,9	30,1 ± 3,9

Средние ранги, полученные при сравнении результатов по всем пользователям критерием Фридмана, представлены в таблице 3.

**Таблица 3.** Средние ранги по критерию Фридмана

Метод	Точность	FRR	FAR	Признаки
MAO	4,8	11,3	7,8	13,9
merge	10,9	6,2	6,1	10,0
MAOm	5,9	8,9	10,5	2,8
F1R1	8,1	9,2	6,4	11,9
F2R1	7,1	9,0	8,4	5,3
F3R1	8,0	8,0	8,1	4,0
F1R2	4,7	12,1	7,5	13,5
F2R2	11,0	<b>4,6</b>	7,3	7,4
F3R2	<b>11,6</b>	5,1	5,7	8,9
F1R1m	8,4	9,7	<b>5,3</b>	11,1
F2R1m	6,9	7,4	11,0	2,2
F3R1m	6,1	7,6	12,0	<b>1,0</b>
F1R2m	5,5	10,2	8,9	14,6
F2R2m	11,1	<b>4,6</b>	8,0	6,0
F3R2m	10,1	6,2	7,1	7,5

Лучшие результаты по общей точности продемонстрировали V-образные функции трансформации с правилом R<sub>2</sub> (11). При этом функция F<sub>2</sub> (8) показала лучший результат в комбинации с операцией слияния, а функция F<sub>3</sub> (9) – индивидуально. Ошибка первого рода оказалась наименьшей при использовании F<sub>2</sub> с правилом R<sub>2</sub> как при отсутствии, так и при наличии операции слияния. Наименьшая ошибка второго рода была получена с помощью F<sub>1</sub> (7) и правила R<sub>1</sub> (10) в комбинации с операцией слияния. Наилучшую способность к сокращению признаков продемонстрировала функция F<sub>3</sub> с правилом R<sub>1</sub> и операцией слияния.

**Заключение.** В работе исследована эффективность пятнадцати модификаций алгоритма прыгающих лягушек для осуществления поиска в бинарном пространстве на примере задачи отбора признаков. Во время эксперимента на наборе данных SVC2004 худшие результаты по совокупности рангов были получены комбинацией «сигмоидальной» фитнес-функции, первого трансформационного правила и операции слияния; этой же функцией, но со вторым правилом, а также методом модифицированных алгебраических операций. Лучшие совокупные результаты продемонстрировал алгоритм прыгающих лягушек при применении V-образных функций и второго трансформационного правила.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-37-90064.

#### СПИСОК ЛИТЕРАТУРЫ

1. Бардамова М.Б., Анфилофьев А.Е., Ковалев В.С., Филимоненко И.В. Сравнительный анализ эффективности метаэвристических алгоритмов при построении нечетких классификаторов // Международная школа-семинар «Интеллектуальные системы и технологии: современное состояние и перспективы»: труды. СПб.: Политехника-сервис. 2017. С. 22-31.
2. Al-Tashi Q., Abdulkadir S.J., Rais H.M., Mirjalili S., Alhussian H., Ragab M.G., Alqushaibi A. Binary Multi-Objective Grey Wolf Optimizer for Feature Selection in Classification // IEEE Access. 2020. vol. 8. Pp. 106247-106263. DOI: 10.1109/ACCESS.2020.3000040
3. Bardamova M. B. Construction of fuzzy classifiers with shuffle frog leaping algorithm for handwritten signature authentication // Сборник избранных статей научной сессии ТУСУР. Т. 1. № 1-2. Томск: В-Спектр. 2019. С. 229-232.
4. Elbeltagi E., Hegazy T., Grierson D. A modified shuffled frog-leaping optimization algorithm: applications to project management // Structure and Infrastructure Engineering. 2007. vol. 3. № 1. Pp. 53–60. DOI: 10.1080/15732470500254535
5. Eusuff M. M., Lansey K.E. Optimizing of water distribution network design using the shuffled frog leaping algorithm // Journal of Water Resources Planning and Management. 2003. vol. 129. № 3. Pp. 210-225.
6. Farsangi M., Barati M. Solving unit commitment problem by a binary shuffled frog leaping algorithm // IET Generation Transmission & Distribution. 2014. vol. 8. № 6. Pp. 1050-1060. DOI: 10.1049/iet-gtd.2013.0436
7. Hodashinsky I., Sarin K., Shelupanov A., Slezkin A. Feature selection based on swallow swarm optimization for fuzzy classification // Symmetry. 2019. vol. 11. № 11. Pp. 1423. DOI: 10.3390/sym11111423
8. Jia D., Duan X., Khan M. K. Binary artificial bee colony optimization using bitwise operation // Computers and Industrial Engineering. 2014. vol. 76. Pp. 360–365. DOI: 10.1016/j.cie.2014.08.016
9. Kiran M. S. Gunduz M. XOR-Based Artificial Bee Colony Algorithm for Binary Optimization // Turkish Journal of Electrical Engineering and Computer Sciences. 2013. vol. 21. 2307–2328. DOI: 10.3906/elk-1203-104
10. Mirjalili S., Lewis A. S-shaped versus V-shaped transfer functions for binary Particle Swarm Optimization // Swarm and Evolutionary Computation. 2013. vol. 9. Pp. 1–14. DOI: 10.1016/j.swevo.2012.09.002
11. Mirjalili S., Mohd Hashim S.Z. BMOA: binary magnetic optimization algorithm // International Journal of Machine Learning and Computing. 2012. vol. 2. № 3. Pp. 204–208. DOI: 10.7763/IJMLC.2012.V2.114
12. Vakil Baghmisheh M.T., Madani K., Navarbaф A. A discrete shuffled frog optimization algorithm // Artificial Intelligence Review. 2011. vol. 36. Pp. 267–284. DOI: 10.1007/s10462-011-9213-2
13. Wolpert D.H., Macready W.G. No Free Lunch Theorems for Optimization // IEEE Transactions on Evolutionary Computation. 1997. Vol.1. Pp. 67-82.

14. Yeung D.-Y., Chang H., Xiong Y., George S., Kashi R., Matsumoto T., Rigoll G.. SVC2004: First International Signature Verification Competition // Conference: Biometric Authentication, First International Conference. Proceedings. Lecture Notes in Computer Science. 2004. Pp. 16-22. DOI: 10.1007/978-3-540-25948-0\_3
  15. Zhang H., Liu Y.-A. Binary Shuffled Frog Leaping Algorithm in Population Diversity for Opportunistic Spectrum Assignment // Journal of Beijing University of Posts Ans Telecom. 2013. vol. 36. № 3. Pp. 92-96. DOI: 10.13190/jbupt.201303.96.zhanghg
- 

**UDK 004.8: 004.056.5**

## **FEATURE SELECTION FOR FUZZY CLASSIFIERS USING THE RANKING AND CROSS-VALIDATION**

**Ilya A. Hodashinsky**

Dr., Professor, e-mail: [hodashn@rambler.ru](mailto:hodashn@rambler.ru),

**Marina B. Bardamova**

Postgraduate student, e-mail: [722bmb@gmail.com](mailto:722bmb@gmail.com),

Tomsk State University of Control Systems and Radioelectronics,  
634050 Russia, Tomsk, Prospect Lenina 40

**Annotation.** The paper compares the effectiveness of modifications of the shuffled frog leaping algorithm that allow metaheuristics to function in the binary search space. Methods based on modified algebraic operations, transformation functions and fusion operations, as well as their combinations, have been tested for the task of selecting features in the fuzzy classifier. The SVC2004 data set containing a large number of features for user authentication based on dynamic handwritten signature characteristics was used in the experiment.

**Keywords:** fuzzy classifier, shuffled frog leaping algorithm, metaheuristics, feature selection, binarization

### **Reference**

1. Bardamova M. B., Anfilofiev A.E., Kovalev V.S., Filimonenko I.V. Sravnitelnyj analiz effektivnosti metaevristicheskikh algoritmov pri postroenii nechetkih klassifikatorov [Comparative analysis of the efficiency of metaheuristic algorithms for constructing fuzzy classifiers] // Mezhdunarodnaya shkola-seminar «Intellektualnye sistemy i tekhnologii: sovremennoe sostoyanie i perspektivy»: trudy = International school-seminar "Intelligent systems and technologies: current state and prospects". 2017. P. 22-31. (in Russian).
2. Al-Tashi Q., Abdulkadir S.J., Rais H.M., Mirjalili S., Alhussian H., Ragab M.G., Alqushaibi A. Binary Multi-Objective Grey Wolf Optimizer for Feature Selection in Classification // IEEE Access. 2020. vol. 8. Pp. 106247-106263. DOI: 10.1109/ACCESS.2020.3000040
3. Bardamova M. B. Construction of fuzzy classifiers with shuffle frog leaping algorithm for handwritten signature authentication // TUSUR Scientific Session: Collection of selected papers. 2019. vol. 1. № 1-2. C. 229-232.

4. Elbeltagi E., Hegazy T., Grierson D. A modified shuffled frog-leaping optimization algorithm: applications to project management // *Structure and Infrastructure Engineering*. 2007. vol. 3. № 1. Pp. 53–60. DOI: 10.1080/15732470500254535
5. Eusuff M. M., Lansey K.E. Optimizing of water distribution network design using the shuffled frog leaping algorithm // *Journal of Water Resources Planning and Management*. 2003. vol. 129. № 3. Pp. 210-225.
6. Farsangi M., Barati M. Solving unit commitment problem by a binary shuffled frog leaping algorithm // *IET Generation Transmission & Distribution*. 2014. vol. 8. № 6. Pp. 1050-1060. DOI: 10.1049/iet-gtd.2013.0436
7. Hodashinsky I., Sarin K., Shelupanov A., Slezkin A. Feature selection based on swallow swarm optimization for fuzzy classification // *Symmetry*. 2019. vol. 11. № 11. Pp. 1423. DOI: 10.3390/sym11111423
8. Jia D., Duan X., Khan M. K. Binary artificial bee colony optimization using bitwise operation // *Computers and Industrial Engineering*. 2014. vol. 76. Pp. 360–365. DOI:10.1016/j.cie.2014.08.016
9. Kiran M. S. Gunduz M. XOR-Based Artificial Bee Colony Algorithm for Binary Optimization // *Turkish Journal of Electrical Engineering and Computer Sciences*. 2013. vol. 21. 2307–2328. DOI: 10.3906/elk-1203-104
10. Mirjalili S., Lewis A. S-shaped versus V-shaped transfer functions for binary Particle Swarm Optimization // *Swarm and Evolutionary Computation*. 2013. vol. 9. Pp. 1–14. DOI: 10.1016/j.swevo.2012.09.002
11. Mirjalili S., Mohd Hashim S.Z. BMOA: binary magnetic optimization algorithm // *International Journal of Machine Learning and Computing*. 2012. vol. 2. № 3. Pp. 204–208. DOI: 10.7763/IJMLC.2012.V2.114
12. Vakil Baghmishah M.T., Madani K., Navarbah A. A discrete shuffled frog optimization algorithm // *Artificial Intelligence Review*. 2011. vol. 36. Pp. 267–284. DOI: 10.1007/s10462-011-9213-2
13. Wolpert D.H., Macready W.G. No Free Lunch Theorems for Optimization // *IEEE Transactions on Evolutionary Computation*. 1997. Vol.1. Pp. 67-82.
14. Yeung D.-Y., Chang H., Xiong Y., George S., Kashi R., Matsumoto T., Rigoll G.. SVC2004: First International Signature Verification Competition // *Conference: Biometric Authentication, First International Conference. Proceedings. Lecture Notes in Computer Science*. 2004. Pp. 16-22. DOI: 10.1007/978-3-540-25948-0\_3
15. Zhang H., Liu Y.-A. Binary Shuffled Frog Leaping Algorithm in Population Diversity for Opportunistic Spectrum Assignment // *Journal of Beijing University of Posts and Telecom*. 2013. vol. 36. № 3. Pp. 92-96. DOI: 10.13190/jbupt.201303.96.zhangh

## ПРИМЕНЕНИЕ ИНТЕГРИРОВАННОГО ПОКАЗАТЕЛЯ ОТПЕЧАТКОВ БРАУЗЕРА В ЗАДАЧЕ АДАПТИВНОЙ АУТЕНТИФИКАЦИИ СУБЪЕКТОВ ДОСТУПА

Саломатин Александр Александрович

Аспирант,

e-mail: [karateka30@mail.ru](mailto:karateka30@mail.ru),

Исхаков Андрей Юнусович

К.т.н., старший научный сотрудник,

e-mail: [iskhakovandrey@gmail.com](mailto:iskhakovandrey@gmail.com),

Институт проблем управления им. В.А. Трапезникова,  
117997, г. Москва, ул. Профсоюзная, 65.

**Аннотация.** В статье предлагается подход к аутентификации пользователя в веб-пространстве с помощью сравнения интегрированных показателей отпечатков браузера. Интегрированный показатель вычисляется дифференцированно в зависимости от класса защищаемой системы, имеющих факторов аутентификации и аппаратных характеристик вычислительных систем, используемых субъектами доступа. Также в статье рассмотрены группы идентифицирующих пользователя статических и поведенческих признаков, приведены параметры, рассчитываемые с помощью JavaScript библиотеки Fingerprint.js. Проведенный эксперимент подтвердил успешное применение предложенного подхода для нескольких наборов отпечатков браузера, полученных экспериментально.

**Ключевые слова:** технология цифровых отпечатков браузера, адаптивная аутентификация, fingerprinting, субъект доступа.

**Цитирование:** Саломатин А.А., Исхаков А. Ю. Применение интегрированного показателя отпечатков браузера в задаче адаптивной аутентификации субъектов доступа // Информационные и математические технологии в науке и управлении. 2020. № 4(20). С. 84-92. DOI:10.38028/ESI.2020.20.4.008

**Введение.** В настоящее время особенно актуальными являются вопросы кибербезопасности, связанные с защитой учетных данных субъектов доступа в критически важных инфраструктурах. Низкий уровень защищенности подсистем аутентификации и авторизации способствует повышению рисков компрометации учетных данных и кражи конфиденциальной информации посредством атак повторного воспроизведения, рефлексии с параллельным протоколом, фишинга и других. Для обеспечения защиты разрабатываются новые методы риск-ориентированной аутентификации с динамической конфигурацией факторов проверки [2], позволяющие эффективно определить, является ли пользователь легитимным.

Одним из подходов к реализации подобных систем является определение уникального многокритериального цифрового следа пользователя, основанного на оценке различных статических и поведенческих признаков профиля субъекта доступа. В рамках построения такого профиля зачастую используются следующие группы признаков: данные о браузере (версия, составляющие user-agent, доступные шрифты, шаблоны заголовков запросов,

спецификации алгоритмов шифрования, использующихся в TLS и др.), данные об операционной системе и аппаратном обеспечении (разрешение и диагональ экрана, версия операционной системы, параметры видеоадаптера, MAC-адрес и др.), данные, относящиеся к запросу (IP-адрес и сопутствующие сведения о провайдере, страна/город, RTT запроса и др.), данные о пользователе (домен электронной почты, оператор мобильной связи, профили в социальных сетях и других веб-приложениях, типичное время активности, другая информация, относящаяся к доменной модели веб-приложения) и т.д. [5]. Деление является условным, поэтому один и тот же поведенческий признак может быть вычислен в двух разных группах. Однако, глубинный анализ совокупности разнородных признаков с применением методов машинного обучения позволяет строить робастные персонифицированные модели субъектов доступа [3].

Очевидно, что определение перечня информативных статических характеристик и поведенческих признаков пользователя, а также алгоритмы определения факторов проверки необходимо адаптивно определять, исходя из специфики защищаемого ресурса и выполняемых операций [6]. В данной работе рассматривается метод определения подлинности пользователя на основе первой группы признаков - данных о браузере (отпечатков браузера).

**1. Обзор характеристик браузера для получения отпечатков.** Отпечатки браузера представляют собой уникальные значения, отражающие настройки веб-обозревателя субъекта. Ниже представлены примеры собираемых характеристик.

User-Agent – уникальная строка-идентификатор [1, 7]. Ниже приведен пример разбора строки User-Agent: "Mozilla/5.0 (X11; Linux x86\_64; en-US; rv:57.0) Gecko/20100101 Firefox/57.0":

“Mozilla/5.0”: оригинальное кодовое имя Navigator;

“X11; Linux x86\_64”: операционная система и аппаратная платформа компьютера;

“en-US”: язык локализации;

“rv:57.0”: версия верстки;

“Gecko/20100101”: кодовое наименование программного обеспечения, преобразующего содержимое веб-страниц и информацию о форматировании в интерактивное изображение форматированного содержимого на экране / сборка;

“Firefox/57.0”: имя браузера и версию.

TimeZone – название временной зоны и часовой пояс, определяемый числовым параметром смещения относительно нулевой зоны.

ScreenResolution – разрешение экрана - длина и ширина экрана в пикселях. По данным этой характеристики вычисляется параметр «Диагональ экрана». Стоит отметить, что при переводе пикселей в дюймы часть информации теряется. В связи с тем, что новый идентификатор представляется одним числом, в некоторых случаях такое переопределение может быть эффективным.

Canvas – уникальная строка-идентификатор, которая образуется путем конвертирования градиентного цветного объекта с помощью кодировки base64. Разные вычислители характеризуются различными механизмами обработки изображений, параметров экспорта и уровня сжатия. Эти функции относятся к аппаратным средствам и позволяют проводить идентификацию различных пользователей [10]. Данный признак характеризуется низкой вероятностью повторяемости и коллизий в задаче идентификации.

WebGL – уникальная строка-идентификатор, которую получают конвертацией градиентного объекта с шейдерами с помощью кодировки base64 с учетом всех исключений и возможностей WebGL (например, сглаживания, рендеринга, фильтрации).

DeviceMemory – размер оперативной памяти на устройстве.

HardwareConcurrency – максимально возможно задействованное число потоков на компьютере.

AdBlock – параметр, определяющий факт работоспособности модуля adBlock.

TouchSupport – параметр, определяющий характеристики touchscreen.

Webdriver – параметр, позволяющий определить, управляется ли клиентское приложение, использующее определенный сетевой протокол, автоматически.

Language – язык пользовательского интерфейса браузера.

Colordepth – число бит, определяющих глубину цвета для одного пикселя.

AvailableScreenResolution – доступное разрешение для окна.

SessionStorage – параметр, позволяющий определить возможность сохранения данных в сессионном хранилище.

LocalStorage – параметр, позволяющий определить возможность сохранения данных в локальном хранилище;

IndexedDb – параметр, позволяющий определить возможность сохранения данных в NOSQL хранилище на стороне клиента.

AddBehavior – параметр, позволяющий определить возможность использования поведенческих признаков.

OpenDatabase – параметр, позволяющий определить возможность создания объекта базы данных SQL Lite Database.

CpuClass – класс центрального процессора операционной системы пользователя.

Platform – название платформы браузера, которая представляет прямой способ взаимодействия приложений пользователя с операционной системой Windows.

Plugins – массив сведений о плагинах, установленных в приложении. Плагин представляет собой самостоятельный программный модуль, подключаемый к основной программе и предназначенный для расширения используемых возможностей. Данный параметр включает наименования плагинов, описания плагинов и их типы.

HasLiedLanguages – параметр, проверяющий факт совпадения языка пользовательского интерфейса браузера (отпечаток «Language») с первым языком в списке наиболее предпочитаемых языков пользователя.

HasLiedResolution – параметр, проверяющий факт совпадения разрешение экрана («screenResolution») с доступным разрешением экрана («availableScreenResolution»).

HasLiedOs – параметр, проверяющий факт соответствия между данными об ОС (операционной системе), извлеченными из «userAgent», данными о платформе, классе ЦПУ (central processing unit – центрального процессорного устройства), поддержке touch-screen.

HasLiedBrowser – параметр, проверяющий факт соответствия между данными о браузере, извлеченными из «User-Agent», и данными о том, как браузер справляется с искусственно созданной ошибкой.

Fonts – список шрифтов, доступных браузеру.

Audio – число, отражающее сумму буферных значений. Для получения данного параметра веб-сайт отправляет запрос браузеру на моделирование синусоида, основанного

на результатах аудио-стека устройства. Затем результат отправляется серверу и используется аналогично с энтропией для уникальной идентификации.

## 2. Метод сравнения отпечатков браузера на основе вычисления интегрированного параметра уникальности пользователя.

В качестве интегрированного показателя уникальности пользователя в предлагаемом методе используется параметр Fingerprinting ID. Схема его получения выглядит следующим образом:

- 1) получение информации по отпечаткам браузера [8];
- 2) определение весов отпечатков;
- 3) объединение полученной информации в интегрированную строку;
- 4) выполнение процедуры хеширования [9, 10].

Для определения весов используется аналитический-иерархический процесс (АИП), в котором используют три критерия: информативность, скорость получения и сложность получения [11].

АИП генерирует вес для каждого критерия в соответствии с попарными сравнениями критерия, а затем присваивает оценку каждому варианту для получения фиксированного критерия. Далее веса критериев и оценки опций объединяются, чтобы определить глобальные оценки для каждого варианта, являющиеся весами соответствующих объектов. На следующем этапе полученные веса в тех же пропорциях переопределяют таким образом, чтобы их сумма была равна 1.

Представим АИП в данном случае с точки зрения математической формализации. Пусть даны  $N$  браузерных отпечатков и 3 критерия. Тогда веса критериев представлены вектором  $\lambda = [\lambda_1 \lambda_2 \lambda_3]^T$ ;  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ , а оценки опций по критериям представлены матрицей  $W = [w_{ij}]$ , где  $w_{1j} + w_{2j} + \dots + w_{Nj} = 1$ ;  $1 \leq i \leq N, 1 \leq j \leq 3; i, j \in Z$ . Для данной матрицы  $i$  - номер отпечатка браузера,  $j$  - номер критерия. Также обозначим  $I = [w_{i1}], F = [w_{i2}], D = [w_{i3}]$ . Для получения весов отпечатков браузера используется формула:

$$W^* = W\lambda \quad (1)$$

Затем, с целью обеспечения различимости отпечатков пользователей используется основной алгоритм подобию. Алгоритм подобию основан на сравнении частей цифрового следа пользователя между собой. Если они совпадают, то присваивается значение 1, иначе – 0. В итоге формируется вектор  $s$  из нулей и единиц. Далее вычисляется число подобию

$$P = sW', \quad (2)$$

определяющее вероятность того, что пользователь является легитимным. Здесь  $W'$  – вектор-столбец весов браузерных отпечатков при их сравнении.

Заданное число подобию сравнивается с пороговым показателем. В случае не достижения показателя данная операция помечается, как подозрение на инцидент информационной безопасности и выполняется активная реакция системы защиты в соответствии с политиками безопасности: адаптивный подбор дополнительного фактора аутентификации, блокировка субъекта и др.

**3. Описание эксперимента.** На первом этапе для получения информации о отпечатках браузера была использована библиотека fingerprint2.js. С помощью программного кода на языке программирования JavaScript были получены 28 параметров (примеры представлены на рисунке 1).

```

▶ 0: {key: "userAgent", value: "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit, like Gecko) Chrome/85.0.41...
▶ 1: {key: "webdriver", value: "not available"}
▶ 2: {key: "language", value: "ru-RU"}
▶ 3: {key: "colorDepth", value: 24}
▶ 4: {key: "deviceMemory", value: 8}
▶ 5: {key: "hardwareConcurrency", value: 4}
▶ 6: {key: "screenResolution", value: Array(2)}
▶ 7: {key: "availableScreenResolution", value: Array(2)}
▶ 8: {key: "timezoneOffset", value: -180}
▶ 9: {key: "timezone", value: "Europe/Moscow"}
▶ 10: {key: "sessionStorage", value: true}
▶ 11: {key: "localStorage", value: true}
▶ 12: {key: "indexedDb", value: true}
▶ 13: {key: "addBehavior", value: false}
▶ 14: {key: "openDatabase", value: true}
▶ 15: {key: "cpuClass", value: "not available"}
▶ 16: {key: "platform", value: "win32"}
▶ 17: {key: "plugins", value: Array(3)}
▶ 18: {key: "canvas", value: Array(2)}
▶ 19: {key: "webgl", value: Array(65)}
▶ 20: {key: "webglVendorAndRenderer", value: "Google Inc.-ANGLE (Intel(R) HD Graphics 630 Direct3D11 vs_5_0 ps_5...
▶ 21: {key: "adBlock", value: false}
▶ 22: {key: "hasLiedLanguages", value: false}
▶ 23: {key: "hasLiedResolution", value: false}
▶ 24: {key: "hasLiedOs", value: false}
▶ 25: {key: "hasLiedBrowser", value: false}
▶ 26: {key: "touchSupport", value: Array(3)}
▶ 27: {key: "fonts", value: Array(33)}
▶ 28: {key: "audio", value: "124.04347527516074"}

```

**Рис. 1.** Вычисление отпечатков браузера в fingerprint2.js

На следующем этапе необходимо определить веса отпечатков браузера с помощью АИП.

Рассмотрим первый критерий – информативность. Значение браузерного отпечатка по данному критерию будем задавать как число попарных сравнений, в которых данный отпечаток браузера является более информативным. Считаем, что строки-идентификаторы уникальнее других параметров, не являющихся строковыми, а результаты других видов сравнений равноценны. Тогда этот набор определяется массивом  $I$ , тогда  $I = [0.190 \ 0.008 \ 0.008 \ \dots \ 0.206 \ 0.214 \ 0.198 \ 0.008 \ \dots \ 0.008]^T$ .

Рассмотрим второй критерий – время получения. Значение отпечатка браузера по данному критерию будем задавать как число попарных сравнений, в которых данный отпечаток браузера был в среднем получен быстрее других. Пусть этот набор определяется массивом  $F$ , где время получения отпечатков одинаковое, тогда  $F = [0.036 \ 0.036 \ \dots \ 0.036]^T$ .

Рассмотрим третий критерий – сложность получения. Так как отпечатки браузера были получены одним способом комплексно, то зададим сложность получения для каждого отпечатка браузера как равную. Пусть набор определяется массивом  $D$ , тогда  $D = [0.036 \ 0.036 \ \dots \ 0.036]^T$ .

Предположим, что все три критерия являются равноценными, тогда после перемножения на коэффициенты важности критериев  $\lambda = [0.33; 0.33; 0.33]^T$  можно будет получить показатели, определяющие значимость каждого отпечатка браузера  $W^* = [0.08712 \ 0.027 \ 0.027 \ \dots \ 0.092 \ 0.095 \ 0.089 \ 0.027 \ \dots \ 0.027]^T$ .

С помощью хеш-функции `md5` был сформирован показатель `fingerprintID` (с 28 элементами). Рассмотрим другие условия аутентификации. Допустим, что изменилось значение двух наиболее информативных признаков – `canvas` и `webGL`, а остальные значения не подвергались изменению. В результате проведения аналогичных вычислений был получен

показатель fingerprintID, отличающийся по двум параметрам. Проведем сравнение и получим следующие результаты вектора подобия:  $s = [1, 1, \dots, 0, 0, 1, \dots, 1]$ . Веса для вычисления степени подобия зададим из  $W' = W^*$ . Вычислим степень подобия:

$P = 1 * 0.027 + 1 * 0.027 + \dots + 0 * 0.092 + 0 * 0.095 + 1 * 0.089 + 1 * 0.027 + 1 * 0.027 = 1 - 0.092 - 0.095 = 0.81$ . Считаем заданным пороговое значение, равное 0,75. Следовательно, пользователь является легитимным с вероятностью 0,81 (81%).

**Заключение.** В статье предложен подход к аутентификации пользователя в веб-пространстве на основе интегрированного показателя отпечатков браузера.

Первый раздел посвящен описанию технологии отпечатков браузера, рассмотрены группы идентифицирующих пользователя статических и поведенческих признаков, приведены параметры, рассчитываемые с помощью JavaScript библиотеки fingerprint2.js. Рассмотрены основные типы полученных признаков и приведена интерпретация некоторых значений. Для наиболее информативных признаков были приведены примеры значений и раскрыта практическая значимость оцениваемых параметров. Представлено формализованное описание предложенного метода. Показано, что основой для определения подлинности пользователя является сравнение степени подобия цифровых следов, в качестве которых выступают наборы отпечатков браузера, с заданным пороговым значением. Также описана поэтапная процедура вычисления интегрированных показателей отпечатков для заданных начальных условий. В третьем разделе представлены результаты эксперимента, в котором рассмотренный метод применялся при сравнении двух цифровых следов, один из которых был выявлен практически с помощью программного кода и математических вычислений, а второй – на основе вычислений и предложенных изменений значений нескольких отпечатков браузера. Результат эксперимента подтвердил эффективность предложенного подхода для конкретного практически полученного набора отпечатков браузера.

Исследование выполнено при частичной финансовой поддержке гранта Президента Российской Федерации в рамках научного проекта № МК-2421.2020.9.

#### СПИСОК ЛИТЕРАТУРЫ

1. Агафонов Ю.М. Деанонимизация пользователей на основе цифровых отпечатков браузера // Безопасность информационного пространства: сборник трудов XVI Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых. Екатеринбург: Уральский федеральный университет имени первого Президента России Б.Н. Ельцина. 2018. С. 3-5
2. Бессонова Е.Е. Метод идентификации пользователей в сети Интернет с использованием компонентного профиля: дис. ... канд. тех. наук: 05.13.19.- Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики. Спб.. 2014. . 115 с.
3. Исхаков А.Ю. Схемы аутентификации пользователя в СКУД с использованием QR кодов и передачи данных по технологии // Информационное противодействие угрозам терроризма. 2014. № 22. С. 11-15.
4. Исхаков А.Ю., Исхакова А.О., Мещеряков Р.В., Бендрау Р., Мелехова О. Использование тепловой карты поведения пользователя в задаче идентификации субъекта инцидента информационной безопасности // Труды СПИИРАН. 2018. № 6 (61). С. 147-171.

5. Платонов Т.С., Оголюк А.А. Межсетевые экраны уровня веб-приложения в современном мире // Программная инженерия и компьютерная техника (майоровские чтения). Спб.. 2019. С. 106-119
6. Решетников С.Ю. Разработка метода идентификации анонимных пользователей сети TOR // Молодежь и современные информационные технологии: сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых. Томск: Национальный исследовательский Томский политехнический университет. 2016. С. 58–60.
7. ElBanna A., Abdelbaki N. Browsers Fingerprinting Motives, Methods, and Countermeasures // 2018 International Conference on Computer, Information and Telecommunication Systems (CITS). Colmar. 2018. Pp. 1-5
8. N. I. Daud, G. R. Haron and S. S. S. Othman, "Adaptive authentication: Implementing random canvas fingerprinting as user attributes factor" 2017 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE). Langkawi. 2017. Pp. 152-156. DOI:10.1109/ISCAIE.2017.8074968.
9. Luangmaneerote S., Zaluska E., Carr L. Inhibiting Browser Fingerprinting and Tracking // IEEE 3rd International Conference on Big Data Security on Cloud. Beijing. 2017. Pp. 63-68
10. Nair K., RoseLalson E. The Unique Id's You Can't Delete: Browser Fingerprints // 2018 International Conference on Emerging Trends and Innovations in Engineering and Technological Research (ICETIETR). Ernakulam. 2018. Pp. 1-5
11. Tracking Your Browser with High-Performance Browser Fingerprint Recognition Model / [Jiang W. and etc.] // China communications. № 17(3). Beijing, 2020. Pp. 168-175

---

**UDK 004.056.53**

**APPLICATION OF THE INTEGRATED INDICATOR OF BROWSER  
FINGERPRINTING IN THE PROBLEM OF ADAPTIVE AUTHENTICATION OF  
ACCESS SUBJECTS**

**Alexander A. Salomatín**

Graduate student,

e-mail: [karateka30@mail.ru](mailto:karateka30@mail.ru),

**Andrey Yu. Iskhakov**

PhD, senior researcher,

e-mail: [iay@ipu.ru](mailto:iay@ipu.ru),

V.A.Trapeznikov Institute of Control Sciences  
of Russian Academy of Sciences  
117997, Profsoyuznaya Str., 65, Moscow, Russia.

**Abstract.** This article presents an approach to user authentication in the web space by comparing the integrated metrics of browser fingerprints. The integrated indicator is calculated differentially depending on the class of the protected system, the available

authentication factors and the hardware characteristics of the computing systems used by the access subjects. The experiment carried out confirmed the successful application of the method for a specific practically obtained set of browser fingerprints.

**Keywords:** digital browser fingerprint technology, fingerprinting, adaptive authentication, access subject

### References

1. Agafonov U.M. Deanonimizaciya pol'zovatelej na osnove cifrovyyh otpechatkov brauzera [Deanonymization of users based on digital fingerprint] // Bezopasnost' informacionnogo prostranstva: sbornik trudov XVI Vserossijskoj nauchno-prakticheskoy konferencii studentov, aspirantov i molodyh uchenyh = Information Space Security: Collection of Proceedings of the XVI All-Russian Scientific and Practical Conference of Students, Postgraduates and Young Scientists. Ekaterinburg. Ekaterinburg: Ural'skij federal'nyj universitet imeni pervogo Prezidenta Rossii B.N. El'cina = Ural Federal university named after the first President of Russia B.N. Elcin. 2018. Pp. 3-5
2. Bessonova E.E. Metod identifikacii pol'zovatelej v seti Internet s ispol'zovaniem komponentnogo profilya: dis. ... kand. tekh. nauk: 05.13.19 [A method for identifying users on the Internet using a component profile: diss....of the cand. of tech. sc: 05.13.19]. - Sankt-Peterburgskij nacional'nyj issledovatel'skij universitet informacionnyh tekhnologij, mekhaniki i optiki = Saint Petersburg National Research University of Information Technologies, Mechanics and Optics. Spb. 2014. 115 p.
3. Iskhakov A.YU. Skhemy autentifikacii pol'zovatelya v SKUD s ispol'zovaniem QR kodov i peredachi dannyh po tekhnologii NFC [User authentication schemes in ACS using QR codes and data transfer using NFC technology] / A. YU. Iskhakov, R. V. Meshcheryakov. - Tekst: neposredstvennyj // Informacionnoe protivodejstvie ugrozam terrorizma = Information countermeasures against terrorism threats. 2014. № 22. Pp. 11-15.
4. Iskhakov A.YU., Iskhakova A.O., Meshcheryakov R.V., Bendrau R., Melekhova O. Ispol'zovanie teplovoj karty povedeniya pol'zovatelya v zadache identifikacii sub"ekta incidenta informacionnoj bezopasnosti [Using a heatmap of user behavior in the problem of identifying the subject of an information security incident] // Trudy SPIIRAN = Proceedings of the SPIIRAS. 2018. № 6 (61). Pp. 147-171.
5. Platonov T.S., Ogolyuk A.A. Mezhssetevye ekrany urovnya veb-prilozheniya v sovremennom mire [Web Application Layer Firewalls in the Modern World] // Programmnyaya inzheneriya i komp'yuternaya tekhnika (majorovskie chteniya): konf = Software engineering and computer technology (major's readings): conf. Spb. 2019. Pp. 106-119
6. Reshetnikov S.YU. Razrabotka metoda identifikacii anonimnyh pol'zovatelej seti TOR [Development of a method for identifying anonymous users in the TOR network] // Molodezh' i sovremennye informacionnye tekhnologii: sbornik trudov XIV Mezhdunarodnoj nauchno-prakticheskoy konferencii studentov, aspirantov i molodyh uchenyh = Youth and modern information technologies: collection of works of the XIV International scientific-practical conference of students, graduate students and young scientists, Tomsk. - Tomsk: Nacional'nyj issledovatel'skij Tomskij politekhnicheskij universitet = National Research Tomsk Polytechnic University. 2016. Pp. 58–60.

7. ElBanna A., Abdelbaki N. Browsers Fingerprinting Motives, Methods, and Countermeasures // 2018 International Conference on Computer, Information and Telecommunication Systems (CITS). Colmar. 2018. Pp. 1-5
8. N. I. Daud, G. R. Haron and S. S. S. Othman, "Adaptive authentication: Implementing random canvas fingerprinting as user attributes factor," 2017 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE). Langkawi. 2017. Pp. 152-156. DOI:10.1109/ISCAIE.2017.8074968.
9. Luangmaneerote S., Zaluska E., Carr L. Inhibiting Browser Fingerprinting and Tracking // IEEE 3rd International Conference on Big Data Security on Cloud. Beijing, 2017. Pp. 63-68
10. Nair K., RoseLalson E. The Unique Id's You Can't Delete: Browser Fingerprints // 2018 International Conference on Emerging Trends and Innovations in Engineering and Technological Research (ICETIETR). – Ernakulam. 2018. Pp. 1-5
11. Tracking Your Browser with High-Performance Browser Fingerprint Recognition Model / [Jiang W. and etc.] // China communications. - № 17(3). Beijing, 2020. Pp. 168-175
12. This work was partially funded by Russian Federation President Grant for the young scientists [MK-2421.2020.9]

УДК 519.71+ 681.5+ 62.752

## СИСТЕМНЫЕ ПРЕДСТАВЛЕНИЯ ДИНАМИЧЕСКИХ ПРОЦЕССОВ В МЕХАНИЧЕСКИХ КОЛЕБАТЕЛЬНЫХ СИСТЕМАХ С ОСОБЕННОСТЯМИ

Елисеев Андрей Владимирович

К.т.н., доцент кафедры «Математика», e-mail: [eavsh@ya.ru](mailto:eavsh@ya.ru),

Иркутский государственный университет путей сообщения

664074, г.Иркутск, ул. Чернышевского 15.

**Аннотация.** Рассматриваются подходы в формировании методологического базиса системного анализа колебательных структур на примере упруго-диссипативных механических систем с двумя степенями свободы. Целью исследования является разработка метода оценки свойств механических колебательных систем с учетом сил вязкого трения на основе частотных функций и функции демпфирования в зависимости от коэффициентов форм связности. Для построения математических моделей используется формализм Лагранжа, матричные методы, элементы теории функций комплексной переменной. Представлены понятия частотной функции и функции демпфирования, отражающие особенности соотношения потенциальной, кинетической энергии системы с учетом сил вязкого трения, представленных диссипативной функцией. Разработан метод построения частотной функции и функции демпфирования, отражающих динамические особенности свободных движений с учетом сил трения в зависимости от коэффициента форм связности. На основе полученных общих аналитических выражений частотных функций и функций демпфирования проведен анализ особых вариантов механических колебательных систем, представляющих интерес на начальном этапе исследования. Разработанный метод построения частотной функции и функции демпфирования может быть использован для отображения динамических форм связности движений механических колебательных систем. Предложенный метод построения частотной функции и функции демпфирования может быть обобщен на механические колебательные системы, рассматриваемые в различных системах координат.

**Ключевые слова:** механическая система, динамические связи, частотная функция, функция демпфирования, связность движений, экстремальные свойства, колебание, вязкое трение.

**Цитирование:** Елисеев А. В. Системные представления динамических процессов в механических колебательных системах с особенностями // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 93-107. DOI: 10.38028/ESI.2020.20.4.009

**Введение.** Оценке динамических свойств технических объектов, работающих в условиях интенсивных вибрационных нагрузок, уделяется существенное внимание [1, 2, 9-11]. К числу известных подходов можно отнести идеи оценки особенностей динамических свойств систем через отношение потенциальной и кинетической энергии [6, 7]. Использование частотной функции для исследования динамики механических колебательных систем без

учета сил трения инициирует интерес к расширению области её применения в задачах динамики, связанных с учетом так называемых коэффициентов форм связности [3-5, 8].

Вместе с тем, для механических колебательных систем с учетом сил вязкого трения методы построения частотной функции требуют детализации в зависимости от величины вязкого трения, в частности, для систем с апериодическим движением понятие частоты колебания может терять смысл, и тогда функцией, отражающей особенности движения системы, становится функция демпфирования, также зависящая от коэффициента форм связности движения.

Предлагаемая работа посвящена развитию метода оценки свойств механических колебательных систем с учетом сил вязкого трения на основе частотных функций и функции демпфирования в зависимости от коэффициентов форм связности.

**I. Основные положения. Постановка задачи.** Рассматриваются свободные движения механической упруго-диссипативной системы с сосредоточенными параметрами с двумя степенями свободы. Принципиальная схема системы представлена на рисунке 1.

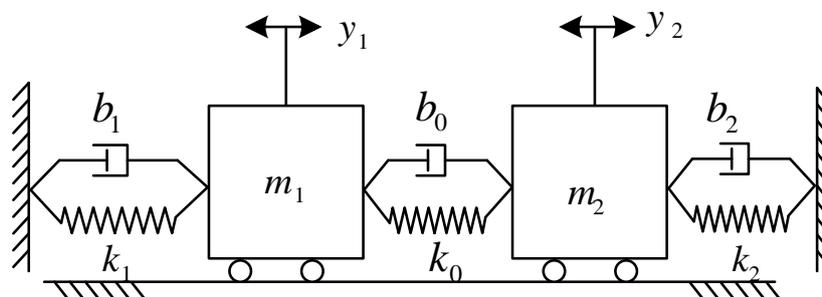


Рис. 1. Механическая колебательная система с учетом вязкого трения

Обобщенные координаты  $y_1, y_2$  обозначают смещение массо-инерционных элементов  $m_1, m_2$  относительно положения статического равновесия. Кинетическая энергия  $T$ , потенциальная энергия  $\Pi$  и функция рассеяния  $F$  имеют вид:

$$T = \frac{1}{2} m_1 \dot{y}_1^2 + \frac{1}{2} m_2 \dot{y}_2^2, \quad (1)$$

$$\Pi = \frac{1}{2} k_1 y_1^2 + \frac{1}{2} k_2 y_2^2 + \frac{1}{2} k_0 (y_2 - y_1)^2, \quad (2)$$

$$F = \frac{1}{2} b_1 \dot{y}_1^2 + \frac{1}{2} b_0 (\dot{y}_2 - \dot{y}_1)^2 + \frac{1}{2} b_2 \dot{y}_2^2. \quad (3)$$

Система уравнений Лагранжа второго рода имеет вид:

$$\begin{cases} \frac{d}{dt} \frac{\partial T}{\partial \dot{y}_1} + \frac{\partial \Pi}{\partial y_1} + \frac{\partial F}{\partial \dot{y}_1} = 0; \\ \frac{d}{dt} \frac{\partial T}{\partial \dot{y}_2} + \frac{\partial \Pi}{\partial y_2} + \frac{\partial F}{\partial \dot{y}_2} = 0. \end{cases} \quad (4)$$

После подстановки выражений  $T, \Pi, F$  система дифференциальных уравнений (4) принимает вид:

$$\begin{cases} m_1 \ddot{y}_1 + (b_0 + b_1) \dot{y}_1 - b_0 \dot{y}_2 + (k_0 + k_1) y_1 - k_0 y_2 = 0; \\ m_2 \ddot{y}_2 + (b_0 + b_2) \dot{y}_2 - b_0 \dot{y}_1 + (k_0 + k_2) y_2 - k_0 y_1 = 0. \end{cases} \quad (5)$$

Формы свободных движений представленной системы (5) в общем случае определяются собственными числами характеристического уравнения с учетом их кратности. Рассматривается случай простых корней. Таким образом, пусть решение  $y_1 = y_1(t)$ ,  $y_2 = y_2(t)$  системы (5) представимо в виде:

$$\vec{y} = \vec{Y}e^{pt}, \quad (6)$$

где  $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$  – вектор-решение,  $\vec{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$  – числовой вектор,  $p = \sigma + j\omega$  – комплексный параметр,  $t$  – переменная времени. Полагается, что начальные условия согласованы с видом искомого решения (6).

Задача заключается в построении и оценке экстремальных свойств функций, отображающих характеристики собственных движений системы с учетом сил вязкого трения.

**II. Построение частотной функции и функции диссипации на основе энергетического соотношения.** Система (5) в обозначениях (6) имеет вид:

$$\begin{bmatrix} m_1 p^2 + (b_0 + b_1)p + k_0 + k_1 & -b_0 p - k_0 \\ -b_0 p - k_0 & m_2 p^2 + (b_0 + b_2)p + k_0 + k_2 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = 0. \quad (7)$$

Введем обозначения:

$$A = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \quad B = \begin{bmatrix} b_0 + b_1 & -b_0 \\ -b_0 & b_0 + b_2 \end{bmatrix}, \quad C = \begin{bmatrix} k_0 + k_1 & -k_0 \\ -k_0 & k_0 + k_2 \end{bmatrix}. \quad (8)$$

С учетом (8) матричное соотношение (5) принимает вид:

$$(p^2 A + pB + C)\vec{Y} = 0. \quad (9)$$

На основе матричного соотношения (9) могут быть получены различные скалярные уравнения, с учетом экстремальных свойств которых, в свою очередь, могут быть определены свойства решения  $p$  уравнения (9). В частности, скалярное умножение левой и правой части равенства (9) на вектор  $\vec{Y}$  приводит к скалярному выражению:

$$p^2 \langle A\vec{Y}, \vec{Y} \rangle + p \langle B\vec{Y}, \vec{Y} \rangle + \langle C\vec{Y}, \vec{Y} \rangle = 0. \quad (10)$$

С помощью подстановки  $p = \sigma + j\omega$  приведем (10) к виду:

$$(\sigma^2 - \omega^2 + 2j\sigma\omega) \langle A\vec{Y}, \vec{Y} \rangle + (\sigma + j\omega) \langle B\vec{Y}, \vec{Y} \rangle + \langle C\vec{Y}, \vec{Y} \rangle = 0, \quad (11)$$

Пусть для координат вектора  $\vec{Y}$  выполнено соотношение:

$$Y_2 = \alpha Y_1, \quad (12)$$

где  $\alpha$  – коэффициент формы связности. В этом случае вектор  $\vec{Y}$  может быть представлен в виде:

$$\vec{Y} = Y_1 \vec{\alpha}, \quad (13)$$

где  $\vec{\alpha} = \begin{bmatrix} 1 \\ \alpha \end{bmatrix}$  – вектор, определяемый коэффициентом связности  $\alpha$ . После подстановки (13)

выражение (11) принимает вид:

$$(\sigma^2 - \omega^2 + 2j\sigma\omega)A_\alpha + (\sigma + j\omega)B_\alpha + C_\alpha = 0, \quad (14)$$

где  $A_\alpha = \langle A\vec{\alpha}, \vec{\alpha} \rangle$ ,  $B_\alpha = \langle B\vec{\alpha}, \vec{\alpha} \rangle$ ,  $C_\alpha = \langle C\vec{\alpha}, \vec{\alpha} \rangle$  – скалярные функции коэффициента формы связности  $\alpha$ . Уравнение (14) может быть представлено в эквивалентном виде:

$$\begin{cases} \omega^2 A_\alpha = \sigma^2 A_\alpha + \sigma B_\alpha + C_\alpha; \\ 2\sigma\omega A_\alpha + \omega B_\alpha = 0. \end{cases} \quad (15)$$

Найдём решение системы (15) относительно  $\sigma, \omega$ , как функций  $\alpha$ . Функция  $\omega(\alpha)$  рассматривается как частотная функция,  $\sigma(\alpha)$  – функция демпфирования. Особенности решения системы определяются знаком выражения  $\sigma^2 A_\alpha + \sigma B_\alpha + C_\alpha$ .

*Учет диссипации.* Уровень диссипации энергии может быть охарактеризован дискриминантом  $B_\alpha^2 - 4A_\alpha C_\alpha$ . При условии:

$$B_\alpha^2 < 4A_\alpha C_\alpha, \quad (16)$$

понимаемом, как малость трения, решение (15) может быть представлено в виде:

$$\begin{cases} \omega^2 = \frac{C_\alpha}{A_\alpha} - \left(\frac{B_\alpha}{2A_\alpha}\right)^2; \\ \sigma = -\frac{B_\alpha}{2A_\alpha}. \end{cases} \quad (17)$$

Необходимо отметить, что при выполнении условий малости сил вязкого трения (16) выполнено равенство:

$$\omega^2 + \sigma^2 = \frac{C_\alpha}{A_\alpha}. \quad (18)$$

При условии «больших сил вязкого трения»:

$$B_\alpha^2 > 4A_\alpha C_\alpha, \quad (19)$$

решение (15) может быть представлено в виде:

$$\omega = 0, \quad (20) \quad \sigma_1(\alpha) = -\frac{B_\alpha}{2A_\alpha} - \sqrt{\left(\frac{B_\alpha}{2A_\alpha}\right)^2 - \frac{C_\alpha}{A_\alpha}}, \quad \sigma_2(\alpha) = -\frac{B_\alpha}{2A_\alpha} + \sqrt{\left(\frac{B_\alpha}{2A_\alpha}\right)^2 - \frac{C_\alpha}{A_\alpha}}. \quad (21)$$

Функции демпфирования имеют две компоненты  $\sigma_1(\alpha)$ ,  $\sigma_2(\alpha)$ , такие, что:

$$\sigma_1(\alpha) + \sigma_2(\alpha) = -\frac{B_\alpha}{A_\alpha}, \quad \sigma_1(\alpha) \cdot \sigma_2(\alpha) = \frac{C_\alpha}{A_\alpha}. \quad (22)$$

Таким образом, в зависимости от уровня сил вязкого трения возможны различные виды представления частотной функции и функции демпфирования. Если  $B_\alpha^2 - 4A_\alpha C_\alpha < 0$ , то определена частотная функция  $\omega^2(\alpha)$  и одна компонента функции демпфирования  $\sigma(\alpha)$ . Если  $B_\alpha^2 - 4A_\alpha C_\alpha > 0$ , то предполагается, что частотная функция  $\omega^2(\alpha)$  принимает нулевые значения, а функция демпфирования имеет две различные отрицательные компоненты  $\sigma_1(\alpha)$ ,  $\sigma_2(\alpha)$ . Что касается условия  $B_\alpha^2 - 4A_\alpha C_\alpha = 0$ , то оно требует отдельного анализа. Вместе с тем, условие  $B_\alpha^2 - 4A_\alpha C_\alpha = 0$  может быть интерпретировано как граница двух различных режимов движения механической системы.

Представленные аналитические выражения частотной функции, функции демпфирования и условий «малых» и «больших» сил вязкого трения могут быть детализированы при рассмотрении частных вариантов механических колебательных систем, полученных на основе системы с двумя степенями свободы.

**III. Частотная функция и функция демпфирования для механической системы с двумя степенями свободы.** Рассматриваются варианты параметров для механической системы, представленной на рисунке 1. Предполагается, что множество граничных параметров, разделяющих режимы движения при малых и больших силах вязкого трения, определяется уравнением:

$$B_{\alpha}^2 = 4A_{\alpha}C_{\alpha}, \quad (23)$$

где:

$$A_{\alpha} = m_1 + m_2\alpha^2, \quad (24) \quad B_{\alpha} = (b_0 + b_2)\alpha^2 - 2\alpha b_0 + b_0 + b_1, \quad (25) \quad C_{\alpha} = (k_0 + k_2)\alpha^2 - 2\alpha k_0 + k_0 + k_1. \quad (26)$$

Условия малости сил вязкого трения имеют форму:

$$B_{\alpha}^2 < 4A_{\alpha}C_{\alpha}. \quad (27)$$

В этом случае частотная функция  $\omega^2(\alpha)$  и функция демпфирования  $\sigma(\alpha)$ :

$$\begin{cases} \omega^2(\alpha) = \frac{(k_0 + k_2)\alpha^2 - 2\alpha k_0 + k_0 + k_1}{m_1 + m_2\alpha^2} - \left(\frac{1}{2} \frac{(b_0 + b_2)\alpha^2 - 2\alpha b_0 + b_0 + b_1}{m_1 + m_2\alpha^2}\right)^2 \\ \sigma(\alpha) = -\frac{1}{2} \frac{(b_0 + b_2)\alpha^2 - 2\alpha b_0 + b_0 + b_1}{m_1 + m_2\alpha^2} \end{cases}. \quad (28)$$

Условия больших сил вязкого трения имеют форму:

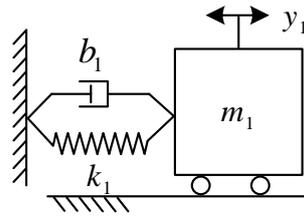
$$B_{\alpha}^2 > 4A_{\alpha}C_{\alpha}. \quad (29)$$

При условиях (29) функции  $\omega^2$  и  $\sigma(\alpha)$  имеют вид:

$$\begin{cases} \omega^2 = 0; \\ \sigma_1(\alpha) = -\frac{1}{2} \frac{(b_0 + b_2)\alpha^2 - 2\alpha b_0 + b_0 + b_1}{m_1 + m_2\alpha^2} - \sqrt{\left(\frac{1}{2} \frac{(b_0 + b_2)\alpha^2 - 2\alpha b_0 + b_0 + b_1}{m_1 + m_2\alpha^2}\right)^2 - \frac{(k_0 + k_2)\alpha^2 - 2\alpha k_0 + k_0 + k_1}{m_1 + m_2\alpha^2}}; \\ \sigma_2(\alpha) = -\frac{1}{2} \frac{(b_0 + b_2)\alpha^2 - 2\alpha b_0 + b_0 + b_1}{m_1 + m_2\alpha^2} + \sqrt{\left(\frac{1}{2} \frac{(b_0 + b_2)\alpha^2 - 2\alpha b_0 + b_0 + b_1}{m_1 + m_2\alpha^2}\right)^2 - \frac{(k_0 + k_2)\alpha^2 - 2\alpha k_0 + k_0 + k_1}{m_1 + m_2\alpha^2}}. \end{cases} \quad (31)$$

Представленные выражения отражают движения в форме экспоненциального убывания при отсутствии колебаний. На основе общего варианта механической колебательной системы могут быть построены частные случаи.

**IV. Частотная функция и функция демпфирования для механической системы с одной степенью свободы.** Особенным, в некотором роде вырожденным, случаем является система с одной степенью свободы, так как представляет собою вариант парциальной системы. Для рассмотрения системы с одной степенью свободы положим, что в исходной системе на рис. 1  $b_0 = 0$ ,  $b_2 = 0$ ,  $k_0 = 0$ ,  $k_2 = 0$ ,  $m_2 = 0$ . Принципиальная схема системы представлена на рис. 2. Необходимо отметить, что рассматриваемая система (рис. 2) не является системой с двумя степенями свободы. Вместе с тем, соответствующие системе частотные функции и функции демпфирования представляют интерес на начальном этапе исследования в качестве базы для сравнения.



**Рис. 2.** Принципиальная схема системы с одной степенью свободы

Система дифференциальных уравнений имеет вид:

$$m_1 \ddot{y}_1 + b_1 \dot{y}_1 + k_1 y_1 = 0. \quad (32)$$

Величины  $A_\alpha$ ,  $B_\alpha$ ,  $C_\alpha$  имеют вид:

$$A_\alpha = m_1, \quad (33)$$

$$B_\alpha = b_1, \quad (34)$$

$$C_\alpha = k_1. \quad (35)$$

Множество граничных параметров определяется уравнением:

$$\left(\frac{b_1}{2m_1}\right)^2 = \frac{k_1}{m_1}. \quad (36)$$

Условие малости сил трения имеет вид:

$$\left(\frac{b_1}{2m_1}\right)^2 < \frac{k_1}{m_1}. \quad (37)$$

Частотная функция  $\omega^2(\alpha)$  и функция демпфирования  $\sigma(\alpha)$  принимают вид:

$$\begin{cases} \omega^2(\alpha) = \frac{k_1}{m_1} - \left(\frac{1}{2} \frac{b_1}{m_1}\right)^2; \\ \sigma(\alpha) = -\frac{1}{2} \frac{b_1}{m_1}. \end{cases} \quad (38)$$

Условий больших сил трения имеют вид:

$$\left(\frac{b_1}{2m_1}\right)^2 > \frac{k_1}{m_1}. \quad (39)$$

Частотная функция  $\omega^2(\alpha)$  и функция демпфирования  $\sigma(\alpha)$ :

$$\begin{cases} \omega^2 = 0; \\ \sigma_1(\alpha) = -\frac{b_1}{2m_1} - \sqrt{\left(\frac{b_1}{2m_1}\right)^2 - \frac{k_1}{m_1}}; \\ \sigma_2(\alpha) = -\frac{b_1}{2m_1} + \sqrt{\left(\frac{b_1}{2m_1}\right)^2 - \frac{k_1}{m_1}}, \end{cases} \quad (40)$$

где функция демпфирования  $\sigma(\alpha)$  представлена компонентами  $\sigma_1(\alpha)$ ,  $\sigma_2(\alpha)$  для  $\alpha$ , для которого выполнено условие (39).

В случае граничных параметров:

$$\left(\frac{b_1}{2m_1}\right)^2 = \frac{k_1}{m_1}. \quad (41)$$

частотная функция  $\omega^2(\alpha)$  и функция демпфирования  $\sigma(\alpha)$  имеют вид:

$$\begin{cases} \omega^2(\alpha) = 0; \\ \sigma(\alpha) = -\frac{b_1}{2m_1}. \end{cases} \quad (42)$$

При этом необходимо заметить, что данный вариант условий (41) выходит за рамки настоящего исследования.

Отметим, что граничный вариант отсутствия упругого элемента  $k_1 = 0$ ,  $b_1 > 0$  соответствует условию большого трения:

$$\begin{cases} \omega^2 = 0; \\ \sigma_1(\alpha) = -\frac{b_1}{m_1}; \\ \sigma_2(\alpha) = 0. \end{cases} \quad (43)$$

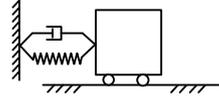
Вместе с тем, граничный вариант отсутствия диссипативного элемента  $k_1 > 0$ ,  $b_1 = 0$  приводит к выражению:

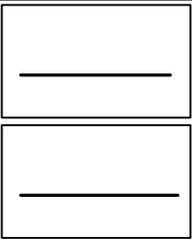
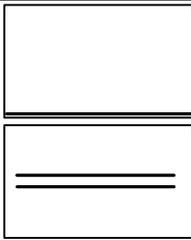
$$\begin{cases} \omega^2(\alpha) = \frac{k_1}{m_1}; \\ \sigma(\alpha) = 0. \end{cases} \quad (44)$$

Представленная система функций может быть интерпретирована как наличие гармонических форм движения.

Сводная информация о характеристиках системы представлена в таблице 1.

**Таблица 1.** Сводная информация

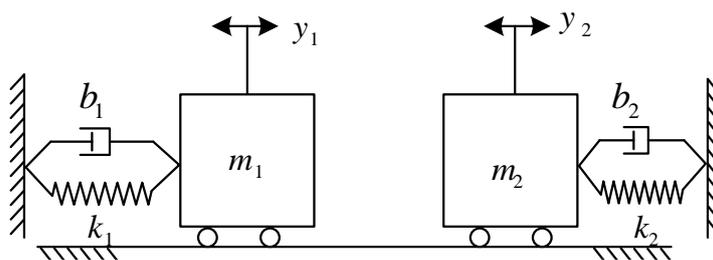
Вариант	Механическая колебательная система с одной степенью свободы
Параметры	$b_0 = 0, b_2 = 0, k_0 = 0, k_2 = 0,$
Схема	
Частотная функция	$\omega^2(\alpha) = \frac{k_1}{m_1} - \left(\frac{1}{2} \frac{b_1}{m_1}\right)^2$
Функция демпфирования	$\sigma(\alpha) = -\frac{1}{2} \frac{b_1}{m_1}$
Двухкомпонентная функция демпфирования	$\begin{cases} \sigma_1(\alpha) = -\frac{b_1}{2m_1} - \sqrt{\left(\frac{b_1}{2m_1}\right)^2 - \frac{k_1}{m_1}}; \\ \sigma_2(\alpha) = -\frac{b_1}{2m_1} + \sqrt{\left(\frac{b_1}{2m_1}\right)^2 - \frac{k_1}{m_1}}, \end{cases}$
Параметризирующее неравенство	$\left(\frac{b_1}{2m_1}\right)^2 < \frac{k_1}{m_1}$

Характерные формы сочетания частотной функции и функции демпфирования		
	(1.1)	(1.2)

В строке «Характерные формы сочетания частотной функции и функции демпфирования» таблицы 1 приведены парные пиктограммы, образованные верхним прямоугольником, отображающим схематически неотрицательную частотную функцию, и нижним прямоугольником – неположительные функции демпфирования, имеющие либо одну, либо две компоненты в зависимости от условий малости сил вязкого трения. В частности, пиктограмма (1.1) отображает тот факт, что частотная функция принимает постоянное положительное значение, а функция демпфирования имеет одну компоненту, принимающую постоянное отрицательное значение. Вместе с тем, пиктограмма (1.2) отображает тот факт, что существует набор параметров, для которых частотная функция принимает нулевое значение, а функция демпфирования имеет две компоненты, принимающие отрицательные постоянные значения.

**V. Механическая система с двумя несвязанными парциальными системами.**

Рассматривается механическая колебательная система (рис. 1) с параметрами  $b_0 = 0, k_0 = 0$ . Расчетная схема представлена на рисунке 3. Рассматриваемая система представляет собой также вырожденный вариант системы с двумя степенями свободы, главной особенностью которого является отсутствие упруго-диссипативной связи между массо-инерционными элементами. Вместе с тем, система (рис.1) представляет интерес в плане базы для сравнения на начальном этапе оценки разнообразия видов частотных функций и функций демпфирования.



**Рис. 3.** Механическая колебательная система с двумя несвязанными парциальными системами  $b_0 = 0, k_0 = 0$

Система дифференциальных уравнений принимает вид:

$$\begin{cases} m_1 \ddot{y}_1 + b_1 \dot{y}_1 + k_1 y_1 = 0; \\ m_2 \ddot{y}_2 + b_2 \dot{y}_2 + k_2 y_2 = 0. \end{cases} \quad (45)$$

Компоненты  $A_\alpha, B_\alpha, C_\alpha$  имеют вид четных функций:

$$A_\alpha = m_1 + m_2 \alpha^2, \quad (46)$$

$$B_\alpha = b_2 \alpha^2 + b_1, \quad (47)$$

$$C_\alpha = k_2 \alpha^2 + k_1. \quad (48)$$

На основе представленных компонент могут быть построены частотная функция, функция демпфирования и условия на величину диссипации.

*Учет диссипации.* Граничные условия могут быть представлены в виде:

$$\frac{1}{4} \cdot \left( \frac{b_2 \alpha^2 + b_1}{m_1 + m_2 \alpha^2} \right)^2 = \frac{k_2 \alpha^2 + k_1}{m_1 + m_2 \alpha^2}. \quad (49)$$

При малых силах вязкого трения функции  $\omega^2(\alpha)$  и  $\sigma(\alpha)$  принимают вид:

$$\begin{cases} \omega^2(\alpha) = \frac{k_2 \alpha^2 + k_1}{m_1 + m_2 \alpha^2} - \left( \frac{1}{2} \frac{b_2 \alpha^2 + b_1}{m_1 + m_2 \alpha^2} \right)^2 \\ \sigma(\alpha) = -\frac{1}{2} \frac{b_2 \alpha^2 + b_1}{m_1 + m_2 \alpha^2} \end{cases}. \quad (50)$$

При больших силах трения функции  $\omega^2(\alpha)$ ,  $\sigma(\alpha)$  имеют значения:

$$\begin{cases} \omega^2(\alpha) = 0; \\ \sigma_1(\alpha) = -\frac{1}{2} \frac{b_2 \alpha^2 + b_1}{m_1 + m_2 \alpha^2} - \sqrt{\left( \frac{1}{2} \frac{b_2 \alpha^2 + b_1}{m_1 + m_2 \alpha^2} \right)^2 - \frac{k_2 \alpha^2 + k_1}{m_1 + m_2 \alpha^2}}; \\ \sigma_2(\alpha) = -\frac{1}{2} \frac{b_2 \alpha^2 + b_1}{m_1 + m_2 \alpha^2} + \sqrt{\left( \frac{1}{2} \frac{b_2 \alpha^2 + b_1}{m_1 + m_2 \alpha^2} \right)^2 - \frac{k_2 \alpha^2 + k_1}{m_1 + m_2 \alpha^2}}. \end{cases} \quad (51)$$

*Экстремальные и предельные свойства частотной функции и функции демпфирования.* Определим производные частотной функции и функции демпфирования:

$$\begin{cases} \omega^2(\alpha)' = \frac{2\alpha}{(m_1 + m_2 \alpha^2)^2} [(m_2 k_1 - m_1 k_2) + \frac{b_2 \alpha^2 + b_1}{m_1 + m_2 \alpha^2} \cdot (b_2 m_1 - b_1 m_2)]; \\ \sigma(\alpha)' = -\frac{\alpha(b_2 m_1 - b_1 m_2)}{(m_1 + m_2 \alpha^2)^2}. \end{cases} \quad (52)$$

В точке  $\alpha = 0$  производные  $\omega^2(\alpha)'$ ,  $\sigma(\alpha)'$  равны нулю. При условии, что функции монотонны на интервале  $(0, \infty)$  можно полагать, что экстремальные значения функции достигают при  $\alpha = 0$  и  $\alpha \rightarrow \infty$ . Если  $\alpha = 0$ , то:

$$\begin{cases} \omega^2(0) = \frac{k_1}{m_1} - \left( \frac{1}{2} \frac{b_1}{m_1} \right)^2; \\ \sigma(0) = -\frac{1}{2} \frac{b_1}{m_1}. \end{cases} \quad (53)$$

Если  $\alpha \rightarrow \infty$ , то

$$\begin{cases} \omega^2(\alpha) \rightarrow \frac{k_2}{m_2} - \left( \frac{1}{2} \frac{b_2}{m_2} \right)^2; \\ \sigma(\alpha) \rightarrow -\frac{1}{2} \frac{b_2}{m_2}. \end{cases} \quad (54)$$

*Характерные формы частотных функций и функций демпфирования.* Рассмотрим экстремальные значения частотной функции  $\omega^2(\alpha)$  и функции демпфирования  $\sigma(\alpha)$  при больших силах трения. Полагаем, что для любого  $\alpha$  выполнено:

$$\frac{1}{4} \cdot \left( \frac{b_2 \alpha^2 + b_1}{m_1 + m_2 \alpha^2} \right)^2 > \frac{k_2 \alpha^2 + k_1}{m_1 + m_2 \alpha^2}. \quad (55)$$

Рассмотрим семейство механических систем с параметрами  $k_1 = 4$ ,  $k_2 = 6$ ,  $m_1 = 1$ ,  $m_2 = 5$  в предположении, что выполнено соотношение  $b_2 = \lambda b_1$ ,  $\lambda = 0.35$ . В этом случае условие малости сил трения принимает вид:

$$\gamma_0 < M(\alpha), \quad (56)$$

где параметр  $\gamma_0 = \frac{b_2^2}{4k_2}$  отражает упруго-диссипативные свойства системы,

$$M(\alpha) = \frac{(m_1 + m_2 \alpha^2)(\alpha^2 + \frac{k_1}{k_2})}{(\alpha^2 + \lambda)^2}$$

параметризирующая функция коэффициента формы  $\alpha$ .

Функция  $M(\alpha)$  представлена на рисунке 4(а). Значение параметра  $b_1$  определяет области, в которых выполняется условие малости сил трения. Критические значения  $b_1$  определяются экстремальными значениями функции  $M(\alpha)$  и предельными значениями при  $\alpha \rightarrow \infty$ . Положим, что  $M_0$  – значение локального минимума в точке  $\alpha_0 = 0$ ,  $M_1$  – величина локального максимума в точке  $\alpha_1, \alpha_2$ .

Можно указать четыре варианта относительного положения коэффициента  $\gamma_0$

**1.** Пусть  $\gamma_0 < m_2$ . Пусть  $b_2 = 1$ ,  $\gamma_0 \approx 0.42$ . В этом случае для любого коэффициента формы  $\alpha \in (-\infty, \infty)$  выполнены условия малости сил трения. На рисунках 4(б) и 4(г) представлены частотная функция  $\omega^2(\alpha)$  и функция демпфирования  $\sigma(\alpha)$ . Корни уравнения, представляющего собой равенство нулю определителя:

$$|Ap^2 + Bp + C| = 0 \quad (57)$$

составляют  $p_i = \omega_i + j\sigma_i$ ,  $i = 1..4$ , где вещественные части  $\sigma_1 \approx -0.1$ ,  $\sigma_2 \approx -0.1$ ,  $\sigma_3 \approx -0.18$ ,  $\sigma_4 \approx -0.18$ , комплексные части  $\omega_1 \approx 1.09$ ,  $\omega_2 \approx -1.09$ ,  $\omega_3 \approx 1.99$ ,  $\omega_4 \approx -1.99$ . На рисунке 4(б) частотная функция достигает экстремальных значений, равных квадратам частот  $\omega_3^2 \approx 3.97$  и  $\omega_1^2 \approx 1.19$  в точках  $\alpha_1^* = 0$  и  $\alpha_2^* = \infty$  соответственно. На рисунке 4(г) функция демпфирования достигает экстремальные значения, которые составляют  $\sigma_3 \approx -0.18$  и  $\sigma_1 \approx -0.1$  в точках  $\alpha_1^* = 0$  и  $\alpha_2^* = \infty$  соответственно.

**2.** Пусть  $m_2 < \gamma_0 < M_0$ . Рассматриваются параметры  $b_2 = 11.2$ ,  $\gamma_0 \approx 5.23$ . Условие реализации малости сил трения нарушаются на интервале  $(-\infty, \alpha_1) \cup (\alpha_2, \infty)$ , где  $\alpha_1 \approx -1.74$ ,  $\alpha_2 \approx 1.74$  - корни уравнения  $\gamma_0 = M(\alpha)$ . Для рассматриваемых параметров корни уравнения (51) имеют соответствующие вещественные части  $\sigma_1 \approx -0.89$ ,  $\sigma_2 \approx -1.35$ ,  $\sigma_3 \approx -1.96$ ,

$\sigma_4 \approx -1.96$  и мнимые части  $\omega_1 \approx 0$ ,  $\omega_2 \approx 0$ ,  $\omega_3 \approx 0.40$ ,  $\omega_4 \approx -0.40$ . Соответствующие частотная функция и функция демпфирования приведены на рисунках 4(с) и 4(г). На интервале  $(-\infty, \alpha_1) \cup (\alpha_2, \infty)$  частотная функция равна нулю, а на интервале  $(\alpha_1, \alpha_2)$  положительна и достигает локальный минимум в точке  $\alpha_2^* = 0$  равный  $\omega_3^2 \approx 0.16$ . В свою очередь, функция демпфирования в интервале  $(-\infty, \alpha_1) \cup (\alpha_2, \infty)$  двузначна и достигает экстремальных значений при  $\alpha \rightarrow \infty$ , которые составляют  $\sigma_1 \approx -0.89$ ,  $\sigma_2 \approx -1.35$ . В области  $(-\infty, \alpha_1) \cup (\alpha_2, \infty)$  функция затухания однозначна и имеет один локальный экстремум  $\sigma_4 \approx -1.96$  в точке  $\alpha_2^* = 0$ .

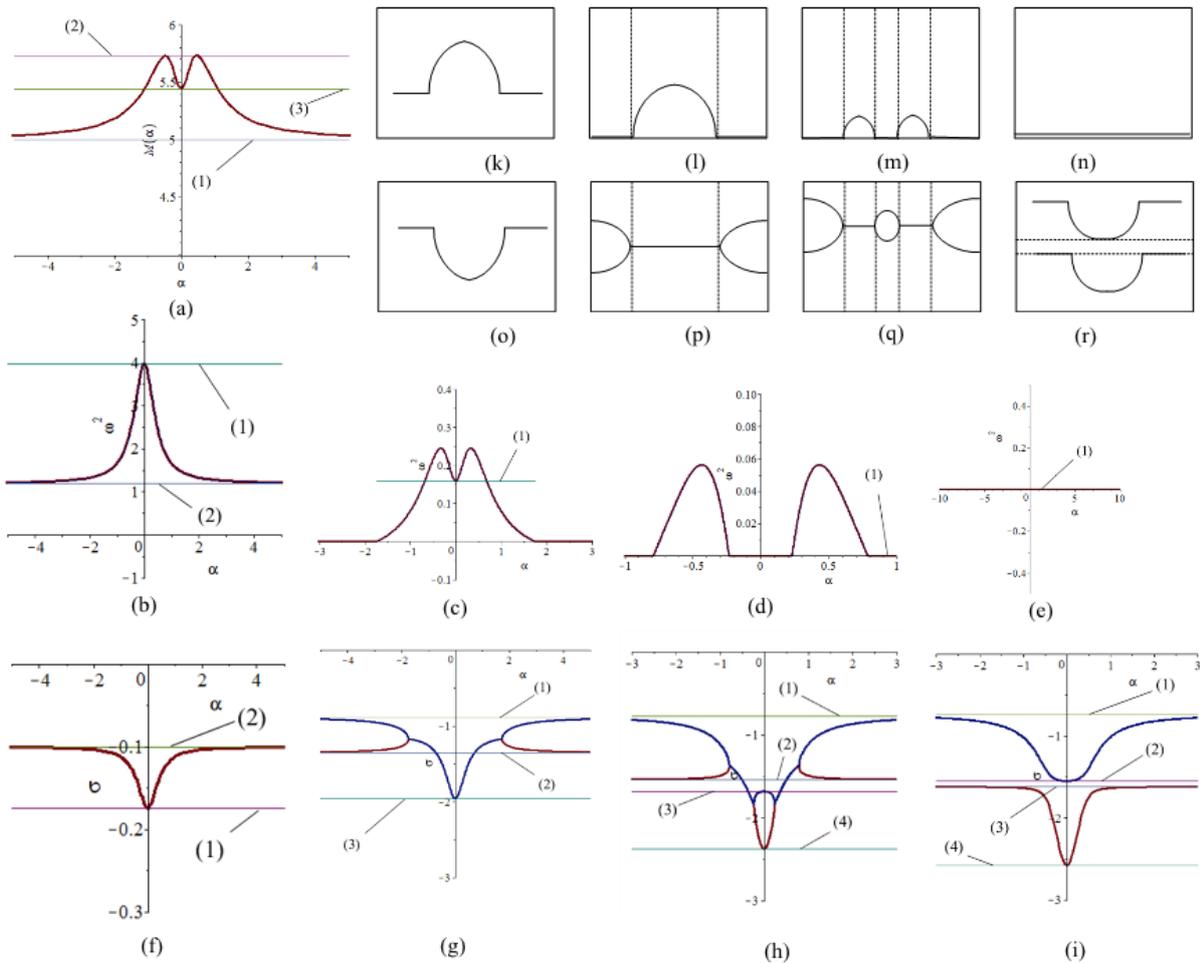
**3.** Пусть  $M_0 < \gamma_0 < M_1$ ,  $b_2 = 11.6$ ,  $\gamma_0 \approx 5.61$  Условие реализации малости колебаний нарушаются на множестве  $(-\infty, \alpha_1) \cup (\alpha_2, \alpha_3) \cup (\alpha_4, \infty)$ ,

где  $\alpha_1 \approx -0.79$ ,  $\alpha_2 \approx -0.23$ ,  $\alpha_3 \approx 0.23$ ,  $\alpha_4 \approx 0.79$  – корни уравнения  $\gamma_0 = M(\alpha)$ .

Для рассматриваемых параметров корни уравнения (61) имеют соответствующие вещественные  $\sigma_1 \approx -0.78$ ,  $\sigma_2 \approx -1.54$ ,  $\sigma_3 \approx -1.68$ ,  $\sigma_4 \approx -2.38$  и мнимые части  $\omega_1 \approx 0$ ,  $\omega_2 \approx 0$ ,  $\omega_3 \approx 0$ ,  $\omega_4 \approx 0$ . Соответствующие частотная функция и функция демпфирования приведены на рисунках 4(d) и 4(h). На множестве  $(-\infty, \alpha_1) \cup (\alpha_2, \alpha_3) \cup (\alpha_4, \infty)$  частотная функция равна нулю, а на множестве  $(\alpha_1, \alpha_2) \cup (\alpha_3, \alpha_4)$  положительна, имеет экстремальные значения, которые не совпадают с нулевой собственной частотой  $\omega_1 \approx 0$ . В свою очередь, функция демпфирования в интервале  $(-\infty, \alpha_1) \cup (\alpha_2, \alpha_3) \cup (\alpha_4, \infty)$  двузначна и достигает экстремальных значений  $\sigma_3 \approx -1.68$ ,  $\sigma_4 \approx -2.38$  в точке  $\alpha_1^* = 0$  и при  $\alpha \rightarrow \infty$  приближается к  $\sigma_1 \approx -0.78$ ,  $\sigma_2 \approx -1.54$ . В области  $(\alpha_1, \alpha_2) \cup (\alpha_3, \alpha_4)$  функция затухания однозначна.

**4.** Пусть  $M_1 < \gamma_0$ ,  $b_2 = 11.8$ ,  $\gamma_0 \approx 5.8$  Условие реализации малости сил трения нарушаются на всей вещественной оси. Для рассматриваемых параметров корни уравнения (51) имеют вещественные  $\sigma_1 \approx -0.58$ ,  $\sigma_2 \approx -2.07$ ,  $\sigma_3 \approx -1.15$ ,  $\sigma_4 \approx -3.49$  и мнимые части  $\omega_1 \approx 0$ ,  $\omega_2 \approx 0$ ,  $\omega_3 \approx 0$ ,  $\omega_4 \approx 0$ . Частотная функция и функция демпфирования приведены на рисунках 4(е) и е(i). Частотная функция равна нулю. Функция демпфирования в интервале двузначна на всей вещественной оси и достигает экстремальные значения  $\sigma_4 \approx -3.49$ ,  $\sigma_2 \approx -2.07$  в точке  $\alpha_1^* = 0$  и при  $\alpha \rightarrow \infty$  приближается к  $\sigma_1 \approx -0.58$ ,  $\sigma_3 \approx -1.15$ .

На рисунка 4(k)-4(r) приведены парные пиктограммы, образованные верхним прямоугольником (рис. 4(k)-4(n)), отображающим схематически неотрицательную частотную функцию, и нижним прямоугольником (рис. 4(o)-4(r)), отображающим неположительные функции демпфирования, имеющие либо одну, либо две компоненты в зависимости от выполнения условий малости сил вязкого трения. Приведенные парные пиктограммы отображают существенные свойства частотной функции и функции демпфирования, представленные на рисунках 4(b)-4(i). К существенным свойствам следует отнести равенство нулю частотной функции на интервале и наличие двух или одной компоненты у функции демпфирования.



**Рис. 4.** Особенности формирования частотной функции и функции демпфирования: (а) – график параметризующей функции семейства механических систем; (б)-(е) – частотные функции; (ф)-(и) – функции демпфирования; (к)-(н) – пиктограммы соответствующих частотных функций (б)-(е); (о)-(р) – пиктограммы соответствующих функций демпфирования (ф)-(и)

Таким образом, для рассмотренных механических систем можно предположить, что экстремальные значения совокупности частотной функции и функции демпфирования отражают такие динамические параметры, как собственные характеристики системы.

**Заключение.** По результатам представленных исследований возможны следующие выводы.

1. Для механической колебательной системы с двумя степенями свободы разработан метод построения частотной функции и функции демпфирования, отражающих динамические особенности свободных движений с учетом сил трения в зависимости от коэффициента форм связности.

2. Показано, что определенная совокупность частотной функции и функции демпфирования для механической колебательной с двумя степенями системы с учетом вязкого трения может быть представлена, в общем случае, тремя вещественными функциями, значения которых определяются уровнем диссипации энергии или величиной сил вязкого трения для фиксированного значения коэффициента форм.

3. Показано, что для условий малых сил вязкого трения значения частотной функции принимают положительные значения, а функция демпфирования имеет одну отрицательную компоненту. Если выполнены условия больших сил вязкого трения, то частотная функция рассматривается, как нулевая функция, а функция демпфирования имеет две отрицательные компоненты.

4. Показано, что совокупность экстремальных значений частотной функции и функции демпфирования отображает собственные характеристики упруго-диссипативной механической системы.

5. Предложенный матричный метод построения частотно-демпфирующей функции для системы с двумя степенями свободы может быть распространен на механические колебательные системы, рассматриваемые в различных системах координат.

#### СПИСОК ЛИТЕРАТУРЫ

1. Елисеев А. В. Динамика машин. Системные представления, структурные схемы и связи элементов: монография / А. В. Елисеев, Н. К. Кузнецов, А. О. Московских. // Москва: Инновационное машиностроение. 2019. 381 с.
2. Елисеев С. В. Прикладная теория колебаний в задачах динамики линейных механических систем / Елисеев С. В., Артюнин А. И.- Наука, Новосибирск. 2016. с. 459
3. Елисеев С.В. О возможностях использования дополнительных связей инерционного типа в задачах динамики технических систем / Елисеев С.В., Кузнецов Н.К., Большаков Р.С., Нгуен Д.Х. // Вестник Иркутского государственного технического университета. 2016. № 5 (112). С. 19-36.
4. Елисеев С.В. Определение частот собственных колебаний механических колебательных систем: особенности использования частотной энергетической функции. Часть I. / Елисеев С.В., Большаков Р.С., Нгуен Д.Х., Вьюнг К.Ч. // Вестник Иркутского государственного технического университета. 2016. № 6 (113). С. 26-33.
5. Елисеев С.В. Определение частот собственных колебаний механических колебательных систем: особенности использования частотной энергетической функции. Часть II. / Елисеев С.В., Большаков Р.С., Нгуен Д.Х., Вьюнг К.Ч. // Вестник Иркутского государственного технического университета. 2016. № 7 (114). С. 10-23.
6. Елисеев С.В. Прикладной системный анализ и структурное математическое моделирование (динамика транспортных и технологических машин: связность движений, вибрационные взаимодействия, рычажные связи): монография / С.В. Елисеев; отв. ред. А.И. Артюнин. – Иркутск: ИрГУПС. 2018. 692 с.
7. Стретт Дж. В. Теория звука. М.: ГИТТЛ. 1955. Т.1. 503 с.
8. Хоменко А.П. Развитие энергетического метода определения частот свободных колебаний механических систем / Хоменко А.П., Елисеев С.В. // Современные технологии. Системный анализ. Моделирование. 2016. № 1 (49). С. 8-19.
9. Clarence W. de Silva. Vibration. Fundamentals and Practice. Boca Raton, London, New York, Washington, D.C.: CRC Press. 2000. 957 p.
10. Eliseev S.V., Eliseev A.V. Theory of Oscillations. Structural Mathematical Modeling in Problems of Dynamics of Technical Objects. Series: Studies in Systems, Decision and Control, Vol.252, Springer International Publishing. Cham. 2020. 521 p.

11. Karnovsky I.A., Lebed E. Theory of Vibration Protection, Springer International Publishing, Switzerland. 2016. 708 p.

УДК 519.71+ 681.5+ 62.752

## SYSTEM REPRESENTATIONS OF DYNAMIC PROCESSES IN MECHANICAL OSCILLATORY SYSTEMS WITH SPECIAL FEATURES

**Eliseev Andrey Vladimirovich**

Candidate of Technical Sciences, Associate Professor of the Department of Mathematics,

e-mail: eavsh@ya.ru,

Irkutsk State Transport University

664074, Irkutsk, 15 Chernyshevsky st.

**Abstract.** Approaches to the formation of a methodological basis for the system analysis of oscillatory structures on the example of elastic-dissipative mechanical systems with two degrees of freedom are considered. The aim of the study is to develop a method for evaluating the properties of mechanical oscillatory systems with account for viscous friction forces based on frequency functions and damping functions depending on the coefficients of connectivity forms. To build mathematical models, we use Lagrange formalism, matrix methods, and elements of the theory of functions of a complex variable. The concepts of the frequency function and the damping function are presented, which reflect the features of the ratio of the potential and kinetic energy of the system, taking into account the viscous friction forces represented by the dissipative function. A method is developed for constructing the frequency function and damping function that reflect the dynamic features of free movements, taking into account the friction forces depending on the coefficient of connectivity forms. Based on the obtained General analytical expressions of frequency functions and damping functions, special variants of mechanical oscillatory systems that are of interest at the initial stage of research are analyzed. The developed method for constructing the frequency function and the damping function can be used to display dynamic forms of connectivity of movements of mechanical oscillatory systems. The proposed method for constructing the frequency function and the damping function can be extended to mechanical oscillatory systems considered in different coordinate systems.

**Keywords:** mechanical system, dynamic connections, frequency function, damping function, connectivity of movement, extreme properties, oscillation, viscous friction.

### References

1. Eliseev A.V., Kuznecov N. K., Moskovskih A. O. Dinamika mashin. Sistemnye predstavlenija, strukturnye shemy i svjazi jelementov [Dynamics of machines. System representations, structural schemes and connections of elements]. Moskva: Innovacionnoe mashinostroenie=(Innovative engineering). 2019. 381 p.

2. Eliseev S. V., Artjunin A. I. Prikladnaja teorija kolebanij v zadachah dinamiki linejnyh mehanicheskikh sistem [Applied theory of vibrations in problems of dynamics of linear mechanical systems]. Novosibirsk: Nauka = Science. 2016. 459 p.
3. Eliseev S.V., Kuznecov N.K., Bol'shakov R.S., Nguen D.H. O vozmozhnostjakh ispol'zovanija dopolnitel'nyh svjazej inercionnogo tipa v zadachah dinamiki tehničeskikh sistem [On the possibilities of using additional inertial type connections in problems of dynamics of technical systems] Vestnik Irkutskogo gosudarstvennogo tehničeskogo universiteta = Bulletin of the Irkutsk State Technical University. 2016. № 5 (112). Pp. 19-36.
4. Eliseev S.V., Bol'shakov R.S., Nguen D.H., Vyong K.Ch. Opredelenie chastot sobstvennyh kolebanij mehanicheskikh kolebatel'nyh sistem: osobennosti ispol'zovanija chastotnoj jenergetičeskoj funkcii. Chast' I. [Determination of natural oscillation frequencies of mechanical oscillatory systems: features of using the frequency energy function. Part I.] Vestnik Irkutskogo gosudarstvennogo tehničeskogo universiteta = Bulletin of the Irkutsk State Technical University. 2016. № 6 (113). Pp. 26-33.
5. Eliseev S.V., Bol'shakov R.S., Nguen D.H., Vyong K.Ch. Opredelenie chastot sobstvennyh kolebanij mehanicheskikh kolebatel'nyh sistem: osobennosti ispol'zovanija chastotnoj jenergetičeskoj funkcii. Chast' II [Determination of natural oscillation frequencies of mechanical oscillatory systems: features of using the frequency energy function. Part II]. Vestnik Irkutskogo gosudarstvennogo tehničeskogo universiteta = Bulletin of the Irkutsk State Technical University. 2016. № 7 (114). Pp. 10-23.
6. Eliseev S.V. Prikladnoj sistemnyj analiz i strukturnoe matematičeskoe modelirovanie (dinamika transportnyh i tehničeskikh mashin: svjaznost' dvizhenij, vibracionnye vzaimodejstvija, ryčazhnye svjazi) [Applied system analysis and structural mathematical modeling (dynamics of transport and technological machines: connectivity of movements, vibration interactions, lever connections)] Irkutsk: IrGUPS = Irkutsk State Transport University. 2018. 692 p.
7. Strett Dzh. V. Teorija zvuka [Theory of sound]. M.: GITTL = State publishing house of technical and theoretical literature. 1955. V. 1. 503 p.
8. Homenko A.P., Eliseev S.V. Razvitie jenergetičeskogo metoda opredelenija chastot svobodnyh kolebanij mehanicheskikh sistem [Development of the energy method for determining the frequencies of free vibrations of mechanical systems] Sovremennye tehnologii. Sistemnyj analiz. Modelirovanie = Modern technologies. System analysis. Modeling. 2016. № 1 (49). Pp. 8-19.
9. Clarence W. de Silva. Vibration. Fundamentals and Practice. Boca Raton, London, New York, Washington, D.C.: CRC Press. 2000. 957 p.
10. Eliseev S.V., Eliseev A.V. Theory of Oscillations. Structural Mathematical Modeling in Problems of Dynamics of Technical Objects. Series: Studies in Systems, Decision and Control. Vol.252. Springer International Publishing. Cham. 2020. 521 p.
11. Karnovsky I.A., Lebed E. Theory of Vibration Protection, Springer International Publishing, Switzerland. 2016. 708 p.

УДК 004.6::574.5

**ИНФОРМАЦИОННАЯ ИНФРАСТРУКТУРА ДЛЯ ПОДДЕРЖКИ  
ИССЛЕДОВАНИЙ МИКРОБИОМА БАЙКАЛА**

**Черкашин Евгений Александрович**

Институт динамики систем и теории управления  
Сибирского отделения Российской Академии наук  
email: [eugeneai@icc.ru](mailto:eugeneai@icc.ru),

**Шигаров Алексей Олегович**

Иркутский научный центр Сибирского отделения Российской Академии наук  
e-mail: [shigarov@icc.ru](mailto:shigarov@icc.ru),

664033, Иркутск, ул. Лермонтова, 134.

**Христюк Василий Владимирович**

Национальный исследовательский Иркутский государственный технический университет  
e-mail: [khr@icc.ru](mailto:khr@icc.ru)

664074, Иркутск, Лермонтова, 83.

**Аннотация.** Рассмотрена проблема построения исследовательской среды для обработки данных секвенирования нового поколения (NGS – Next Generation Sequencing). Среда включает облачное хранилище данных (DaaS) и вычислительные службы (SaaS и PaaS), а также службы визуализации и интеграции данных. Осуществляется интеграция технологий с открытым исходным кодом для поддержки MiSeq SOP (стандартная операционная процедура), которая позволяет специалистам в предметной области – биологам независимо от программистов самостоятельно обрабатывать данные. Для реализации интеграции конструируются формальные модели SOP, позволяющие автоматически порождать исходный код компонентов среды. Технология преобразования основана на принципах архитектуры, управляемой моделями (Model driven architecture), и логическом выводе структур производных моделей и модулей. Представлены текущие результаты и задачи на ближайшую перспективу.

**Ключевые слова:** секвенирование нового поколения, большие данные, архитектура, управляемая моделями, открытые связанные данные, планирование действий

**Цитирование:** Черкашин Е.А., Шигаров А., Христюк В. Информационная инфраструктура для поддержки исследований микробиома Байкала // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 108-123. DOI: 10.38028/ESI.2020.20.4.010

**Введение.** В последнее десятилетие, после изобретения методов секвенирования нового поколения (NGS) и внедрения их в практику исследования биологических систем, формируется новое направление молекулярной генетики, получившее название *метагеномики*. Его основной объект изучения выходит за рамки отдельных микроскопических культивируемых организмов и перемещается к их сообществам, *микробиомам*. Из образца извлекается ДНК, в результате чего накапливается информация, представляющая данные для всего образца микробиома. Этот метод позволяет описать значительное количество новых групп организмов

на всех таксономических уровнях. Подробный обзор существующих проблем и современных подходов к секвенированию представлен в [22].

Один из видов метагеномных исследований – анализ ампликонов, применяющийся для изучения микробиоты различных сред озера Байкал [4]. Для выполнения анализа и интерпретации результатов требуются значительные вычислительные ресурсы, а также навыки в области биоинформатики. Исследователь проектирует вычислительный процесс, комбинируя различные модули биоинформатического программного обеспечения, преобразования данных, анализа данных и визуализации. Для проведения исследований специалисты-предметники должны обладать навыками программирования сценариев командной оболочки операционной системы (Linux, Windows), запуска распределенных вычислительных сред, кластерных вычислительных системы, а также программирования на языках общего и специального назначения, например, Python и R.

Целью исследования является разработка математического и программного обеспечения процессов анализа результатов NGS. Перспективная задача авторов – разработать и внедрить методы и программное обеспечение для визуального представления вычислительного процесса анализа ампликонов, чтобы специалисты предметной области могли составлять вычислительные конвейеры, выполняющиеся в распределенных гетерогенных вычислительных ресурсах (облаках). Конечная цель – создание облачной инфраструктуры, построенной на моделях вычислительного процесса в виде набора операций, структур и функций вычислительных ресурсов, а также алгоритмов управления вычислительными ресурсами.

В [11] проблема использования облачных хранилищ и вычислений сформулирована как основная проблема NGS, так как экспоненциальный рост емкости хранилищ генетических данных происходит медленнее по сравнению с ростом объема данных, генерируемых NGS. Передача данных между хранилищами и компьютерными системами для их обработки приводит к исчерпанию емкости сети. Для обработки данных, в общем случае реконструкции всего генома, требуются терабайты оперативной памяти, а в случае использования кластерных вычислений – специальные высокопроизводительные параллельные алгоритмы. Выявлены два класса пользователей: *опытный пользователь (power user)*, который производит самостоятельно анализ генома, и *предметный исследователь с невысокой квалификацией пользователя в биоинформатике (casual user)*, осуществляющий обработку результатов NGS, интегрируя их с результатами других исследований.

Область исследований авторов в области информационных технологий (ИТ) связана с обработкой данных в рамках естественнонаучных исследований, которые характеризуются разнообразием задач, методов и мультидисциплинарных целей. Новые данные сравниваются со всеми данными, полученными в предыдущие годы. Увеличивается и количество научных задач. В рамках исследования предлагается построить облако PaaS и DaaS, состоящее из независимых сетевых SaaS-сервисов, адаптированных для стандартной операционной процедуры MiSeq (MiSeq SOP, стандартная процедура обработки данных), используемой в исследовании микробиома озера Байкал. PaaS позволит биологам самим анализировать и исследовать данные, DaaS позволит специалистам по биоинформатике работать с данными по запросу при разработке новых методов обработки данных, а сервис SaaS будет поддерживать отдельные операции для PaaS. К настоящему моменту уже существует программное обеспечение SaaS для обработки данных NGS, статья [11] содержит его подробный обзор.

**1. Стандартная операционная процедура MiSeq.** В качестве введения в проблемную область опишем пример реализации MiSeq SOP, выполненный при помощи прикладного пакета Mothur [15].

Процесс анализа данных NGS состоит из отдельных операций с генетическими данными, хранящихся в файлах. Для ознакомления с методикой авторами вручную проведена обработка лимнологических данных из исследования [18] в соответствии с процедурой, представленной на сайте Mothur. Исходные данные прочтения генов получены на секвенаторе GS FLX454.<sup>1</sup>

Все начинается с объединения левых и правых прочтений исходного fasta-файла, содержащего 44934 контига (смежные последовательности генов). Следующая операция – это обрезка праймеров олигонуклеотидов, т.е. частей последовательностей, идентифицирующих образцы для каждого контига. После операции обрезки, фильтрации нужно изучить сводную статистику по fasta-данным «склеенных и обрезанных» прочтений. Свод (summary) представлен в виде статистических критериев минимальной, максимальной, средней и четырехквартильной оценки. На основе полученной информации задаются значения параметров для следующих операций.

После обрезки олигонуклеотидов необходимо удалить последовательности, которые короче или длиннее некоторого среднего значения. Степень отклонения обычно определяется процентом прочтений, которые разрешено удалить. В нашем случае удалено 11052 (25%) коротких прочтений. Следующим шагом – распознавание уникальных последовательностей и подсчет их реплик (точных копий), что позволяет далее выполнять следующие шаги по одному разу для каждой уникальной последовательности, а это, в свою очередь, сокращает время вычислений. Получено 26744 уникальных последовательности, из них удаляются прочтения с характеристикой гомополимера более 8, т.к. более длинные последовательности одних и тех же пар-оснований в ряду гомополимера в природе не существуют.

Основным вычислительно сложным этапом анализа является сопоставление (выравнивание) прочтений с данными, хранящимися в базе данных референсных генов, например, SILVA [23]. Применение выравнивания к имеющимся 21170 последовательностям привело к изменению длины прочтений со средней 407 пар-оснований до 871. Пропуски символов заполняются пробелами. После обрезки «висящих» символов и фильтрации невыровненных прочтений, надо опять выбирать уникальные последовательности, в результате получилось 21052 уникальных прочтения. Для дополнительного сужения количества последовательностей применено «мягкое» сравнение (фильтрация), допускающее одну мутацию на 100 пар-оснований, количество уникальных прочтений уменьшилось до 14234.

Следующий шаг – обнаружение и удаление химер. Химерные последовательности представляют собой контиги, построенные из частичных считываний рРНК (справа налево, слева направо) разных видов микроорганизмов, но имеющих общие части. Этот этап основан на построении таксономии и отнесении прочтений к таксонам (классам). Очень редкие случаи считаются химерными и удаляются.

Следующий этап является обязательным - необходимо оценить количество отклоненных последовательностей после распознавания химер. Имея на входе 11119

---

<sup>1</sup>Исследование поддержано грантом Иркутского научного центра СО РАН № 4.2. [https://www.mothur.org/wiki/MiSeq\\_SOP](https://www.mothur.org/wiki/MiSeq_SOP)

уникальных последовательностей, удалено химер 1.3% от общего количества, что считается разумным количеством. В целом, осуществлено сокращение объема обрабатываемой информации до 25% от исходного объема 44939 прочтений.

После этого этапа выполняется одна заключительная фильтрация для удаления митохондриальных и хлоропластных рРНК. Для этого необходимо классифицировать последовательности. Классификация проводится при помощи предобученной байесовской сети. В нашем примере удалено 1828 последовательностей, и уже окончательный набор содержал 9291 уникальное прочтение.

**1.1. Анализ микробного сообщества.** Для обработанных наборами последовательностей (групп, образцов) выполняются, в основном, четыре вида анализов. Последовательности организуются в операционные таксономические единицы (OTU), которые можно рассматривать как абстракцию понятия «вид». OTU содержит последовательности, имеющих заранее определенный уровень мутации, обычно не превышающий 3%. По нашим данным получено около тысячи OTU.

Первый вид анализа микробного сообщества это определение численности распознанных бактерий по таксономическим уровням в каждой OTU. Второй филогенетический анализ, направленный на определение расстояний между OTU, выражаемый в количестве мутаций. Третий, называемый «Альфа-разнообразие», измерение сходства групп на основе OTU или индивидуальных последовательностей, а также сходства OTU на основе состава групп. В результате получают тепловые карты, пример которых изображен на рисунке 1.

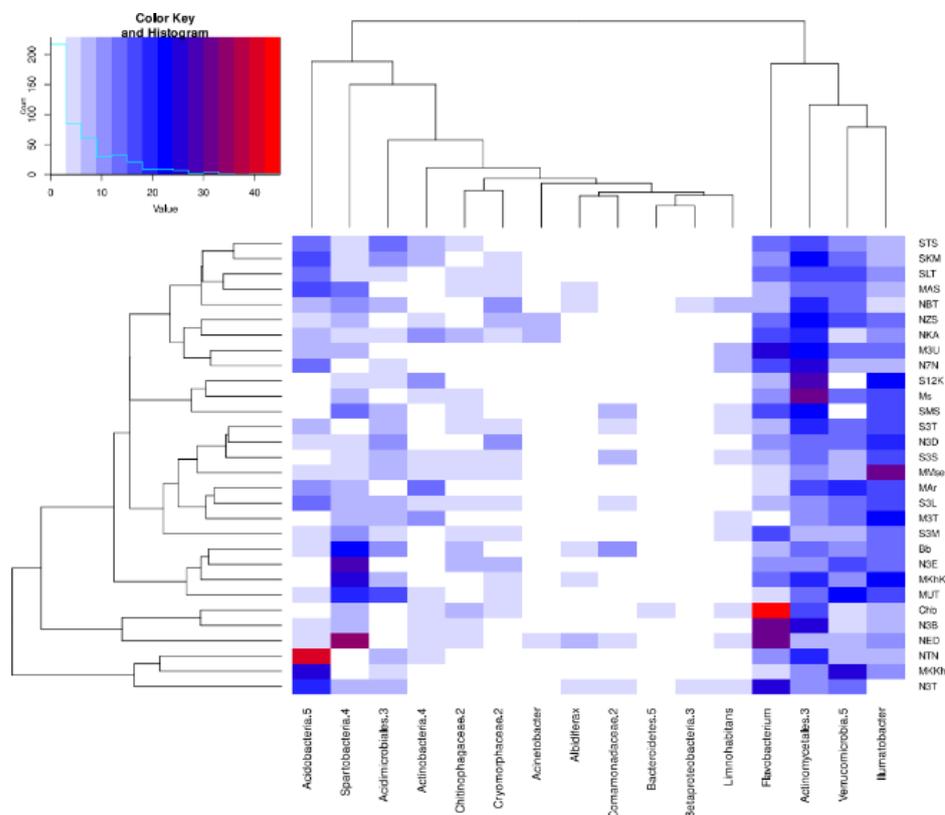


Рис.1. Пример тепловой карты

Четвертый анализ называется «Бета-разнообразие» и основывается на методе главных компонент и его аналогах, в результате строятся диаграммы групп OTU по сходству в пространстве осей главных компонент. Количество осей определяется последовательно, добавляя одну за другой и анализируя изменение R- и stress-критериев.

**1.2. Особенности ручной реализации технологии.** После применения технологии вручную выявился ряд проблем. Структура параметров и вывод команды Mothur достаточно сложные: необходимо отслеживать изменение имен файлов от команды к команде. После каждого выполнения команда Mothur добавляет суффиксы к именам входных файлов. Именование файлов зависит от входных параметров команды, например, от метода обработки данных. В частности, после применения команды align.seq к файлу с именем НХН779К01.shhh.trim.good.unique.fasta получаем НХН779К01.shhh.trim.good.unique.align, НХН779К01 ... trim.good.unique.align.report, и НХН779К01 ... im.good.unique.flip.acnos<sup>2</sup>.

Повторное применение операции к данным приводит к повторяющимся суффиксам. В Mothur включены инструменты отслеживания и передачи правильных имен файлов между командами, который поддерживает стандартные процедуры, реализуемые в виде сценариев. Вернувшись к предыдущему шагу для уточнения коэффициента, исследователь уже должен вручную вводить исходные имена файлов. После получения результатов в виде таблиц и диаграмм биологи обычно повторяют некоторые этапы фильтрации, исключая дополнительные OTU, например, которые похожи на митохондрии и хлоропласты. Иногда пользователи хотят заменить команду на аналогичную из другого пакета, например QUIIME2 или Usearch, где реализованы другие алгоритмы для некоторой процедуры. В этом случае необходимо преобразовывать данные из форматов Mothur в формат QUIIME2, Usearch и обратно.

Другое возможное, но менее частое отклонение от методики – это использование ранее обработанных OTU из ранних исследований, например, из образцов прошлых лет, отобранных на тех же локациях в рамках экологического мониторинга. Здесь требуется реализация процедуры для объединения содержимого OTU из разных исследований, т. е. программирование новых сценариев.

Визуализация результатов частично представлена в наборе команд Mothur, можно создавать векторные изображения SVG, но вид изображений, как правило, не настраивается. Приходится использовать внешнее программное обеспечение, такое, как R, для построения диаграмм желаемого качества. Наш опыт показывает, что, несмотря на время, потраченное на изучение методов построения диаграмм, большая часть времени уходит на преобразование и фильтрацию входных данных и уточнение параметров команд, генерирующих диаграммы.

**1.3. Исследования-аналоги.** Основная деятельность НИОКР в области NGS разделена на три основных направления: а) разработка новых эффективных алгоритмов, реализующих операции обработки данных и построения диаграмм, б) организация стандартизированных конвейеров облачных вычислений на кластерах класса HPC (High performance computing), в) представление конвейеров в виде рабочих процессов, а также пользовательских интерфейсов для поддержки интерактивной обработки и оценки данных (моделирование вычислительного процесса).

---

<sup>2</sup> Имена файлов в начале усечены, чтобы соответствовать размерам текста.

Рассмотрению свойств вычислительных сред, как инструментов реализации методов бионформатики, посвящена работа [8], где языки программирования Go, C++ и Java оценены с точки зрения простоты реализации алгоритмов NGS, потребления памяти и общей производительности вычислений; в результате выбран Go. В [9] разработан конвейер NGS для анализа ДНК вирусов в организме человека. Этот анализ позволил медицинским инженерам сосредоточить внимание на разработке вакцины. В статье [10] рассматривается реализация эвристического алгоритма для построения структур данных для упорядочения, перемещения и ориентации контигов с использованием дополнительной информации, что на следующих этапах позволяет создавать более длинные последовательности. В [21] рассматривается разработка пользовательского интерфейса для визуализации вычислительного процесса, управляемого данными. Результаты находят применение в системе поддержки принятия клинических решений. Пользовательский интерфейс предоставляет врачу возможность принятия решений и выдачи пояснений для пациентов. Пояснение представляется в виде списка структур различного типа (медицинские записи, результаты NGS, аутопсии ткани и т. д.), представленных в виде портлетов HTML5. Подробный обзор контроля качества, обнаружения и исправления ошибок при обработке данных NGS представлен в [1].

Обзор методов НРС начнем с применения технологии BOINC для выравнивания последовательностей, представленной в [25], где отмасштабирован алгоритм Novoalign. В справочнике [11] рассмотрена проблемная область, описаны существующие подходы и уже реализованные методики, но нет упоминаются действительно реализованных сред облачных вычислений. В статье [16] содержится превосходный обзор текущих достижений в области NGS и смежных областях. Авторы сомневаются в возможности организации лабораторного НРС-центра на базе кластерных вычислений пользователями программного обеспечения NGS и предлагают заняться облачными вычислениями IaaS. В статье содержится также обзор существующих коммерческих платформ и платформ с открытым исходным кодом, позволяющих создавать конвейеры вычислительных процессов. Коммерческое программное обеспечение, в основном, реализует предопределенные конвейеры и является негибким, в то время как программное обеспечение с открытым исходным кодом позволяет реализовывать как стандартизированные конвейеры, так и предоставлять модули для отдельных операций и реализации облачных сервисов, т. е. инструментарий разработчика.

В [26] рассматривается программа Rainbow, поддерживающая облачную обработку данных NGS. По сути, Rainbow – это сценарий Perl, реализующий операцию map (разделение) над входными данными и reduce (соединение) для получения выходных агрегированных данных; разделенные части распределяются между узлами облачного сервиса Amazon EC2. Облачные узлы выполняют исключительно операцию выравнивания. В статье представлен также хороший обзор дистрибутивов виртуальных машин Linux для организации облачных сервисов и биоинформатических пакетов. Другой интересный обзор применения облачных вычислений представлен в [17]. Авторы обращают внимание на облачную технологию с открытым исходным кодом Open Stack и её инструментарий – Common Workflow Language (CWL) [2], используемый для представления вычислительного процесса в облаке. Подробный обзор технологий есть и в [3].

Существуют визуальные инструменты для генетического анализа, например, Galaxy [5], реализующие популярный подход «интерактивной веб-страницы», где данные импортируются и обрабатываются модулями. В Galaxy реализованы инструменты анализа существующих

скриптов (сценариев) и визуального представления модели потока данных (dataflow). Его основная цель – научить биологов проведению анализов данных NGS. Набор функций является расширяемым, что позволяет проводить специализированные исследования NGS. Проект находится в стадии активной разработки, авторы считают необходимым провести его интеграцию в разрабатываемую платформу. Другой полезный инструмент UGENE [24] – приложение с открытым исходным кодом, основанное на платформе QT5, также находится в активной разработке. UGENE визуализирует вычислительный процесс и генетические данные.

Основная критика существующих инструментов представлена в [19], где справедливо утверждается, что утилиты командной строки поддерживают больше функций и обладают большей гибкостью, чем визуальные инструменты. Авторы предлагают собственный визуальный инструмент VisPro, подключенный к облаку. Инструмент реализует гибкий подход к проектированию вычислений, предполагающий принципиальное участие пользователя в процессе построения, настройки и выполнения процесса анализа данных.

Подводя итоги этого краткого обзора, приходим к выводу, что для построения инфраструктуры анализа NGS есть хорошие открытые технологические наработки, и методы, используемые в исследованиях микробиома Байкала в Лимнологическом институте СО РАН, необходимо адаптировать к этим наработкам. Вышеупомянутые технологии позволяют действовать в рамках наших задач, решение которых требуют большей гибкости процесса вычислений и адаптации опыта пользователя-предметника.

**2. Подход к автоматизации MiSeq SOP.** Лимнологи выполняют действия как опытных биоинформатиков, так и пользователей-предметников [11], т. е. обрабатывают как исходные данные секвенирования, так и результаты обработки последовательностей, визуализируя, сравнивая и обобщая результаты. Высокопроизводительные вычисления (HPC) обычно основаны на двух популярных моделях программирования [11]: Map-Reduce и программировании сети задач. Первая подразумевает, что данные разделяемы на подмножества, обрабатываемые независимо, затем результаты параллельной обработки объединяются в агрегированный объект. Вторая модель, сеть задач, запускает модули на узлах кластера при готовности их входных данных, при этом модули также выполняются независимо.

Простые команды фильтрации Mothur запускаются параллельно, используя отдельные ядра процессора, но вычислительная сложность не кажется достаточно высокой, чтобы было рационально тратить время на их разделение на отдельные узлы кластера и последующее объединение. Основной причиной использования облачных сред здесь является объединение ресурсов оперативной памяти, если данные не помещаются в памяти одной рабочей станции. Некоторые алгоритмы фильтрации основаны на классификации, которая обрабатывает все генетические данные, используя случайные выборки (bagging). Переход соответствующих алгоритмов на SaaS потребует их замены на распределенную версию.

В целом, чтобы упростить создаваемую архитектуру облачных вычислений, на первом этапе НИОКР решено использовать модель выполнения очереди задач (dataflow), в которой вычислительные ресурсы выполняют отдельные задачи из сети модулей, представляющих вариант MiSeq SOP. В этой сети каждый узел является модулем пакета Mothur. Конкретный вычислительный процесс проектируется при помощи приложения Rapidminer studio, являющегося визуальным редактором сети dataflow-модулей. На рис. 2 представлена начальная часть процесса MiSeq SOP. Объединение с облачной инфраструктурой потребует передачи

данных между DaaS и SaaS, хранения объектов с метаданными, что также учитывается в нашей модели.

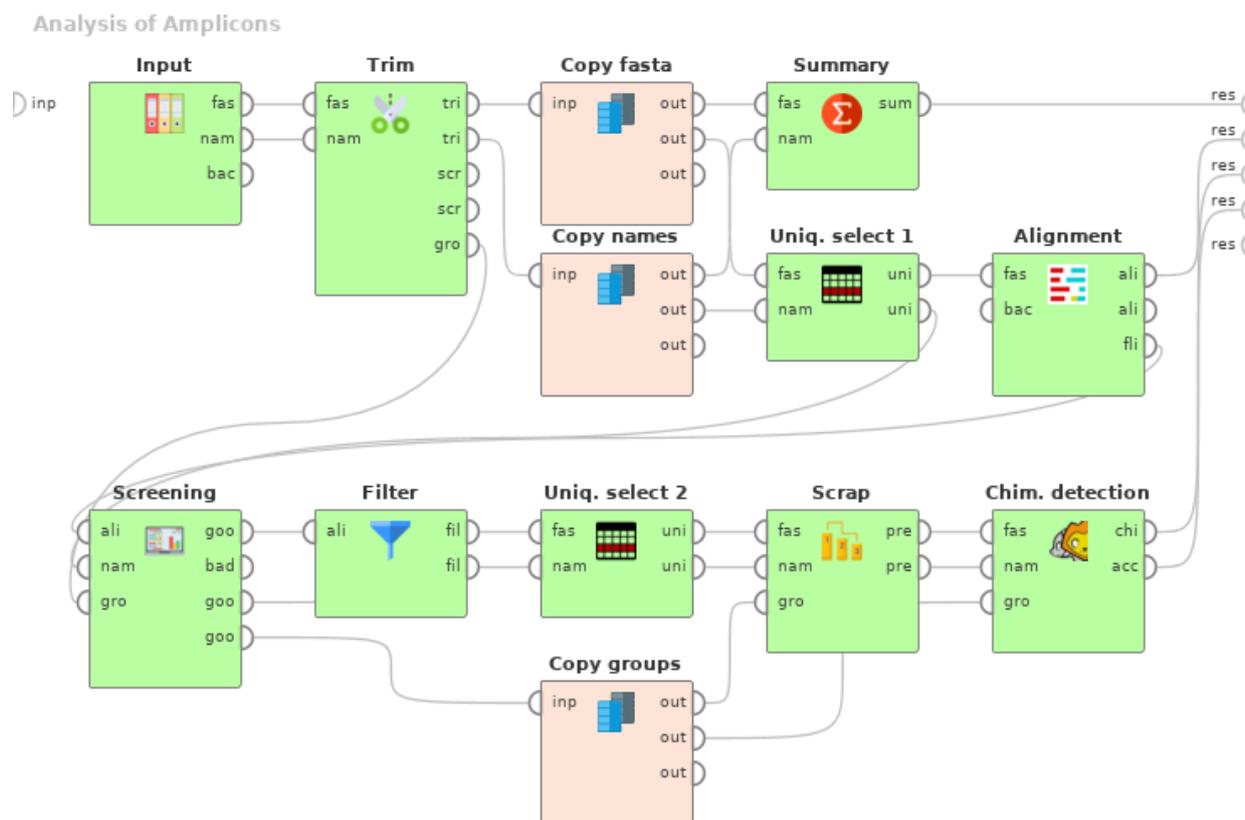


Рис. 2. Представление первых этапов MiSeq в виде модулей потока данных [7]

**3. Концепция реализации.** Обзор литературы показывает, что в настоящее время в активной разработке находятся два open-source проекта, связанных с автоматизацией анализа данных NGS: Galaxy и UGENE. Если взять Galaxy в качестве основного метода визуализации обработки данных, то необходимо реализовать модули Galaxy, адаптирующие команды Mothur, а также адаптировать визуализацию Galaxy к MiSeq SOP. С другой стороны, то же самое реализуемо и для UGENE, что позволит пользователю работать с более отзывчивым динамическим интерфейсом настольного приложения UGENE.

В [7] и [8] авторами предложена и частично реализована методика представления потока данных для всех команд Mothur с использованием Model Driven Architecture (MDA) для создания модулей Rapidminer studio. Согласно MDA, исходный код модулей генерируется на основе платформенно-зависимой модели (PSM), которая представляет разрабатываемое программное обеспечение в нотации, позволяющей прямую генерацию программного кода с помощью шаблонов и других императивных процедур. В нашем случае PSM представляет собой исходный код модулей потока данных на языке программирования Java.

Модель PSM порождается на основе платформенно-независимой модели (PIM), представляющей программное обеспечение на более абстрактном уровне, чем PSM. Она описывает отношения между сущностями, их объектные структуры, метаданные и т.д. Переход от PIM к PSM (трансформация модели) осуществляется при помощи логического вывода свойств PSM на основе фактов, представляющих PIM, и свойств платформы реализации, называемой моделью платформы (PM), в нашем случае это свойства среды программирования Java.

Некоторые свойства PIM, например, список полей объекта, создаются путем преобразования вычислительно независимой модели (CIM), еще более абстрактной модели, чем PIM. Модель CIM представляет программное обеспечение в виде объектов – команд Mothur. Преобразование – логический вывод, реализующий распознавание образов над данными CIM. Модель CIM также получается автоматически в результате анализа исходного кода C++ пакета Mothur. Процедура реализована в среде Python, где происходит сканирование исходного кода и выявление заданных структур. Каждая реализация команды анализируется с помощью набора регулярных выражений, организованных в сценарии.

**3.1. Представление модельных данных в RDF.** Данные исходной модели представлены на основе технологий Семантической паутины (Semantic Web). Модель и составляющие ее структуры идентифицируются глобально как ресурсы. Отношения между ресурсами и литералами выражаются с помощью предикатов стандартных и специально разработанных онтологий. Использование онтологий позволило стандартизировать структуры метаданных и, используя опыт разработчиков онтологий, сузить пространство поиска решений.

В облачном DaaS предполагается хранить файлы и их содержимое как объект с его метаданными. Современные стандарты OMG описывают спецификации преобразования реляционных метаданных UML, SysML в представление RDF. Таким образом, данные хранятся в обычных реляционных базах данных или базах данных «ключ-значение» совместно с их моделями. При проектировании облачного хранилища используется декларативный язык Hipster Domain Language [12] и его инструменты для создания структур баз данных, преобразования метаданных, формальных представлений онтологий для хранимых данных.

Метаданные хранимых объектов базы данных, в основном, описывают отношения между ресурсом и его атрибутами. Некоторые атрибуты, а именно внешние ключи, являются ссылками на другие ресурсы, которые также отражаются в метаданных. Между ресурсами встречаются редкие отношения, которые не хранятся в обычных базах данных. Эти отношения отражают, например, происхождение данных, дополнительные специальные атрибуты для конкретного объекта. Редкие связи встречаются нечасто и появляются в отдельных исследованиях, поэтому изменение структуры реляционной базы данных для каждого такого случая не имеет смысла. Для хранения таких отношений будем использовать RDF-хранилище ClioPatria.

В формализации данных адаптированы следующие стандартизированные онтологии: онтология Friend-of-a-friend (foaf), которая используется для представления информации об агентах: физических, юридических лицах, программах; Provenance (prov), используемая для ссылок между документами; Dublin Core (dc), которая используется для разметки метаданных опубликованных ресурсов; ресурсы DBpedia (dbp) относят данные к внешним глобально используемым классам и объектам-экземплярам; Open annotation (oa), которая используется в качестве онтологии представления содержания опубликованного документа; Bibliographic Ontology (bibo), используемая для разметки литературы. Для представления Mothur CIM и PIM разработаны две онтологии mothur и uml. Онтологии CIM и PIM используются для представления отношений между хранимыми объектами, как субъектами входных и выходных данных команд Mothur.

Используемые инструменты трансформации позволили авторам решить множество технических проблем, в том числе укомплектовать систему dataflow-визуализации актуальным набором команд Mothur, реализовать абстрактный механизм отображения свойств команд Mothur в программную среду. В рамках разработки получен набор сценариев преобразования, выраженных в виде объектов на языке программирования Logtalk [20]. Эти результаты можно использовать для генерации PSM и исходных кодов для внедрения команд Mothur в новые средства вычислений и визуализации.

**3.2. Интеграция данных: вывод метаданных.** Общепринятая форма представления полученных результатов NGS – это публикация в виде научных статей, отчетов и таблиц в различных форматах, таких, как HTML, Word, Excel, PDF. Дополнительная семантическая разметка RDF/RDFa в этих файлах позволит как исследователю, так и программному агенту использовать ранее полученные результаты в новых задачах. Разметка RDF документов является частью сервисов DaaS, поддерживающих также LOD (Открытые связанные данные), и обеспечивающих интеграцию с другими Интернет-ресурсами NGS. К данному моменту не найдено стандартизованного способа интеграции: есть только прототипы аннотационных ресурсов, такие, как BioSearch [13], реализованные на основе технологий BIO2RDF и LOD.

Служба LOD и необходимая гибкость программного обеспечения для научных исследований требуют, чтобы метаданные были связаны со всеми элементами данных NGS. Метаданные входных данных MiSeq SOP преобразуются при каждом применении команд в метаданные, описывающие объекты выходных данных. Для Mothur синтезируются правила автоматического вывода метаданных на основе анализа его исходного кода C++ и алгоритмов преобразования имен файлов. Для экономии памяти, расходуемой на метаданные, решено реализовать динамическое восстановление метаданных по структурам баз данных. Поскольку тысячи последовательностей организованы в файлы, группы и OTU, метаданные последовательностей дополняются всеми метаданными файла/группы/OTU. Метаданные каждой последовательности генерируются с использованием контекста своего хранилища, а именно имени файла fasta и отношения к своей группе, происхождению файла и т. д. Текущая архитектура сервисов обработки метаданных показана на рисунке 3.

**Заключение.** Предложен подход к созданию ИТ-инфраструктуры поддержки исследований микробиома озера Байкал на основе секвенирования следующего поколения. Хорошая база алгоритмов и программного обеспечения, уже созданных разработчиками, позволяет авторам реализовать среду, применяя адаптацию существующих методов, используемых биологами, к инструментарию проектирования вычислительных процессов. Для этого реализуется модель потока данных методики MiSeq, поддерживаемая прикладным пакетом Mothur. Модель интерпретируется в виде программных модулей различных визуальных сред и облачных сервисов. Преобразование осуществляется с использованием архитектуры, управляемой моделями, где трансформация модели реализована как система логического вывода, что позволяет быстро переходить от одной платформы к другой, сохраняя полученный формализованный опыт. На данном этапе существующая реализация ограничена спецификой программного обеспечения Mothur, также пока не было необходимости автоматизировать преобразование форматов данных.

Следующий этап НИРОКР связан с реализацией преобразования данных в форматы, поддерживаемые внешними по отношению к пакету Mothur программами, например,

открытым QIME2 или коммерческим программным обеспечением Usearch. Эти пакеты дополняют средства Mothur в части функций визуализации и модулями обработки данных на отдельных стадиях MiSeq.

Набор задач, требующих решения, включает оптимизацию использования вычислительных ресурсов кластера, используемого биологами в настоящее время; планирование выполнения параллельных вычислений на основе анализа структуры процесса и свойств алгоритмов Mothur, а также реализацию контрольных точек для служб.

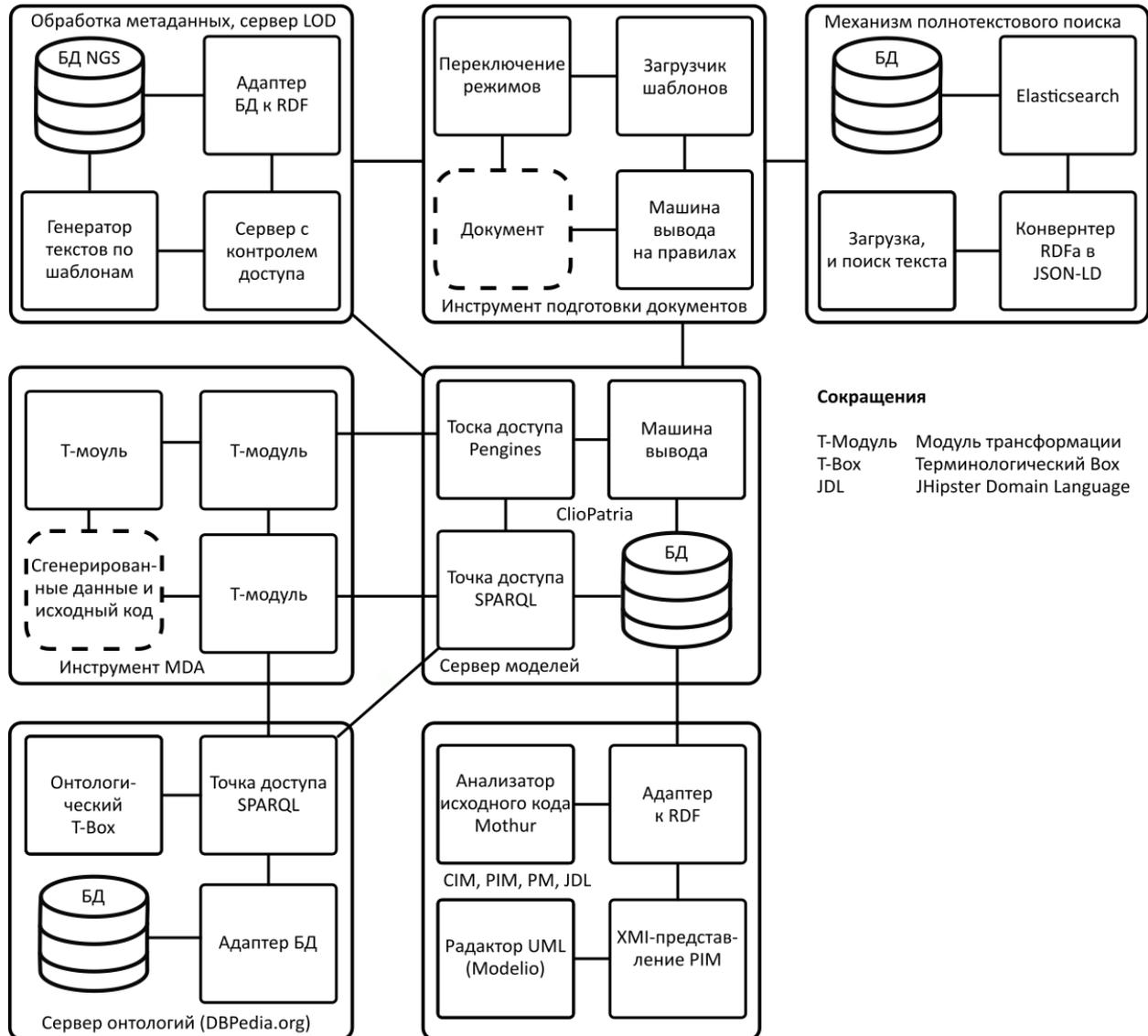


Рис.3. Архитектура облачных сервисов обработки метаданных

### СПИСОК ЛИТЕРАТУРЫ

1. Боекхорст Р, Науменко Ф. М., Орлова Н. Г., Галиева Э. Р., Спицина А. М., Чадаева И. В., Орлов Ю. Л., Абнизова И. И. Вычислительные проблемы анализа ошибок коротких прочтений ДНК при секвенировании следующего поколения // Вавиловский журнал генетики и селекции. 2016. №6 (20). С. 746-755.
2. Amstutz P., Crusoe M. R., Tijanic N., Chapman B. Common workflow language. V. 1.0. 2016. DOI:10.6084/m9.figshare.3115156.v2
3. Baker Q. B., Al-Rashdan W., Jararweh Y. Cloud-based tools for next-generation sequencing data analysis // Procs. of Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS). Valencia. 2018. Pp. 99–105. DOI:10.1109/SNAMS.2018.8554515
4. Bashenkaeva M. V., Zakharova Y. R., Petrova D. P. Sub-ice microalgal and bacterial communities in freshwater lake Baikal // Environmental Microbiology. 2015. Vol.70. № 3. Pp. 751–765. DOI:10.1007/s00248-015-0619-2
5. Batut B., Hiltemann S., Bagnacani A., Baker D., Bhardwaj V. Community-driven data analysis training for biology cell systems. 2018. DOI:10.1016/j.cels.2018.05.012
6. Cherkashin E., Shigarov A., Paramonov V. Representation of MDA transformation with logical objects // International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). Novosibirsk. Russia. 2019. Pp. 0913–0918. DOI:10.1109/SIBIRCON48586.2019.8958008/
7. Cherkashin E., Shigarov A., Malkov F., Morozov A. An instrumental environment for metagenomic analysis // Information Technologies in the Research of Biodiversity. Springer Proceedings in Earth and Environmental Sciences. 2019. Pp. 151–158. DOI:10.1007/978-3-030-11720-7.
8. Costanza P., Herzeel C., Verachtert W. A comparison of three programming languages for a full-fledged next generation sequencing tool // BMC Bioinformatics. 2019. Vol. 20. № 1. 301p. DOI:10.1186/s12859-019-2903-5
9. Gong Y.-N., Chen G.-W., Yang S.-L., Lee C.-J. A next-generation sequencing data analysis pipeline for detecting unknown pathogens from mixed clinical samples and revealing their genetic diversity. 2016. DOI:10.1371/journal.pone.0151495
10. Gritsenko A. A., Nijkamp J. F., Reinders M.J.T., D. de Ridder. GRASS: a generic algorithm for scaffolding next-generation sequencing assemblies // Bioinformatics. 2012. Vol.28. № 11. Pp. 1429–1437. DOI:10.1093/bioinformatics/bts175
11. Guo X., Yu N., Li B., Pan Y. Cloud computing for next-generation sequencing data analysis. in Computational Methods for Next Generation Sequencing Data Analysis. John Wiley & Sons. 2016. Pp. 3-24.
12. Halin A., Nuttinck A., Acher M., Devroey X., Perrouin G., Heymans P. Yo. JHipster: a playground for web-apps analyses // Procs. of the Eleventh international workshop on variability modelling of software-intensive systems. VAMOS'17. ACM. New York. 2017. Pp. 44–51. DOI:10.1145/3023956.3023963.
13. Hu W., Qiu H., Huang J., Dumontier M. BioSearch: a semantic search engine for Bio2RDF. Database. Vol. 2017. bax059. DOI:10.1093/database/bax059.

14. Johnston W. M., Hanna J. R. P., Millar R. J. Advances in dataflow programming languages // *ACM Computing Surveys*. 2004. Vol. 36. Pp 1–34. DOI:10.1145/1013208.1013209
15. Kozich J. J., Westcott S. L., Baxter N. T., Highlander S. K., Schloss P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*. 2013. Vol. 79. № 17. Pp. 5112–5120. DOI:10.1128/AEM.01043-13
16. Kwon T., Yoo W. G., Lee W. Next-generation sequencing data analysis on cloud computing // *Genes Genom*. 2015. Vol. 37. Pp. 489–501. DOI:10.1007/s13258-015-0280-7
17. Langmead B., Nellore A. Cloud computing as a platform for genomic data analysis and collaboration // *Nat Rev Genet*. 2018. Vol. 19. № 4. Pp. 208–219. DOI:10.1038/nrg.2017.113
18. Mikhailov I. S., Zakharova Y. R., Bukin Y. S. Co-occurrence networks among bacteria and microbial eukaryotes of lake Baikal during a spring phytoplankton bloom. *Microbial Ecology*. 2019. Vol. 77. Pp. 96–109. DOI:10.1007/s00248-018-1212-2.
19. Milicchio F., Rose R., Bian J. Visual programming for next-generation sequencing data analytics // *BioData Mining*. 2016. Vol. 9. № 16. DOI:10.1186/s13040-016-0095-3.
20. Moura P. .Programming patterns for Logtalk parametric objects // *Applications of Declarative Programming and Knowledge Management*. A. Abreu. D. Seipel eds. INAP 2009. *Lecture Notes in Computer Science*. Vol. 6547. Berlin. Heidelberg. Pp. 52–69. 2009. DOI:10.1007/978-3-642-20589-7\_4
21. Müller H., Reihls R., Posch A. E., Kremer A., Ulrich D., Zatloukal K. Data driven GUI design and visualization for a NGS based clinical decision support system. // *Procs. of 20th International Conference Information Visualisation*. 2016. Lisbon. Portugal. Pp. 355–360. DOI:10.1109/IV.2016.79
22. Pereira R., Oliveira J., Sousa M. Bioinformatics and computational tools for next-generation sequencing analysis in clinical genetics. *J. Clin. Med.*. Vol. 9. № 1. Pp. 1–30. 2020. DOI:10.3390/jcm9010132
23. Quast Ch., Pruesse E., Yilmaz P., Gerken J. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*. 2013. Vol. 41. D590–D596. DOI:10.1093/nar/gks1219.
24. Rose R., Golosova O., Sukhomlinov D., Tiunov A., Prospero M. Flexible design of multiple metagenomics classification pipelines with UGENE // *Bioinformatics*. 2019. Vol. 35. № 11. Pp. 1963–1965. DOI:10.1186/s13040-016-0095-3
25. Srimani J. K., Wu P., Phan J. H., Wang M. D. A distributed system for fast alignment of next-generation sequencing data // *2010 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*. Hong Kong. 2010. Pp. 579-584. DOI:10.1109/BIBMW.2010.5703865
26. Zhao S., Watrous K., Zhang Ch., Zhang B. Cloud computing for next-generation sequencing data analysis // *Cloud Computing–Architecture and Applications*. IntechOpen Limited. 2017. Pp. 29–51. DOI:10.5772/66732

UDK 004.6::574.5

## INFORMATION INFRASTRUCTURE FOR SUPPORTING BAIKAL MICROBIOME RESEARCH

**Evgeny A. Cherkashin**

Matrosov Institute for System Dynamics and Control theory of  
Siberian Branch of Russian Academy of Sciences

e-mail: [eugeneai@icc.ru](mailto:eugeneai@icc.ru)

**Alexey O. Shigarov**

Irkutsk Scientific Center of Siberian Branch of Russian Academy of Sciences  
Irkutsk, Russia

e-mail: [shigarov@icc.ru](mailto:shigarov@icc.ru)

664033, Irkutsk, st. Lermontov, 134.

**Vasily V. Khristyuk**

National Research Irkutsk State Technical University

e-mail: [khr@icc.ru](mailto:khr@icc.ru)

664074, Irkutsk, st. Lermontov, 83.

**Abstract.** A problem of construction of a research environment for Next Generation Sequencing data processing is considered. The environment comprises cloud data storage (DaaS) and computational services (SaaS and PaaS), as well as visualization, and data integration services. We are integrating existing open-source technologies to support MiSeq SOP (standard operational procedure), which is to allow domain specialists, biologists, to process data independently. For the realization of the integration, formal models of the SOP are constructed, automatically processed (transformed) into source code of new components. The technique of the transformation is based on Model Driven Architecture principles and logical inference of the derived models and the code. The current results are presented and discussed.

**Keywords:** next generation sequencing, big data, model driven architecture, linked open data, problem solving

### Referens

1. Boekhorst R., Naumenko F. M., Orlova N. G., Galieva E. R., Spitsina A. M. Computational problems of analysis of short next generation sequencing reads // Vavilovskii Zhurnal Genetiki i Selektzii = Vavilov Journal of Genetics and Breeding. 2016. Vol. 20. № 6. Pp. 746–755. DOI:10.18699/VJ16.191
2. Amstutz P., Crusoe M. R., Tijanic N., Chapman B. Common workflow language. V. 1.0. 2016. DOI:10.6084/m9.figshare.3115156.v2
3. Baker Q. B., Al-Rashdan W., Jararweh Y. Cloud-based tools for next-generation sequencing data analysis // Procs. of Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS). Valencia. 2018. Pp. 99–105. DOI:10.1109/SNAMS.2018.8554515

4. Bashenkhayeva M. V., Zakharova Y. R., Petrova D. P. Sub-ice microalgal and bacterial communities in freshwater lake Baikal, Russia. *Environmental Microbiology*. Vol.70. № 3. 2015. Pp. 751–765. DOI:10.1007/s00248-015-0619-2
5. Batut B., Hiltemann S., Bagnacani A., Baker D., Bhardwaj V. Community-driven data analysis training for biology cell systems. 2018. DOI:10.1016/j.cels.2018.05.012
6. Cherkashin E., Shigarov A., Paramonov V. Representation of MDA transformation with logical objects //International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). Novosibirsk. Russia. 2019. Pp. 0913–0918. DOI:10.1109/SIBIRCON48586.2019.8958008/
7. Cherkashin E., Shigarov A., Malkov F., Morozov A. An instrumental environment for metagenomic analysis // Information Technologies in the Research of Biodiversity. Springer Proceedings in Earth and Environmental Sciences. 2019. Pp. 151–158. DOI:10.1007/978-3-030-11720-7.
8. Costanza P., Herzeel C., Verachtert W. A comparison of three programming languages for a full-fledged next generation sequencing tool // BMC Bioinformatics. 2019. Vol. 20. № 1. 301p. DOI:10.1186/s12859-019-2903-5
9. Gong Y.-N., Chen G.-W., Yang S.-L., Lee C.-J. A next-generation sequencing data analysis pipeline for detecting unknown pathogens from mixed clinical samples and revealing their genetic diversity. 2016. DOI:10.1371/journal.pone.0151495
10. Gritsenko A. A., Nijkamp J. F., Reinders M.J.T., D. de Ridder. GRASS: a generic algorithm for scaffolding next-generation sequencing assemblies // Bioinformatics. 2012. Vol.28. № 11. Pp. 1429–1437. DOI:10.1093/bioinformatics/bts175
11. Guo X., Yu N., Li B., Pan Y. Cloud computing for next-generation sequencing data analysis. in *Computational Methods for Next Generation Sequencing Data Analysis*. John Wiley & Sons. 2016. Pp. 3-24.
12. Halin A., Nuttinck A., Acher M., Devroey X., Perrouin G., Heymans P. Yo. JHipster: a playground for web-apps analyses // Procs. of the Eleventh international workshop on variability modelling of software-intensive systems. VAMOS'17. ACM. New York. 2017. Pp. 44–51. DOI:10.1145/3023956.3023963.
13. Hu W., Qiu H., Huang J., Dumontier M. BioSearch: a semantic search engine for Bio2RDF. Database. Vol. 2017. bax059. DOI:10.1093/database/bax059.
14. Johnston W. M., Hanna J. R. P., Millar R. J. Advances in dataflow programming languages //ACM Computing Surveys. 2004. Vol. 36. Pp. 1–34. DOI:10.1145/1013208.1013209
15. Kozich J. J., Westcott S. L., Baxter N. T., Highlander S. K., Schloss P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*. 2013. Vol. 79. № 17. Pp. 5112–5120. DOI:10.1128/AEM.01043-13
16. Kwon T., Yoo W. G., Lee W. Next-generation sequencing data analysis on cloud computing // *Genes Genom*. 2015. Vol. 37. Pp. 489–501. DOI:10.1007/s13258-015-0280-7
17. Langmead B., Nellore A. Cloud computing as a platform for genomic data analysis and collaboration // *Nat Rev Genet*. 2018. Vol. 19. № 4. Pp. 208–219. DOI:10.1038/nrg.2017.113

18. Mikhailov I. S., Zakharova Y. R., Bukin Y. S. Co-occurrence networks among bacteria and microbial eukaryotes of lake Baikal during a spring phytoplankton bloom. *Microbial Ecology*. 2019. Vol. 77. Pp. 96–109. DOI:10.1007/s00248-018-1212-2.
19. Milicchio F., Rose R., Bian J. Visual programming for next-generation sequencing data analytics // *BioData Mining*. 2016. Vol. 9. № 16. DOI:10.1186/s13040-016-0095-3.
20. Moura P. .Programming patterns for Logtalk parametric objects // *Applications of Declarative Programming and Knowledge Management*. A. Abreu, D. Seipel eds. INAP 2009. *Lecture Notes in Computer Science*. Vol. 6547. Berlin. Heidelberg. Pp. 52–69. 2009. DOI:10.1007/978-3-642-20589-7\_4
21. Müller H., Reihls R., Posch A. E., Kremer A., Ulrich D., Zatloukal K. Data driven GUI design and visualization for a NGS based clinical decision support system. // *Procs. of 20th International Conference Information Visualisation*. 2016. Lisbon. Portugal. Pp. 355–360. DOI:10.1109/IV.2016.79
22. Pereira R., Oliveira J., Sousa M. Bioinformatics and computational tools for next-generation sequencing analysis in clinical genetics. *J. Clin. Med.* Vol. 9. № 1. Pp. 1–30. 2020. DOI:10.3390/jcm9010132
23. Quast Ch., Pruesse E., Yilmaz P., Gerken J. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*. 2013. Vol. 41. D590–D596. DOI:10.1093/nar/gks1219.
24. Rose R., Golosova O., Sukhomlinov D., Tiunov A., Prospero M. Flexible design of multiple metagenomics classification pipelines with UGENE // *Bioinformatics*. 2019. Vol. 35. № 11. Pp. 1963–1965. DOI:10.1186/s13040-016-0095-3
25. Srimani J. K., Wu P., Phan J. H., Wang M. D. A distributed system for fast alignment of next-generation sequencing data // *2010 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*. Hong Kong. 2010. Pp. 579-584. DOI:10.1109/BIBMW.2010.5703865
26. Zhao S., Watrous K., Zhang Ch., Zhang B. Cloud computing for next-generation sequencing data analysis // *Cloud Computing–Architecture and Applications*. IntechOpen Limited. 2017. Pp. 29–51. DOI:10.5772/66732

**ИНФОРМАЦИОННОЕ СОГЛАСОВАНИЕ  
НАРОДНОХОЗЯЙСТВЕННОЙ И ОТРАСЛЕВОЙ МОДЕЛИ  
НА ПРИМЕРЕ ЛЕСНОГО КОМПЛЕКСА РФ**

**Стойлова Алина Сергеевна**

Аспирант, e-mail: [stoylova.as@mail.ru](mailto:stoylova.as@mail.ru),

**Машкина Людмила Вячеславовна**

К.э.н., e-mail: [ludmila.mashkina@gmail.com](mailto:ludmila.mashkina@gmail.com),

**Блам Юрий Шабсович**

К.э.н., e-mail: [blamukel@gmail.com](mailto:blamukel@gmail.com),

Институт экономики и организации промышленного производства

Сибирского отделения Российской академии наук,

630090, Россия, Новосибирск, проспект Академика Лаврентьева, 17

**Аннотация.** Подход к моделированию, описанный в статье, позволяет на основе Оптимизационной Межрегиональной Межотраслевой Модели (ОМММ) сформировать отраслевую модель лесного комплекса производственно-транспортного типа. С помощью этой модели можно учесть заготовку и переработку лесного сырья, производство промежуточной и конечной продукции, действующие и новые мощности, нагрузку на лесной комплекс со стороны народного хозяйства, а также импорт, экспорт и транспортные связи между регионами. Переход от решения задачи на поиск максимума конечного потребления к задаче на нахождение минимума затрат на производство и перевод стоимостных показателей в натуральные позволяет привести модель к виду, наиболее подходящему для анализа отраслевых особенностей и прогнозирования развития лесного комплекса.

**Ключевые слова:** лесной комплекс, производственно-транспортная модель, ОМММ, отраслевая модель

**Цитирование:** Стойлова А.С., Машкина Л.В., Блам Ю.Ш. Информационное согласование народнохозяйственной и отраслевой модели на примере лесного комплекса РФ //Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 124 - 131. DOI: 10.38028/ESI.2020.20.4.011

**Введение.** Построение модификаций прогнозных моделей позволяет получить подходящий инструментарий для исследования проблемы хозяйственного взаимодействия субъектов экономики в условиях ограниченных ресурсов и интерпретировать дальнейшее развитие народного хозяйства, регионов и отдельных отраслей. Для построения прогнозов в рамках конкретной задачи необходимо оптимальным образом упростить сложные взаимодействия между субъектами, введя системы предположений, оценок и ограничений. При этом в ходе построения идеализированной математической модели требуется наиболее точно передать существующие взаимосвязи между компонентами, чтобы по результатам решения на основе полученной информации дать экспертные рекомендации.

**Методы и материалы.** В качестве основы для последующих структурных модификаций (рис. 1) выступает ОМММ (Оптимизационная Межрегиональная

Межотраслевая Модель) [1, 5-8]. В эту модель входят балансовые ограничения по производству и распределению продукции, ограничения на трудовые ресурсы, инвестиции, ограничения на объемы выпуска и прироста продукции, ограничения на региональное внешнеторговое сальдо и внешнеторговый баланс, ограничения на объемы экспорта и импорта и ограничения на территориальную структуру конечного потребления. Исходная модель описывает 40 отраслей и 8 регионов, среди которых только 3 относятся к лесному комплексу:

- Деревообработка
- Целлюлозно-бумажная промышленность
- Охота и лесное хозяйство (лесозаготовки).

В начале работы производится модификация ОМММ с помощью процедур агрегации и дезагрегации. Модель приводится к специализированному виду, в ее информационном блоке сокращается количество отраслей, не относящихся к лесному комплексу, и увеличивается количество отраслей лесного комплекса. После всех начальных преобразований модель содержит 12 лесных и 13 не лесных отраслей. Эта модель получает название ОМММ25 и является основой для следующих преобразований [2-4, 10].

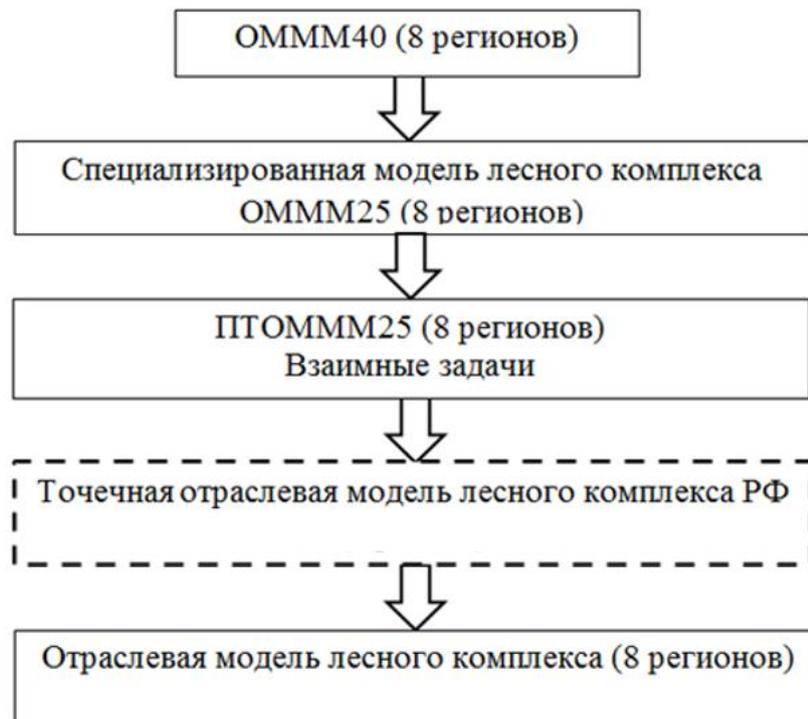


Рис. 1. Схема модификации ОМММ

На следующем этапе происходит переход от задачи на поиск максимума конечного потребления к задаче на нахождение минимума трудовых затрат с использованием теории взаимности (рис. 2) [9].

**Обозначения к рис. 2:**

$x_j, y_j$  – план производства для прямой и взаимной задачи соответственно;

$m$  – виды ресурсов в количестве  $b_1, b_2, \dots, b_m$  соответственно;

$n$  – виды продукции стоимостью  $c_1, c_2, \dots, c_n$  соответственно;

$a_{ij}$  – нормы потребления ресурсов для производства единицы каждого вида продукции;

$d_j$  – дефицитный ресурс.

$\alpha$  – ограничение дефицитного ресурса

$\beta$  – величина максимальной прибыли

Прямая задача	Взаимная задача
$\sum_{j=1}^n c_j x_j \rightarrow \max;$ $\sum_{j=1}^n d_j x_j \leq \alpha \ (\alpha > 0);$ $\sum_{j=1}^n a_{ij} x_j \leq b_i \text{ для } i = 1, \dots, m;$	$\sum_{j=1}^n d_j y_j \rightarrow \min$ $\sum_{j=1}^n a_{ij} y_j \leq b_i \ (1 \leq i \leq m)$ $\sum_{j=1}^n c_j y_j \geq \beta \ (\beta > 0)$ $y_j \geq 0;$

**Рис. 2.** Схема перехода к взаимной задаче

На уровне оптимального решения фиксируется значение фонда конечного потребления, которое в дальнейшем принимается за ограничение снизу. Все ограничения исходной задачи сохраняются. Дополнительно вводится ограничение на трудовые ресурсы по стране в целом, которое является лимитирующим. Новая задача позволяет при фиксированном максимуме конечного потребления найти необходимый минимум трудовых ресурсов для его достижения [12, 13].

Далее совершается переход к отраслевой задаче производственно-транспортного типа. Переменные, не относящиеся к производству лесной продукции, её перевозкам, непроизводственному потреблению, экспорту и импорту, взвешиваются на значения, полученные в оптимальном решении ОМММ25, и переносятся в правую часть балансовых ограничений. Таким образом, формируется нагрузка на лесной комплекс со стороны народного хозяйства. Полученные при решении такой модели двойственные оценки можно интерпретировать, как стоимость продукции в трудовом выражении.

После формирования нагрузки на лесной комплекс из описания исключаются все трудовые затраты, относящиеся к производству лесной продукции не лесными отраслями. В результате среди балансовых ограничений задачи остаются только те, которые непосредственно связаны с производством лесной продукции и использованием трудовых ресурсов лесным комплексом в каждом регионе. Для завершения модификации выделенные показатели взвешиваются на полученные в ходе решения ОМММ25 оценки в трудовом выражении и суммируются с коэффициентами целевой функции.

Нагрузка на отрасль со стороны народного хозяйства переводится из стоимостных показателей в натуральные. Для этого рассчитываются цены за единицу продукции, которые затем умножаются на соответствующие показатели правых частей. После этого формируются новые блоки модели: блок заготовки и переработки сырья и промежуточной продукции. Технологические способы производства рассчитываются, исходя из данных, опубликованных отраслевыми институтами. С помощью статистической информации рассчитываются

коэффициенты целевой функции, затраты на производство продукции лесопромышленного комплекса, средние цены на сырьё.

С использованием полученной отраслевой модели проводятся серии расчетов, которые можно использовать для интерпретации отраслевых документов (например, Стратегии развития лесного комплекса до 2030 года) [11, 14, 15]. В таблице 1 представлено сравнение среднегодовых темпов роста объемов производства, полученных с помощью точечной отраслевой модели, и Стратегии развития лесного комплекса РФ до 2030 года.

**Таблица 1.** Сравнение среднегодовых темпов роста объемов производства в точечной отраслевой модели (ТОМ-лк) и Стратегии-2030

Темпы	ТОМ-лк	Стратегия (инерц.)	Стратегия (страт.)
Пр-во пиломат.	1.004	1.010	1.023
Пр-во фанеры	1.012	1.018	1.029
Пр-во целлюлозы	1.000	1.000	1.125

Таким образом, результаты расчетов по точечной отраслевой модели показывают, что прогнозы Стратегии-2030 завышены.

**Заключение.** Отраслевая модель, построенная на основе ОМММ и информационно согласованная с ней, позволяет получить более детальный прогноз развития лесного комплекса. Модель такого рода содержит в себе одновременно информацию нижнего и верхнего модельных уровней – отраслевого и народнохозяйственного, – что позволяет анализировать отдельную (лесную) отрасль без отрыва от народного хозяйства страны.

#### СПИСОК ЛИТЕРАТУРЫ

1. Бабенко Т.И., Блам Ю.Ш., Машкина Л.В. СОНАР-ЛПК: средства моделирования и анализа лесопромышленного комплекса // Системное моделирование и анализ мезо- и микроэкономических объектов. Новосибирск: Изд-во ИЭОПП СО РАН. 2014. С. 147–183.
2. Блам Ю.Ш., Машкина Л.В., Стойлова А.С. Об одном подходе к детализации народнохозяйственного прогноза развития отрасли (на примере лесного комплекса) // Мир экономики и управления. 2016. Т. 16. № 4. С. 39-47.
3. Блам Ю.Ш., Машкина Л.В., Стойлова А.С. Детализация прогнозов по народнохозяйственной модели в натуральные показатели отраслевой модели (на примере лесного комплекса) // Экономика и управление инновациями. 2017. № 2. С. 66-77.
4. Блам Ю.Ш., Машкина Л.В. Построение иерархического набора моделей: от стоимостной ОМММ к отраслевой модели в натуральных показателях // Мир экономики и управления. 2018. Т. 18. № 4. С. 126-139.
5. Бузулуцков В.Ф., Суслов Н.И. СОНАР-ТЭК: моделирование и анализ проблем энергетического комплекса в системе национальной экономики // Системное моделирование и анализ мезо- и микроэкономических объектов. Новосибирск: Изд-во ИЭОПП СО РАН. 2014. С. 40–111.
6. Гранберг А.Г. Моделирование пространственного развития национальной и мировой экономики: эволюция подходов // Регион: экономика и социология. 2007. № 1. С. 87–107.

7. Гранберг А.Г., Суслов В.И., Суспицын С.А. Многорегиональные системы: экономико-математическое исследование. Новосибирск: Сибирское Научное Издательство. 2007. 371 с.
8. Ершов Ю.С., Ибрагимов Н.М., Мельникова Л.В. Современные постановки прикладных межрегиональных межотраслевых моделей // Исследования многорегиональных экономических систем: опыт применения оптимизационных межрегиональных межотраслевых систем. Новосибирск: ИЭОПП СО РАН. 2007. С.29-59
9. Ицкович И. А. Анализ линейных экономико-математических моделей. Новосибирск: Наука. 1976. 190 с.
10. Мкртчян Г.М., Блам Ю.Ш., Машкина Л.В. Рациональное агрегирование отраслевой номенклатуры народнохозяйственной модели с детализированным лесным комплексом (оценка на основе экспериментальных расчетов) // Вестник Новосибирского государственного университета. 2015. Т.15. вып. 3. С. 56-63.
11. Об утверждении Стратегии развития лесного комплекса до 2030 года. Режим доступа: <http://government.ru/docs/34064/> (дата обращения: 25.10.2020).
12. Стойлова А.С. Моделирование иерархии прогнозов: от стоимостного народнохозяйственного к "натуральному" отраслевому // Экономика Сибири в условиях глобальных вызовов XXI века. Ин-т экон. и организации пром. пр-ва СО РАН. Новосибирск: Изд-во ИЭОПП СО РАН. Т. 4. 2018. С. 228-239.
13. Стойлова А.С. Точечная отраслевая модель лесного комплекса как промежуточный этап комплексного моделирования // Актуальные вопросы экономики и социологии. Новосибирск: ИЭОПП СО РАН. 2019. ISBN 978-5-89665-343-1. С. 589-594.
14. Стойлова А.С. Оценка дефицитности крупномерного лесосырья в рамках стратегии развития лесного комплекса 2030 // Международная научная конференция "Экономическое развитие Сибири и Дальнего Востока. Экономика природопользования, землеустройство, лесоустройство, управление недвижимостью". Новосибирск: СГУГиТ. 2019. DOI: 2618-981X-2019-3-1. Т. 3. № 1. С. 231-238.
15. Стойлова А.С. Количественная оценка прогнозов "Стратегии развития лесного комплекса - 2030" на основе комплекса моделей // Мир экономики и управления. 2019. Т. 9. № 4. С. 114-126. DOI: 10.25205/2542-0429-2019-19-4-114-126

UDK 330.44

**INFORMATION COORDINATION OF THE NATIONAL ECONOMIC AND INDUSTRY  
MODELS ON THE EXAMPLE OF THE RUSSIAN FEDERATION FORESTRY**

**Alina S. Stoylova**

Graduate student, e-mail: [stoylova.as@mail.ru](mailto:stoylova.as@mail.ru),

**Lyudmila V. Mashkina**

PhD in Economics, e-mail: [ludmila.mashkina@gmail.com](mailto:ludmila.mashkina@gmail.com),

**Yuriy Sh. Blam**

PhD in Economics, , e-mail: [blamukel@gmail.com](mailto:blamukel@gmail.com),

Institute of Economics and Industrial Engineering

of Siberian Branch of the Russian Academy of Sciences,

17, Academician Lavrentyev Avenue, Novosibirsk, 630090, Russia

**Abstract.** The modeling approach described in this article allows, on the basis of OIIM, to build the forestry industry model of the production-transport type. Using this model, it is possible to take into account the harvesting and processing of timber, the production of intermediate and final products, existing and new capacities, the load on the forest complex from the national economy, as well as import, export and transport links between regions. The transition from solving the problem of finding the maximum final consumption to the problem of finding the minimum cost of production and converting cost indicators into natural values allows to transform the model to the form most suitable for analyzing industry characteristics and predicting the development of the complex.

**Keywords:** forestry, model of the production-transport type, OIIM, industry model

**References**

1. Babenko T.I., Blam YU.SH., Mashkina L.V. SONAR-LPK: sredstva modelirovaniya i analiza lesopromyshlennogo kompleksa [SONAR-LPK: tools for modeling and analysis of the timber industry complex] // Sistemnoe modelirovanie i analiz mezo- i mikroekonomicheskikh ob"ektov. Novosibirsk: Izd-vo IEOPP SO RAN =System modeling and analysis of meso- and microeconomic objects. Novosibirsk: Publishing house Institute of Economics and Industrial Organization of the Siberian Branch of the Russian Academy of Sciences. 2014. Pp. 147–183.
2. Blam YU.SH., Mashkina L.V., Stojlova A.S. Ob odnom podhode k detalizacii narodnohozyajstvennogo prognoza razvitiya otrasli (na primere lesnogo kompleksa) [On one approach to detailing the national economic forecast for the development of the industry (using the example of the forestry complex)]// Mir ekonomiki i upravleniya = The world of economics and management. 2016 vol. 16. № 4. Pp. 39–47.
3. Blam YU.SH., Mashkina L.V., Stojlova A.S. Detalizaciya prognozov po narodnohozyajstvennoj modeli v natural'nye pokazateli otraslevoj modeli (na primere lesnogo kompleksa) [Detailing forecasts for the national economic model into physical indicators of the industry model (using the example of the forestry complex)] // Ekonomika i upravlenie innovatsiyami = Economics and innovation management. 2017 № 2. Pp. 66–77.

4. Blam YU.SH., Mashkina L.V. Postroenie ierarhicheskogo nabora modelej: ot stoi-mostnoj OMMM k otraslevoj modeli v natural'nyh pokazatelyah [Construction of a hierarchical set of models: from a value OIIM to an industry model in natural terms]// Mir ekonomiki i upravleniya = The world of economics and management. 2018 vol. 18. № 4. Pp. 126–139.
5. Buzuluckov V.F., Suslov N.I. SONAR-TEK: modelirovanie i analiz problem energeticheskogo kompleksa v sisteme nacional'noj ekonomiki [SONAR-TEK: modeling and analysis of the problems of the energy complex in the system of the national economy] // Sistemnoe modelirovanie i analiz mezo- i mikroekonomicheskikh ob"ektov. Novosibirsk: IEOPP SO RAN = System modeling and analysis of meso- and microeconomic objects. Novosibirsk: Publishing house Institute of Economics and Industrial Organization of the Siberian Branch of the Russian Academy of Sciences. 2014. Pp. 40–111.
6. Granberg A.G. Modelirovanie prostranstvennogo razvitiya nacional'noj i mirovoj ekonomiki: evolyuciya podhodov [Modeling the spatial development of the national and world economy: evolution of approaches] // Region: ekonomika i sociologiya = Region: economics and sociology. 2007. № 1. Pp. 87–107.
7. Granberg A.G., Suslov V.I., Suspicyan S.A. Mnogoregional'nye sistemy: ekonomiko-matematicheskoe issledovanie [Multi-regional systems: economic and mathematical research]. Novosibirsk: Sibirskoe Nauchnoe Izdatel'stvo = Novosibirsk: Siberian Scientific Publishing House. 2007. 371p.
8. Ershov YU.S., Ibragimov N.M., Mel'nikova L.V. Sovremennye postanovki prikladnyh mezhregional'nyh mezhotraslevykh modelej [Modern formulations of applied interregional intersectoral models] // Issledovaniya mnogo regional'nyh ekonomicheskikh sistem: opyt primeneniya optimizacionnykh mezhregional'nyh mezhotraslevykh system. Novosibirsk : IEOPP SO RAN = Research of multiregional economic systems: experience in the application of optimized interregional intersectoral systems. Novosibirsk: Publishing House Institute of Economics and Industrial Organization of the Siberian Branch of the Russian Academy of Sciences. 2007. Pp. 29-59
9. Ickovich I. A. Analiz linejnykh ekonomiko-matematicheskikh modelej [Analysis of linear economic and mathematical models]. Novosibirsk. Nauka= Science. 1976. 190 p.
10. Mkrtchyan G.M., Blam YU.SH., Mashkina L.V. Racional'noe agregirovanie otraslevoj nomenklatury narodnohozyajstvennoj modeli s detalizirovannym lesnym kompleksom (ocenka na osnove eksperimental'nykh raschetov) [Rational aggregation of the sectoral nomenclature of the national economic model with a detailed forestry complex (assessment based on experimental calculations)] // Vestnik Novosibirskogo gosudarstvennogo universiteta = Novosibirsk State University Bulletin. 2015. vol. 15. №3 Pp. 56-63.
11. Ob utverzhdenii strategii razvitiya lesnogo kompleksa do 2030 goda [On approval of the Strategy for the development of the forestry complex until 2030]. Available at: <http://government.ru/docs/34064/> (accessed:25.10.2020) (in Russian)
12. Stojlova A.S. Modelirovanie ierarhii prognozov: ot stoimostnogo narodnoho-zyajstvennogo k "natural'nomu" otraslevomu [Modeling the hierarchy of forecasts: from the value of the national economic to the "natural" industry] // Ekonomika Sibiri v usloviyah global'nykh vyzovov XXI veka . Novosibirsk: Izd-vo IEOPP SO RAN = Economy of Siberia in the context of global challenges of the XXI century. Novosibirsk: Publishing House Institute of Economics and Industrial

Organization of the Siberian Branch of the Russian Academy of Sciences. Т. 4. 2018. Pp. 228-239.

13. Stojlova A.S. Tochechnaya otraslevaya model' lesnogo kompleksa kak promezhutochnyj etap kompleksnogo modelirovaniya [The point forestry model as an intermediate step of complex modeling] // Aktual'nye voprosy ekonomiki i sociologii. Novosibirsk: IEOPP SO RAN = Topical issues of economics and sociology. Novosibirsk: Publishing House Institute of Economics and Industrial Organization of the Siberian Branch of the Russian Academy of Sciences. 2019. ISBN 978-5-89665-343-1. Pp. 589-594.

14. Stojlova A.S. Ocenka deficitnosti krupnomernogo lesosyr'ya v ramkah strategii razvitiya lesnogo kompleksa 2030 [Scarcity assessment of heavy timber in the framework of the forestry development strategy 2030] // Mezhdunarodnaya nauchnaya konferenciya "Ekonomicheskoe razvitie Sibiri i Dal'nego Vostoka. Ekonomika prirodopol'zovaniya, zemleustrojstvo, lesoustrojstvo, upravlenie nedvizhimost'yu". Novosibirsk : SGUGiT = International scientific conference "Economic development of Siberia and the Far East. Economics of environmental management, land management, forestry, property management". Novosibirsk: Siberian State University of Geosystems and Technologies. 2019. DOI: 2618-981H-2019-3-1. Vol. 3. № 1. Pp. 231-238.

15. Stojlova A.S. Kolichestvennaya ocenka prognozov "Strategii razvitiya lesnogo kompleksa - 2030" na osnove kompleksa modelej [Model-Based Quantitative Evaluation of 2030 Forestry Development Strategy]. DOI: 10.25205/2542-0429-2019-19-4-114-126 // Mir ekonomiki i upravleniya = The world of economics and management. 2019. vol. 9. № 4. Pp. 114-126.

**О СЕМАНТИЧЕСКОЙ МОДЕЛИ ПРЕДМЕТНОЙ ОБЛАСТИ  
«УРАВНЕНИЯ МАТЕМАТИЧЕСКОЙ ФИЗИКИ»<sup>1</sup>**

**Тучкова Наталия Павловна**

к.ф.-м.н., с.н.с.,

Вычислительный центр им. А.А. Дородницына

Федерального исследовательского центра «Информатика и управление»

Российской академии наук,

e-mail: [natalia\\_tuchkova@mail.ru](mailto:natalia_tuchkova@mail.ru),

ORCID [0000-0001-6518-5817],

г. Москва, 119333, ул. Вавилова, 40

**Аннотация.** Анализируются связи понятий математической предметной области на примере раздела уравнений математической физики. Предлагается вариант статьи тезауруса для терминов и связанных с ними уравнений и формул. Особенность такого тезауруса заключается в использовании контекста формул для их дополнительной идентификации в предметной области. Кроме того, предлагается учитывать индексы авторов и статей, где встречаются термины тезауруса. Предложенный подход способствует уточнению поискового запроса и уменьшению информационного шума при использовании тезауруса в цифровых библиографических коллекциях.

**Ключевые слова:** тезаурус предметной области, уравнения смешанного типа, поисковый запрос.

**Цитирование:** Тучкова Н. П. О семантической модели предметной области «Уравнения математической физики»//Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 132-142. DOI: 10.38028/ESI.2020.20.4.012

**Введение.** Применение тезаурусов в современных цифровых ресурсах является повсеместным средством семантического описания коллекций. Опыт создания лингвистических тезаурусов довольно длинный, начиная от тезауруса английского языка Роже (Roget's Thesaurus) [26], который переиздается уже более 150 лет. Актуальность этого процесса повысилась с появлением языков онтологий консорциума W3C [21, 28]. Технология онтологического проектирования, по определению, позволяет непосредственно воспроизвести структуру и связи, содержащиеся в тезаурусе, средствами программирования [20]. Многочисленные примеры прикладных онтологий и инструментов редактирования данных появились около 2003 года, когда были опубликованы первые рекомендации W3C, в частности, тезаурусы английского языка 21 века [25, 27]. В работе [6] представлена современная модель лингвистической онтологии на основе информационно-поискового тезауруса [5, 18, 22] на примере русскоязычных ресурсов. Опыт создания прикладных тезаурусов по отдельным предметным областям гораздо меньше. Это диктуется спецификой такой работы, которая подразумевает совместные исследования в области информационных

<sup>1</sup> Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований, проект № 18-29-10085мк

технологий и в предметной области. Целый ряд прикладных тезаурусов был создан в 80-е годы прошлого столетия в период расцвета отечественной кибернетики [1, 4, 8, 11-13, 15, 17]. В это же время появился результат работы ведущих ученых-математиков – математическая энциклопедия под редакцией академика И.М. Виноградова [7]. Далее наступила эпоха оцифровки знаний [3] и эта советская энциклопедия стала основой для представления математических знаний в Интернет, где она реализована в русском<sup>2</sup> и английском<sup>3</sup> вариантах, а также как один из источников информации в системе LibMeta<sup>4</sup>.

В настоящей работе развивается онтологический подход к представлению математических знаний в цифровых ресурсах применительно к разделам смешанных задач математической физики.

**1. Математические знания в Интернет.** Представление в Интернет математических знаний – одна из актуальных задач информационных технологий в силу востребованности математики, как в образовании, так и в научных исследованиях. Особенности 21 столетия и дистанционного образования в период пандемии COVID-19 обозначили необходимость достоверных оцифрованных знаний, в том числе и математических. Это не явилось неожиданностью для научного сообщества, поскольку оно уже давно, начиная с первых разработок канадского математика из университета Торонто Кеннета Мэй (Kenneth O. May), занималось систематизацией знаний. Этому ученому принадлежит авторство уникального цифрового проекта по созданию словарей и тезаурусов по математике совместно с издательством McGraw-Hill в 70-е годы прошлого столетия. Тезаурус на основе терминологии математических заметок (Mathematics Notes) и других журналов был буквально собран группой студентов на библиотечных карточках. Современные разработки в этой области относятся, конечно, к использованию онтологического представления словарей и тезаурусов в базах данных и цифровых библиографических коллекциях.

Особенности информационных образов математических предметных областей в контексте задач математической физики были проанализированы в работе [10], там же приведены ссылки на некоторые открытые источники с таблицами иерархических структур и родовидовых отношений в тех случаях, когда это явно представлено в самих источниках. Важным качеством математических и многих других естественно-научных текстов является наличие формул. Математическая формула выступает как свертка информации, которая может быть использована и в описании вторичного документа, и в поисковом образе, также может быть помещена в соответствующую статью словаря или тезауруса. Математическая формула представляет собой уникальный объект предметной области, который несет в себе информацию независимо от языка изложения материала. Это стало поводом для разработки методов использования этой информации в информационно-поисковых системах научных публикаций. Вторичный документ формируется как результат аналитико-синтетической переработки информации. Свертывание информации – это совокупность операций аналитико-синтетической переработки, преследующих цель выразить содержание текста в более короткой форме при сохранении или некотором уменьшении информативности в производном тексте [2, 16]. В качестве сверток информации могут выступать законы, теоремы,

<sup>2</sup> [https://dic.academic.ru/contents.nsf/enc\\_mathematics](https://dic.academic.ru/contents.nsf/enc_mathematics)

<sup>3</sup> [https://encyclopediaofmath.org/wiki/Main\\_Page](https://encyclopediaofmath.org/wiki/Main_Page)

<sup>4</sup> <http://libmeta.ru/>

формулы и тому подобные обобщения, концентрирующие в себе «свертывание» огромного фактического материала.

Одно из первых появлений формул в тезаурусе относится к химии<sup>5</sup>, что вполне оправдано и стало возможным после появления современных средств визуализации. В русскоязычном сегменте в 1972 г. появился «Тезаурус научно-технических наук» [15], в 1978 г. «Тезаурус по теории формальных и алгебраических групп» [4], в 1991 г. «Русско-немецкий информационно-поисковый тезаурус по робототехнике» [1]. В 2005 г. был опубликован «Тезаурус информационно-поисковый по предметной области: обыкновенные дифференциальные уравнения» [9], где были даны основные дескрипторы и связи терминов в этой области. В дальнейшем этот тезаурус был интегрирован в семантическую библиотеку LibMeta<sup>6</sup>.

Можно отметить Zentralblatt MATH (<https://zbmath.org>, Mathematical abstracts and reviews), как информационный ресурс, наиболее адекватно отвечающий запросам пользователей математических библиотек, использующих формулы, как в представлении публикаций, так и в информационном запросе. Здесь содержатся данные о публикациях, начиная с 1826 года по настоящее время. Цель его создания обозначена как сбор, систематизация и распространение библиографических данных, рефератов книг и статей, посвящённых всем разделам математики и её приложениям. В частности, этот проект свидетельствует о том, что интерес к проблеме использования символьной записи в виде формул в поисковом запросе сформировался достаточно давно. Zentralblatt MATH демонстрирует на своей библиографической базе данных преимущества использования символьной информации, формул. Надо отметить, что аналогичный интегрированный проект для отечественных публикаций пока не реализован, хотя ведется много работ в этом направлении. Это – «Общероссийский математический портал» (<http://www.mathnet.ru/>) математического института им. В.А. Стеклова РАН; Интернет-проект (<http://eqworld.ipmnet.ru/indexr.htm>) «Мир математических уравнений» (EqWorld); ресурс «OntoMathpro Ontology classes» (<http://ontomathpro.org/ontology>) Казанского университета, содержащий англо-русский алфавитный указатель терминов с горизонтальными связями, и многие другие.

В направлении интеграции математических ресурсов на основе онтологического представления развивается семантическая библиотека LibMeta, в которую наряду с другими источниками интегрированы математическая энциклопедия, тезаурус обыкновенных дифференциальных уравнений, классификаторы MSC, УДК и др. Расширение предметной области математической физики реализуется через увеличение связей и понятий отдельных разделов, в частности для смешанных задач [14].

**2. Структура статьи тезауруса для смешанных уравнений математической физики.** Уравнения смешанного типа являются частью предметной области уравнений в частных производных (УЧП) и принимают гиперболический тип (hyperbolic type, НТ), параболический тип (parabolic type, РТ) и эллиптический тип (elliptic type, ЕТ), в зависимости от области определения (рассмотрения), как часть УЧП (схема на рис. 1).

На рис. 1 схематически показаны несколько уровней родовидовых отношений для УЧП

<sup>5</sup> <http://www.wssanalytchem.org/ontology/default.aspx>

<sup>6</sup> <http://libmeta.ru/thesaurus/show/19115>

второго порядка с независимыми переменными в самом общем виде. Эта предметная область содержит практически необъятный материал, что затрудняет поиск публикаций конкретной тематики.

Родовой термин	Видовые термины
✓ П второго порядка с независимыми переменными	<ul style="list-style-type: none"> <li>▪ именованные УЧП</li> <li>▪ УЧП                             <ul style="list-style-type: none"> <li>• Линейные УЧП</li> <li>• Линейные УЧП более высоких производных с коэффициентами у более высоких производных                                     <ul style="list-style-type: none"> <li>➤ С постоянными коэффициентами</li> <li>➤ С переменными коэффициентами   <ul style="list-style-type: none"> <li>❖ УЧП гиперболического типа</li> <li>❖ УЧП параболического типа</li> <li>❖ УЧП эллиптического типа</li> </ul> </li> </ul> </li> </ul> </li> </ul>

Рис. 1. Схема родовидовых связей УЧП второго порядка с независимыми переменными

Чем больше признаков можно использовать в поисковом запросе, тем точнее можно осуществить поиск, поэтому состав статьи тезауруса необходимо расширить для того, чтобы было возможно использовать различные точки входа, как минимум, название, автор и формула. Например, для уравнения Трикоми, одного из самых изучаемых в теории смешанных уравнений, статья тезауруса представлена в таблице 1.

Таблица 1. Статья тезауруса для дескриптора «Трикоми уравнение»

Позиции статьи тезауруса	Тип данных, язык	Значения позиций статьи тезауруса
PDE1031(MXT)	Ru	Трикоми уравнение
	En	Tricomi equation
Син1	Ru	Уравнение Т
Син2	En	Equation T
РодMSC	35M10	Equations of mixed type
РодУДК	517.956.6	Уравнения смешанного типа
ВидФорм 1	MSWord	$y \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$
ВидФорм2	LaTex	$\$y \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0\$$
ВидФорм3	LaTex	$\$y u_{xx} + u_{yy} = 0\$$
Лит1	[1ru]	Смирнов М. М. Уравнения смешанного типа. М., 1970. 296 с.
Лит2	[2ru]	Трикоми Ф. Лекции по уравнениям в частных производных. Пер. с итал. Д. А. Райкова с предисл. Б. М. Левитана М.: Изд-во иностр. лит., 1957. 443 с.
Лит3	[1en]	Smirnov M.M. Equations of mixed type. American Translation of the Mathematical monographs. Vol. 51. Mathematical Soc., 31 Dec 1978. p: 232.
Лит4	[2en]	Lupo D., Morawetz C.S., Payne K.R. On closed boundary value problems for equations of mixed elliptic-hyperbolic type // Communications on Pure and Applied Mathematics, 2007, 60 (9), 1319-1348
Род	Ru	PDE1030(MXT)  Уравнения смешанного типа
Род	En	PDE1030(MXT)  Equations of mixed type
Вид 1		PDE1031(CF2,HT,y<0)
Вид 2		PDE1031(CF,ET,y>0)
Вид 3		PDE1031(CF,ET, z = $\eta^\alpha u, \alpha = -\frac{1}{6}$ )
Вид 4		PDE1031(CF,ET, z = $\eta^\alpha u, \alpha = -\frac{1}{3}$ )

КлС		Трикоми задача, обобщение Трикоми задачи, Трикоми метод, задача Коши для уравнения Трикоми, краевая задача для уравнения типа Чаплыгина, краевая задача для уравнения смешанного типа
Асс	Ru	[PDE1037(MXT)] Геллерстедта задача
Асс	En	[PDE1037(MXT)] The Gellerstedt problem
См. также		[PDE1039(MXT)] Газовой динамики уравнение смешанного типа
История		Название термина по имени Трикоми Франческо Джакомо (Tricomi Francesco Giacomo), появилось в [1ru] как ссылка на работу[2ru]
Примечания		[PDE1036(MXT)] Геллерстедта решение задачи Трикоми для уравнения

В Таблице 1 PDE1031(MXT) – это мнемонический идентификатор дескриптора, где PDE – сокращение от Partial Differential Equations (уравнения в частных производных), MXT – сокращение от Mixed Type (смешанный тип), 1031 – номер дескриптора в цифровой библиотеке.

Позиции статьи тезауруса включают родовые, видовые и ассоциативные термины. В качестве видовых добавлены формулы, в качестве родовых – классификаторы. По сравнению со стандартными полями тезауруса ГОСТ [5] добавлены следующие: См. также, История, Примечания. Все эти поля не противоречат стандартам, поскольку могут быть отнесены к ассоциативным терминам и для наглядности вынесены в отдельные позиции.

Основные дескрипторы именных уравнений смешанного типа – это «Трикоми уравнение» и «Геллерстедта задача». Для одного из них выписаны ключевые слова в Таблице 2 со списком публикаций, где они встречаются в Таблице 3.

**Таблица 2.** Ключевые слова дескриптора «Геллерстедта задача»

1	уравнение для Геллерстедта задачи [2, 3, 4,5,7,8,10,11,12,13, 14,15,16,17,18, 19,20]	- equation for Gellerstedt problem - problem G
2	уравнение типа Чаплыгина	- equations of Chaplygin's type - Chaplygin type equation
3	уравнение смешанного типа[2, 3, 4,5,7,8,10,11,12,13, 14,15,16,17,18, 19,20]	- equations of mixed type
4	Трикоми задача [10,12,14,15,16,19,20]	- Tricomi problem - problem T
5	обобщение Трикоми задачи [10,12,14,15,16,19,20]	- generalized Tricomi problem - generalized problem T
6	Трикоми метод [10,12,14,15,16,19,20]	- Tricomi's method - Tricomi method
7	задача Коши для уравнения Трикоми [10,12,14,15,16,19,20]	- Cauchy's problem for Tricomi's equation
8	краевая задача для уравнения типа Чаплыгина	- boundary-value problem Chaplygin type equation
9	краевая задача для уравнения смешанного типа [2, 3, 4,5,7,8,10,11,12,13, 14,15,16,17,18, 19,20]	- boundary-value problem for a mixed-type equation
10	эллиптичность уравнения для Геллерстедта задачи [2, 3, 4,5,7,8,10,11,12,13, 14,15,16,17,18, 19,20]	- ellipticity of the equation for Gellerstedt problem
11	параболичность для уравнения Геллерстедта задачи [2, 3, 4,5,7,8,10,11,12,13, 14,15,16,17,18, 19,20]	- parabolicity of the equation for Gellerstedt problem
12	гиперболичность для уравнения Геллерстедта задачи [2, 3, 4,5,7,8,10,11,12,13, 14,15,16,17,18, 19,20]	- hyperbolicity of the equation for Gellerstedt problem

13	теорема существования и единственности для уравнения Геллерстедта задачи [2, 3, 4,5,7,8,10,11,12,13, 14,15,16,17,18, 19,20]	- existence and uniqueness theorem for the equation for Gellerstedt problem
14	теорема существования и единственности для обобщенной задачи Трикоми	- existence and uniqueness theorems for the generalized Tricomi problem
15	метод Геллерстедта [2, 3, 4,5,7,8,10,11,12,13, 14,15,16,17,18, 19,20]	- Gellerstedt's method
16	уравнение газовой динамики смешанного типа	- gaz dinamycal equation of mixed type
17	приложения уравнения Геллерстедта задачи в околосзвуковой газовой динамике	- applications of the Gellerstedt problem in transsonic gas dynamics
18	уравнение для Геллерстедта задача смешанного типа с оператором Лаврентьева-Бицадзе [2, 3, 4,5,7,8,10,11,12,13, 14,15,16,17,18, 19,20]	- Gellerstedt problem for the mixed type equation with Lavrentyev-Bisadze operator - mixed type problem G with LB operator

**Таблица 3.** Литература к ключевым словам по теме «Геллерстедта задача» из таблицы 2.

1	Математическая Энциклопедия. Т. 1 (А - Г). Ред. коллегия: И. М. Виноградов (глав. ред.) [и др.] - М., «Советская Энциклопедия», 1977, 1152 стб. с илл.
2	Gellerstedt S. Doctoral Thesis, 1935; Jbuch Fortschritte Math. 61, 1259.
3	Rassias J.M. Lecture Notes on Mixed Type Partial Differential Equations. World Scientific, 1990, 144 p.
4	Трикоми Ф.Д., Лекции по уравнениям в частных производных, пер. с итал., М.:Изд-во иностранной литературы, 1957, 446 с.
5	Смирнов М. М. Уравнения смешанного типа. М., 1970. 296 с.
6	Smirnov M.M. Equations of mixed type. American Translation of the Mathematical monographs. Vol. 51. Mathematical Soc., 31 Dec 1978. p: 232.
7	Моисеев Е.И., Таранов Н.О. Решение одной задачи Геллерстедта для уравнения Лаврентьева-Бицадзе // Дифференциальные уравнения. 2009. Т. 45, № 4. С. 543–548.
8	Моисеев Е.И., Таранов Н.О. Интегральное представление решения одной задачи Геллерстедта // Дифференциальные уравнения. 2009. Т. 45, № 11. С. 1554–1559.
9	Моисеев Е.И., Лихоманенко Т.Н. Об одной нелокальной краевой задаче для уравнения Лаврентьева-Бицадзе // Доклады Академии наук. 2012. Т. 446, № 3. С. 256–258.
10	Moiseev E.I., Nefedov P.V. Tricomi problem for the lavrent'ev–bitsadze equation in a 3d domain // Integral Transforms and Special Functions. 2012. Vol. 23, no. 10. P. 761–768.
11	Moiseev E.I., Nefedov P.V. Gellerstedt problem for the Lavrent'ev–Bitsadze equation in a 3D-domain // Integral Transforms and Special Functions, V. 25, Issue 7, 2014 DOI: 10.1080/10652469.2014.883618.
12	Моисеев Е.И., Холомеева А.А., Нефедов П.В. Аналоги задач Трикоми и Франкля в трехмерных областях для уравнения Лаврентьева-Бицадзе // Дифференциальные уравнения, том 50, № 12, 2014. С. 1672-1675. DOI: 10.1134/S0374064114120127
13	Moiseev E.I., Nefedov P.V., Kholomeeva A.A. Analog of the Gellerstedt problem for the Lavrent'ev-Bitsadze equation in a 3D domain Differential Equations // Differential Equations издательство МАИК Nauka/Interpereodika Publishing (Russian Federation), 2015. Т. 51, № 6. С. 827-829.
14	Moiseev E.I., Moiseev T.E., Vafodorova G.O. On an Integral Representation of Neumann-Tricomi Problem for the Lavrent'ev-Bitsadze Equation // Differential Equations, МАИК (Russian Federation), 2015, Т. 51, № 8, С. 1065-1071.
15	Моисеев Е.И., Лихоманенко Т.Н. Собственные функции задачи Трикоми с наклонной линией изменения типа // Дифференциальные уравнения. 2016. Т. 52, № 10, С. 1375-1382.
16	Zarubin A.N., Kholomeeva A.A. Tricomi problem for an advance-delay equation of mixed type with variable deviation of the argument // Differential Equations. 2016. Vol. 52, no. 10. P. 1312–1322.
17	Моисеев Е.И., Моисеев Т.Е., Холомеева А.А. О Разрешимости задачи Геллерстедта с данными на параллельных характеристиках // Дифференциальные уравнения. 2017. Т. 53, № 10. С. 1379–1384.
18	Moiseev E.I., Likhomanenko T.N. Eigenfunctions of the gellerstedt problem with an inclined-type change line // Integral Transforms and Special Functions. 2017. Vol. 28, no. 4. P. 328–335.
19	Moiseev E.I., Likhomanenko T.N. Eigenfunctions of the tricomi problem with an inclined type change line // Differential Equations. 2016. Vol. 52, no. 10. P. 1323–1330.
20	Moiseev E.I., Gulyaev D.A. The completeness of the eigenfunctions of the tricomi problem for the Lavrent'ev–Bitsadze equation with the Frankl gluing condition // Integral Transforms and Special Functions. 2016. Vol. 27, no. 11. P. 893–898.

В Таблицах 2 и 3 содержатся данные, которые входят в соответствующие позиции статей тезауруса (пример – таблица 1). Для краткости изложения они вынесены здесь в отдельные таблицы и не снабжены идентификаторами. Все позиции Таблицы 2 связаны видовыми отношениями с главным термином, дескриптором «Геллерстедта задача», а сам термин «Геллерстедта задача» связан ассоциативными связями с термином-дескриптором «Трикоми уравнение».

Ключевые слова также выступают в роли контекстов формул в поисковом запросе. Например, контекст формулы  $\Delta u = 0$ : «Трикоми задача, обобщение Трикоми задачи, Трикоми метод, задача Коши для уравнения Трикоми, краевая задача для уравнения типа Чаплыгина, краевая задача для уравнения смешанного типа» из таблицы 1 и их английские эквиваленты из таблицы 2. Кроме того, контекстом могут служить ключевые слова из реферативного списка таблицы 3, а также индексы авторов (ORCID и другие). Таким образом, используя связи тезауруса, можно уточнять или расширять поисковый запрос, достигая в результате уменьшение информационного шума.

**Заключение и выводы.** В работе рассмотрены примеры формирования статей информационно-поискового тезауруса для одного из разделов предметной области задач математической физики. Сама предметная область охватывает широкий класс задач, поэтому выбран конкретный раздел, на примере которого можно в дальнейшем распространить онтологическое описание на другие разделы. Данное представление статей тезауруса не противоречит основному математическому ресурсу, математической энциклопедии И.М. Виноградова, а делается в интеграции с ней в рамках одной семантической библиотеки LibMeta [19]. Необходимость дополнения и расширения оцифрованной математической энциклопедии обусловлена тем, что этот цифровой вариант ограничен реферативными ссылками 80-х годов, когда эта энциклопедия была опубликована. Дальнейшая работа предполагается в направлении развития онтологического представления математических знаний в семантических ресурсах и расширения предметной области информационного запроса в соответствии с новыми методами информационных технологий.

#### СПИСОК ЛИТЕРАТУРЫ

1. Барт К., Мицевич А.Т., Петрина А.М., Риглер Ш., Сукачева Е.В., Тош В., Хадзиламбру З.Н. Русско-немецкий информационно-поисковый тезаурус по робототехнике. М.: ВИНТИ. 1991. 318 с.
2. Блюменау Д.И. Проблемы свертывания научной информации. Л.:Наука. 1982.166 с.
3. Гаврилова Т.А. Хорошевский В.Ф. Базы знаний интеллектуальных систем. СПб.: Питер. 2000. 384 с.
4. Глазунов, Н.М., Дриянский, В.М., Комарова, Т.Н., Рыбалко, Ю.А. Тезаурус по теории формальных и алгебраических групп и его реализация на ЕС ЭВМ. АН УССР. Ин-т кибернетики. 1978. 39 с.
5. ГОСТ 7.25.-2001. Тезаурус информационно-поисковый одноязычный: Правила разработки: структура, состав и форма представления // Система стандартов по информации, библиотечному и издательскому делу. Минск: Межгосударственный совет по стандартизации, метрологии и сертификации. 2001. ГОСТ Р 7.0.91-2015 (ИСО 25964-

- 1:2011) Система стандартов по информации, библиотечному и издательскому делу. Тезаурусы для информационного поиска.
6. Лукашевич Н.В., Добров Б.В. Проектирование лингвистических онтологий для информационных систем в широких предметных областях // *Онтология проектирования*. 2015. Т. 5. № 1 (15). С.47-69.
7. Математическая энциклопедия. Гл. ред. И.М. Виноградов. М.: Советская энциклопедия. 1977-1985. Тома 1-5. 1104 с.
8. Мищенко Г.А., Калинина Г.Р., Гладкова Г.И., Губенко Н.Т., Джигирханова А.В. Двухязычный информационно-поисковый тезаурус классов органических соединений (русско-немецкий/немецко-русский). Пер. на немецкий язык Dr. Н.-J.Террег. М.:ВИНИТИ. 1987. 384 с.
9. Моисеев Е.И., Муромский А.А., Тучкова Н.П. Тезаурус информационно-поисковый по предметной области: обыкновенные дифференциальные уравнения. М.:МАКС Пресс. 2005. 116 с.
10. Муромский А.А., Тучкова Н.П. Представление математических понятий в онтологии научных знаний // *Онтология проектирования*. 2019. Т. 9. № 1 (31). С. 50- 69. DOI:10.18287/2223-9537-2019-9-1-50-69.
11. Никитина С.Е. Тезаурус по теоретической и прикладной лингвистике (автоматическая обработка текста). М.:Наука. 1978. 375 с.
12. Пашенко Н.А., Ксенофонтова Е.Б., Скоробогатая В.Ф. Информационно-поисковый тезаурус по информатике. М.: ВИНТИ. 1987. 466 с.
13. Рождественский Ю.В. Словарь терминов: общеобразовательный тезаурус: Общество. Семиотика. Экономика. Культура. Образование. М.: Флинта: Наука. 2002. 112 с.
14. Смирнов М.М. Уравнения смешанного типа. М. 1970. 296 с.
15. Тезаурус научно-технических наук / Под.ред. Ю.И. Шемакина. М.: Воениздат. 1972. 671 с.
16. Шрейдер Ю.А. Об одной модели семантической теории информации // *Проблемы кибернетики*. М.: Наука. 1965. Вып. 18. С. 233-240.
17. Шрейдер Ю.А. Тезаурусы в информатике и теоретической семантике // *Научно-техническая информация*. Сер. 2. 1971. № 3. С. 21–24.
18. ANSI/NISO Z39.19-2005, Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Bethesda. MD: NISO Press. 2005.
19. Ataeva O., Serebryakov V., Tuchkova N. Query Expansion Method Application for Searching in Mathematical Subject Domains // *CEUR Workshop Proceedings*, M. Jeusfeld c/o Redaktion Sun SITE, Informatik V, RWTH Aachen (Germany). 2020. V. 2543. Pp. 38-48. Available at: <http://ceur-ws.org/Vol-2543/rpaper04.pdf> urn:nbn:de:0074-2543-4 (accessed 06.12.2020)
20. Berners-Lee T., Handler J., Lassila O. The Semantic Web // *Scientific American* 2001. V. 284. № 5. P. 28-37.
21. Gomez-Perez A., Corcho O. Ontology languages for the Semantic Web // *IEEE Intelligent Systems*. 2002. Vol. 17. № 1. P. 54-60. DOI: 10.1109/5254.988453
22. ISO 25964-1:2011. Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval / Geneva: International Organization for Standards. 2011.

23. May K.O. *Historiography: A Perspective for Computer Scientists*. Invited address to International Research Conference on the History of Computing. 1976. Los Alamos.
  24. Niles I., Pease A. *Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology* // *Proceedings of the IEEE International Conference on Information and Knowledge Engineering*. 2003. P. 412-416.
  25. Princeton university. *A Lexical Database for English* Available at: <https://wordnet.princeton.edu/> (accessed 06.12.2020)
  26. Roget P.M. *Roget's thesaurus of English words and phrases*. Taylor Anderson (Editor). CreateSpace Independent Publishing Platform. 2017. 660 p.
  27. *Thesaurus.com*. Available at: <https://www.thesaurus.com/> (accessed 06.12.2020)
  28. World Wide Web Consortium (W3C). Available at: <https://www.w3.org> (accessed 06.12.2020)
- 

**UDK 004.89**

**ON SEMANTIC MODEL OF SUBJECT DOMAIN  
"MATHEMATICAL PHYSICS EQUATIONS"**

**Natalia P. Tuchkova**

PhD., senior researcher,

Dorodnicyn Computing Center

of Federal Research Center "Informatics and Control"

of Russian Academy of Science,

e-mail: [natalia\\_tuchkova@mail.ru](mailto:natalia_tuchkova@mail.ru),

ORCID [0000-0001-6518-5817]

Vavilov st. 40, 119333 Moscow, Russia

**Abstract.** The studying is focusing to the relationship of the concepts of the mathematical subject area on the example of the section of mathematical physics equations. A version of the thesaurus article for terms and related equations and formulas is proposed. The peculiarity of such a thesaurus lies in the use of the context of formulas for their additional identification in the subject area. In addition, it is proposed to take into account the indices of authors and articles where thesaurus terms are found. This approach helps to refine the search query and reduce information noise when using the thesaurus in digital bibliographic collections.

**Keywords:** subject domain thesaurus, equations of mixed type, search query.

**References**

1. Bart K., Mitsevich A.T., Petrina A.M., Rigler SH., Sukacheva Ye.V., Tosh V., Khadzilambro Z.N. *Russko-nemetskiy informatsionno-poiskovyy tezaurus po robototekhnike* [Russian-German information retrieval thesaurus on robotics]. M.: VINITI = M.: All-Russian Institute of Scientific and Technical Information. 1991. 318p.
2. Blyumenau D.I. *Problemy svertyvaniya nauchnoy informatsii* [Collapsing scientific information]. L.: Nauka=L.: Science. 1982. 166p.

3. Gavrilova T.A. Khoroshevskiy V.F. Bazy znaniy intellektual'nykh system [Knowledge base of intelligent systems]. SPb.: Piter= SPb.: Piter. 2000. 384p.
4. Glazunov, N.M., Driyanskiy, V.M., Komarova, T.N., Rybalko, YU.A. Tezaurus po teorii formal'nykh i algebraicheskikh grupp i yego realizatsiya na YES EVM [Thesaurus on the theory of formal and algebraic groups and its implementation on an ES computer]. AN USSR. In-t kibernetiki. 1978. 39p.
5. GOST 7.25.-2001. Tezaurus informatsionno-poiskovyy odnoyazychnyy: Pravila razrabotki: struktura, sostav i forma predstavleniya [Monolingual information retrieval thesaurus: Development rules: structure, composition and form of presentation] // Sistema standartov po informatsii, biblioteknomu i izdatel'skomu delu. Minsk: Mezhsudarstvennyy sovet po standartizatsii, metrologii i sertifikatsii = System of standards for information, librarianship and publishing. Minsk: Interstate Council for Standardization, Metrology and Certification. 2001. GOST R 7.0.91-2015 (ISO 25964-1:2011).
6. Lukashovich N.V., Dobrov B.V. Proyektirovaniye lingvisticheskikh ontologiy dlya informatsionnykh sistem v shirokikh predmetnykh oblastiakh [Designing linguistic ontologies for information systems in broad subject areas] // Ontologiya proyektirovaniya= Design Ontology. 2015. V. 5. № 1 (15). P.47-69.
7. Matematicheskaya entsiklopediya. Gl. red. I.M. Vinogradov [Mathematical encyclopedia]. M.: Sovetskaya entsiklopediya =Soviet encyclopedia. 1977-1985. V. 1-5. 1104p.
8. Mishchenko G.A., Kalinina G.R., Gladkova G.I., Gubenko N.T., Dzhigirskhanova A.V. Dvuyazychnyy informatsionno-poiskovyy tezaurus klassov organicheskikh soyedineniy (russko-nemetskiy/nemetsko-russkiy)[ Bilingual information retrieval thesaurus of classes of organic compounds (Russian-German / German-Russian)]/ Per. na nemetskiy yazyk Dr. H.-J.Tepper. M.:VINITI= M.:All-Russian Institute of Scientific and Technical Information. 1987.384p.
9. Moiseyev Ye.I., Muromskiy A.A., Tuchkova N.P. Tezaurus informatsionno-poiskovyy po predmetnoy oblasti: obyknovennyye differentsial'nyye uravneniya [Information retrieval thesaurus in the subject area: ordinary differential equations]. M.:MAKS Press= M.: MAX Press. 2005. 116p.
10. Muromskiy A.A., Tuchkova N.P. Predstavleniye matematicheskikh ponyatiy v ontologii nauchnykh znaniy [Representation of mathematical concepts in the ontology of scientific knowledge] // Ontologiya proyektirovaniya= Ontology of design. 2019.Vol. 9. № 1 (31). Pp. 50-69. doi.org/10.18287/2223-9537-2019-9-1-50-69.
11. Nikitina S.Ye. Tezaurus po teoreticheskoy i prikladnoy lingvistike (avtomaticheskaya obrabotka teksta)[ Theoretical and Applied Linguistics thesaurus (automatic text processing)]. M.:Nauka= M.: Science. 1978. 375p.
12. Pashchenko N.A., Ksenofontova Ye.B., Skorobogataya V.F. Informatsionno-poiskovyy tezaurus po informatike[Information retrieval thesaurus in computer science]. M.: VINITI= M.: All-Russian Institute of Scientific and Technical Information. 1987.466p.
13. Rozhdestvenskiy YU.V. Slovar' terminov: obshcheobrazovatel'nyy tezaurus: Obshchestvo. Semiotika. Ekonomika. Kul'tura. Obrazovaniye [Glossary of terms: general education thesaurus: Society. Semiotics. Economy. Culture. Education]. M.: Flinta: Nauka = M .: Flinta: Science. 2002. 112p.
14. Smirnov M.M. Uravneniya smeshannogo tipa [Mixed type equations]. M. 1970. 296p.

15. Tezaurus nauchno-tekhnicheskikh nauk [Thesaurus of Scientific and Technical Sciences]. M.: Voenizdat = Moscow: Military Publishing. 1972.671p.
16. Shreyder YU.A. Ob odnoy modeli semanticheskoy teorii informatsii [On one model of semantic information theory]// Problemy kibernetiki= Problems of Cybernetics. M.: Nauka = M.: Science. 1965. Issue 18. Pp.233-240.
17. Shreyder YU.A. Tezaurusy v informatike i teoreticheskoy semantike // Nauchno-tekhnicheskaya informatsiya [Thesauri in informatics and theoretical semantics] // Nauchno-tekhnicheskaya informatsiya. Ser. 2. = Scientific and technical information. Ser. 2. 1971. № 3. P.21–24.
18. ANSI/NISO Z39.19-2005, Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Bethesda. MD: NISO Press. 2005.
19. Ataeva O., Serebryakov V., Tuchkova N. Query Expansion Method Application for Searching in Mathematical Subject Domains // CEUR Workshop Proceedings, M. Jeusfeld c/o Redaktion Sun SITE, Informatik V, RWTH Aachen (Germany). 2020. V. 2543. p. 38-48. Available at: <http://ceur-ws.org/Vol-2543/rpaper04.pdf> urn:nbn:de:0074-2543-4 (accessed 06.12.2020)
20. Berners-Lee T., Handler J., Lassila O. The Semantic Web // Scientific American 2001. V.284. № 5. P. 28-37.
21. Gomez-Perez A., Corcho O. Ontology languages for the Semantic Web // IEEE Intelligent Systems. 2002. Vol. 17. № 1. P. 54-60. DOI:10.1109/5254.988453
22. ISO 25964-1:2011. Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval / Geneva: International Organization for Standards. 2011.
23. May K.O. Historiography: A Perspective for Computer Scientists. Invited address to International Research Conference on the History of Computing. 1976. Los Alamos.
24. Niles I., Pease A. Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology // Proceedings of the IEEE International Conference on Information and Knowledge Engineering. 2003. P. 412-416.
25. Prinseton university. A Lexical Database for English Available at: <https://wordnet.princeton.edu/> (accessed 06.12.2020)
26. Roget P.M. Roget's thesaurus of English words and phrases. Taylor Anderson (Editor). CreateSpace Independent Publishing Platform. 2017. 660 p.
27. Thesaurus.com. Available at: <https://www.thesaurus.com/> (accessed 06.12.2020)
28. World Wide Web Consortium (W3C). Available at: <https://www.w3.org> (accessed 06.12.2020)

УДК 004.89

**СЕМАНТИЧЕСКИЕ ТЕХНОЛОГИИ УПРАВЛЕНИЯ ЗНАНИЯМИ ДЛЯ  
ПОДДЕРЖКИ СОВМЕСТНЫХ ЭНЕРГЕТИЧЕСКИХ И СОЦИО-ЭКОЛОГИЧЕСКИХ  
ИССЛЕДОВАНИЙ**

**Ворожцова Татьяна Николаевна**

к.т.н., вед. инж. отдела «Системы искусственного интеллекта в энергетике»,

e-mail: [tnn@isem.irk.ru](mailto:tnn@isem.irk.ru),

**Пестерев Дмитрий Вячеславович**

инж. отдела «Системы искусственного интеллекта в энергетике»,

e-mail: [pesterev.dmitriy@gmail.com](mailto:pesterev.dmitriy@gmail.com),

**Ан Глеб Владимирович**

асп. отдела «Системы искусственного интеллекта в энергетике»,

e-mail: [godblessya95@gmail.com](mailto:godblessya95@gmail.com),

Институт систем энергетики им. Л.А. Мелентьева СО РАН,

664130 г. Иркутск, ул. Лермонтова 130.

**Аннотация.** В статье рассматриваются возможности применения семантических технологий при разработке интеллектуальной информационной системы для поддержки энергетических и социо-экологических исследований. Использование предлагаемых технологий для управления знаниями дает возможность интеграции исследований разных предметных областей. К семантическим технологиям авторы относят традиционное онтологическое представление знаний и когнитивное моделирование, позволяющее формализовать процессы принятия экспертных решений. Использование онтологий предлагается для описания знаний предметных областей междисциплинарных исследований (энергетических и социо-экологических), связанных с оценкой влияния энергетики на качество жизни. Онтологии дают возможность согласовать терминологию между исследователями и используются для описания и структурирования знаний. Когнитивное моделирование применяется для поддержки исследователей в процессе принятия решений, выявления и описания причинно-следственных связей между основными понятиями предметной области, называемыми концептами, их наглядного представления в процессе описания возможных ситуаций. В работе описываются компоненты онтологического пространства знаний для выполнения совместных энергетических и социо-экологических исследований влияния энергетики на экологию и качество жизни. Приводятся примеры онтологии, когнитивной карты и варианта разработанной модели данных для проектирования базы данных исследований.

**Ключевые слова:** семантическое моделирование, когнитивное моделирование, онтология, онтологическое пространство, когнитивная карта, качество жизни, поддержка принятия решений, база данных.

**Цитирование:** Ворожцова Т. Н., Пестерев Д.В., Ан Г.В. Семантические технологии управления знаниями для поддержки совместных энергетических и социо-экологических исследований // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С.143-157. DOI: 10.38028/ESI.2020.20.4.013

**Введение.** В ИСЭМ СО РАН выполняется проект «Методы построения онтологического пространства знаний для интеллектуальной поддержки принятия решений в энергетике и экологии с учетом качества жизни», поддержанный грантом РФФИ № 20-07-00195. Актуальность данной работы обусловлена необходимостью обеспечить поддержку принятия согласованных решений в области энергетики с учетом ее влияния на экологию и качество жизни. Для этого разрабатываются и используются такие типы семантического моделирования, как когнитивное и онтологическое моделирование. В рамках данного проекта предполагается разработка интеллектуальной информационной системы (ИИС), которая интегрирует сложную информационную базу для выполнения исследований, методы математического и семантического моделирования, инструментальные средства оценки влияния энергетики на экологию и качество жизни. Наличие и функционирование объектов энергетики на той или иной территории влияет на качество жизни населения, так как, с одной стороны, обеспечивает потребности в необходимой тепловой и электрической энергии, а с другой стороны, является одним из источников загрязнений окружающей среды. Разрабатывается единое онтологическое пространство знаний, обеспечивающее интеграцию исследований предметных областей энергетики и экологии. Система онтологий обеспечивает информационную поддержку исследований путем интеграции необходимых данных, проектирования и разработки баз данных на основе онтологического описания. Использование методов онтологического и когнитивного моделирования и поддерживающих их инструментальных средств предлагается для согласования энергетических и социо-экологических исследований, а именно, для научного обоснования принятия решений в энергетике с учетом ее влияния на экологию и качество жизни.

**Возможности и цели онтологического моделирования.** Для интеллектуальной поддержки междисциплинарных исследований с учетом пересечения знаний разных предметных областей предлагается использовать семантические технологии для описания и моделирования знаний. Моделирование знаний о предметных областях – главное направление и базовая парадигма искусственного интеллекта. В настоящее время онтологический подход является основным в разработке и внедрении информационных систем и систем управления знаниями [7,10,11]. Онтологическое моделирование – одно из ведущих направлений семантического моделирования [21, 23]. Вопросы применения онтологического моделирования в исследованиях энергетики рассматривались в работах [5, 15, 16]. При этом авторами используются онтологический инжиниринг, как основной метод работы со знаниями, и фрактальный подход для структурирования знаний.

Онтологический инжиниринг – это процесс проектирования и разработки онтологий на основе структурного анализа предметной области [7].

Знания о предметной области представляют собой совокупность сведений об объектах этой предметной области, свойствах этих объектов, отношениях между объектами, а также о процессах и ситуациях, происходящих в данной предметной области. Процесс онтологического инжиниринга включает следующие шаги:

- выделение базовых понятий данной предметной области – концептов,
- построение связей или отношений между концептами,
- выстраивание иерархии концептов,
- описание свойств выбранных понятий.

Основным преимуществом онтологического инжиниринга является то, что при этом достигаются:

- системность – онтология представляет целостный взгляд на предметную область;
- единообразие – материал представляется в единой форме для лучшего восприятия;
- научность – построение онтологии позволяет восстановить недостающие логические связи во всей их полноте.

Возможности применения предложенного Массель Л.В. фрактального подхода к структурированию знаний и построению онтологического пространства описаны в [17]. При использовании фрактального подхода предполагается, что знания об изучаемом объекте представляются в виде совокупности (множества) слоев, каждый из которых характеризует определенный аспект данных, информации или знаний об объекте.

Онтологические модели дают следующие возможности работы со знаниями:

- являются средством представления знаний;
- обеспечивают работу со смыслом информации;
- допускают автоматизированную обработку;
- обеспечивают интеграцию приложений;
- обеспечивают возможность логических выводов.

В данной работе система онтологических моделей предназначена для:

- согласования исследований разных предметных областей (энергетики, экологии, качества жизни) в соответствии с целями проекта,
- согласования понятий этих предметных областей,
- обеспечения доступности и восприятия больших объемов сложной и/или слабоструктурированной информации,
- описания структуры и компонентов ИИС.

**Компоненты онтологического пространства знаний.** Система онтологий разрабатываемого онтологического пространства знаний для исследований обеспечивает представление знаний для взаимосвязанных исследований энергетики, экологии и качества жизни.

Онтологическое пространство знаний базируется на фрактальном подходе, предполагающем введение нескольких слоев – метауровней, и их дальнейшее расслоение, предусматривающее все большую степень детализации на каждом следующем уровне.

Метаонтологии включают базовые понятия соответствующих предметных областей энергетики, экологии и качества жизни, имеющие отношение к совместным исследованиям. Они используются как основа для описания базовых компонентов и структуры ИИС, а также для разработки ее интерфейса.

Выделены следующие уровни (слои) онтологий:

- Онтологическое описание предметных областей исследований – энергетика, экология, качество жизни.
- Онтологии, описывающие используемые данные и информацию, онтологии баз данных (БД).
- Онтологическое описание баз знаний, содержащих не только описание классов и экземпляров, их свойств и взаимосвязей, но и набор правил, обеспечивающих возможность логического вывода.

На рисунке 1 приведена метаонтология, объединяющая базовые понятия предметных областей и отражающая основные разделы исследований – энергетика, экология и качество жизни. Представленные понятия являются базовыми классами системы онтологий. Компоненты ИИС соответствуют этим разделам. Раздел «Энергетика» содержит описание объектов энергетики и их свойств, от которых зависит антропогенное влияние на природную среду и качество жизни. В разделе «Экология» описываются антропогенные факторы, элементы природной среды и методики оценки антропогенного влияния. Понятие «Качество жизни» – это совокупная характеристика уровня и объективных и субъективных условий жизни населения, определяющих физическое, ментальное, социально-культурное развитие человека, группы или сообщества людей [4]. По определению Всемирной организации здравоохранения (ВОЗ) качество жизни – это восприятие индивидами их положения в жизни в контексте культуры и системы ценностей, в которых они живут, в соответствии с целями, ожиданиями, нормами и потребностями. Соответственно раздел «Качество жизни» включает описание индикаторов качества жизни, факторов влияния и методов анализа этого влияния. Раздел «Задача» содержит описание решаемых в данных исследованиях задач. Раздел «Методология», как компонент ИИС, включает совокупность методических подходов в исследованиях, методов и алгоритмов, используемых для оценки влияния объектов энергетики на экологию и качество жизни.

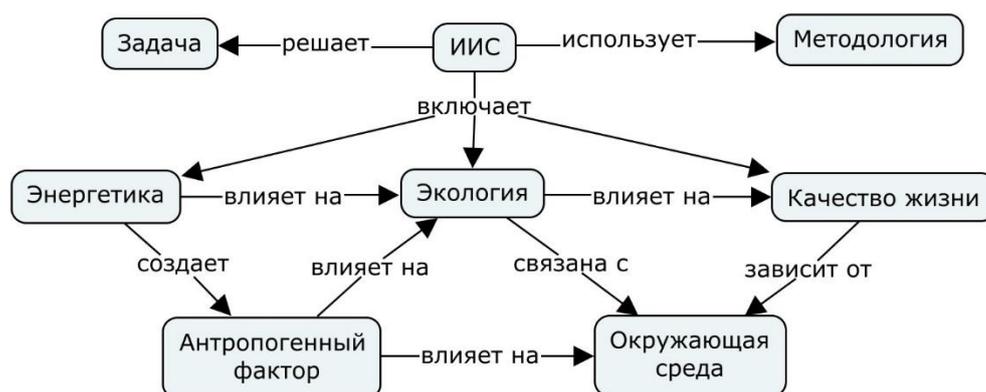


Рис. 1. Метаонтология предметной области исследований

Одним из важных концептов представленной метаонтологии является «Антропогенный фактор» – причина антропогенного воздействия на окружающую среду, обусловленная процессом и условиями функционирования объекта, его характерными особенностями. Применительно к энергетическим объектам под антропогенными факторами понимаются эмиссии, отходы, излучение, шум, вибрация, радиация и т.п. Антропогенный фактор зависит от типа и вида энергоресурса и типа технологий энергетического объекта. Через показатели антропогенного влияния на окружающую среду экспертно оценивается степень отрицательного влияния на качество жизни населения.

**Когнитивное моделирование и его применение в исследованиях.** Когнитивный анализ и моделирование являются современными интеллектуальными технологиями и используются для исследования слабоформализованных и слабоструктурированных систем, к которым относятся экономические, социальные, экологические системы [2, 20].

В настоящее время применение когнитивного моделирования развивается в направлении моделирования и анализа ситуаций и принятия решений. За рубежом наиболее

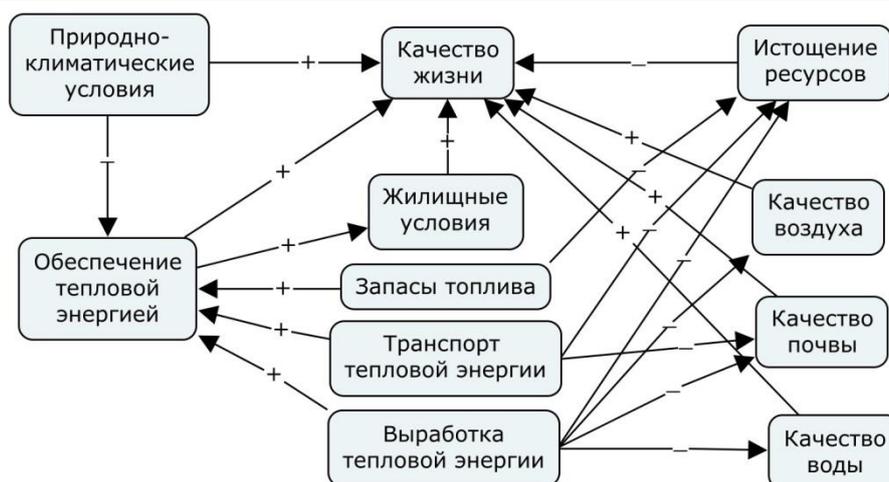
ярко это отражено в работах П. Грумпоса и Х. Стайлиоса [22]. В нашей стране это научная школа Кулинича [12, 13], а также работы сотрудников Института Проблем Управления РАН [1] и др. В ИСЭМ СО РАН предложено использовать когнитивное моделирование как инструмент для исследования энергетической безопасности (ЭБ) для изучения связей между основными составляющими проблемы ЭБ и получения выводов об их взаимовлиянии [14, 16].

В основе когнитивного анализа лежит познавательная-целевая (когнитивная) структуризация знаний об объекте и внешней для него среды. Когнитивная структуризация – это выявление наиболее существенных факторов, влияющих на ситуацию, и причинно-следственных связей между ними. Методология когнитивного моделирования основана на моделировании субъективных представлений экспертов о ситуации. Знания экспертов графически представляются в виде когнитивной карты.

В работе предлагается использовать когнитивное моделирование для анализа факторов взаимовлияния энергетики и экологии на качество жизни населения. Использование онтологий в данном случае необходимо для выявления этих факторов и формирования списка концептов. В онтологическом пространстве, организуемом для изучаемого междисциплинарного взаимодействия, существуют факторы одной предметной области (ПрО), которые могут как положительно, так и отрицательно влиять на факторы другой (ПрО). Например, факторы из области энергетики оказывают как положительное влияние на качество жизни (использование электроэнергии, использование тепловой энергии), так и отрицательное (загрязнение окружающей среды, шумовое воздействие и пр.). Поэтому основная задача когнитивного моделирования – установить взаимовлияние факторов в той или иной конкретной ситуации. Другая задача – визуализация зависимостей между основными факторами исследуемой ПрО.

Для построения когнитивной карты выполняется анализ предметной области, выделяются основные факторы (концепты), которые имеют значения для исследования. Если уже существует онтология предметной области, то концепты выделяются непосредственно на основе этой онтологии. После этого этапа между концептами устанавливаются связи и их характер. В самом простом случае связи могут быть либо положительными, либо отрицательными. В более сложном случае к знакам добавляются весовые коэффициенты, получаемые методом экспертных оценок. При анализе конкретной ситуации когнитивная карта строится либо заново, либо выбирается из числа уже существующих, при условии, что содержит необходимые концепты предметной области.

На рисунке 2 показан пример когнитивной карты, которая описывает основные факторы ситуации «Обеспечение населения тепловой энергией» с позиции основного фактора «Качество жизни».



**Рис. 2.** Когнитивная карта «Обеспечение населения тепловой энергией».

На этой карте видно влияние факторов, оказываемое на целевой фактор «Качество жизни». Это влияние как положительно, так и отрицательно. Чтобы точно установить суммарное влияние, потребуется введение весовых коэффициентов, которые определяются методом экспертных оценок.

**Онтологии и модель данных.** Создание информационных систем с использованием онтологического подхода обеспечивает значительное преимущество с точки зрения корректной подготовки информационных компонентов [19]. Использование семантических и онтологических стандартов помогает наладить обмен и сопоставление данных, выявление коллизий и согласование противоречий, также онтологическая модель является удобным базисом для разработки модели данных [6].

Как правило, проектирование базы данных состоит из последовательных шагов, результатом выполнения которых является создание готовых моделей данных [8]. Процесс проектирования БД состоит из следующих этапов:

- 1) проектирование на уровне концепций – определение общей концепции БД с учетом включения важнейших сущностей (таблиц) и связей между ними, но не зависящей от типа СУБД (иерархической, сетевой, реляционной и т.д.) и физической реализации (целевой СУБД); при использовании семантического подхода создается онтологическая модель;
- 2) логическое проектирование – развитие концептуальной модели данных с учетом типа выбранной СУБД (иерархической, сетевой, реляционной и т.д.);
- 3) проектирование физической БД – конкретизация логической модели данных с учетом выбранной целевой СУБД.

Разрабатываемый программный компонент преобразования онтологии в модель данных является важнейшим компонентом реализации идеи создания интеллектуальной информационной системы на основе онтологий. При анализе онтологических моделей выделяют важные для предметной области концепты и отношения, которые в конечном итоге используются для построения модели типа «сущность-связь» [8]. Для корректного построения модели данных при анализе онтологии выделяют следующие элементы [3]:

- сущности – выделенный набор (класс) однотипных концептов, обобщенных под каким-то одним общим концептом;

- связи – отношения между концептами (связи), кроме случаев, когда связь является «обобщающей» (такие связи правильнее использовать для построения сущности, т.е. таблицы);
- свойства (атрибуты) – дополнительные характеристики концепта, которые в онтологической модели детализируются с помощью других концептов (концептов-характеристик). Основная особенность концептов-характеристик в том, что они предполагают некоторый описательный вид, вариативны – могут принимать различные строковые значения.

В качестве примера преобразования онтологической модели в реляционную базу данных (предполагается, что этап концептуального проектирования выполняется экспертом предметной области) выбрана онтология видов энергоисточников (рисунок 3).

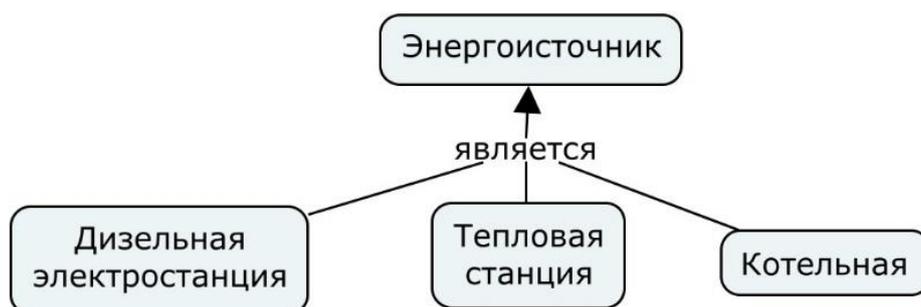


Рис. 3. Онтология «Энергоисточники»

При анализе онтологии можно сделать вывод, что концепт «Энергоисточник» является сущностью (обобщением) объектов «Дизельная э/станция», «Тепловая э/станция», «Котельная». Преобразовав исходную онтологическую модель, получим инфологическую модель данных, представленную на рисунке 4.

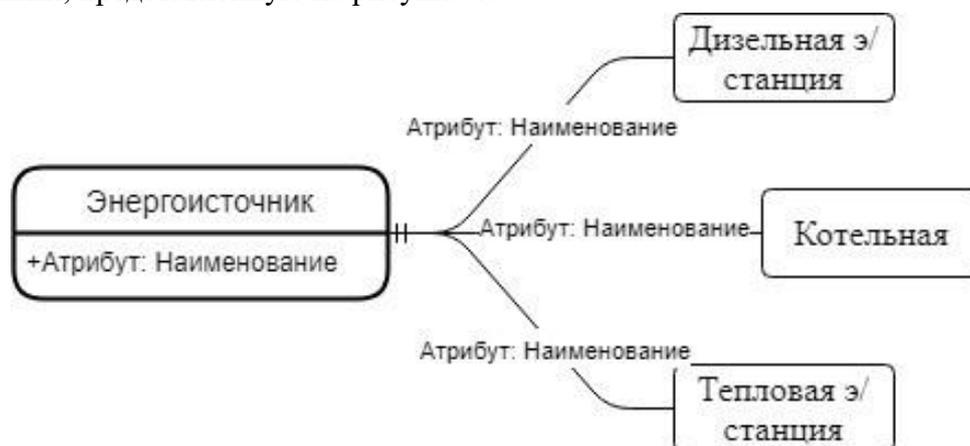


Рис. 4. Инфологическая модель данных

В спроектированной инфологической модели можно выделить следующие элементы: сущность (отношение) «Энергоисточник» и элементы (значения) с наименованиями дизельная/котельная/тепловая станция, которые являются частью спроектированного отношения (данными будущей таблицы) и будут записаны в физическую таблицу под атрибутом «Наименование». Так как в исходной онтологической модели отсутствовали какие-либо дополнительные концепты-характеристики, то нет необходимости в расширении полученной сущности-таблицы.

Завершающий этап проектирования БД – это создание физической модели данных. Физическое проектирование осуществляется путем создания новой базы с именем, соответствующем названию онтологии, в нашем примере БД получит имя «Описание видов энергоисточников». При преобразовании сущности «Энергоисточник» в физическую таблицу БД необходимо сделать уникальной каждую запись в таблице с помощью добавления уникального атрибута таблицы (первичного ключа) `id_station`. На рисунке 5 показана спроектированная таблица данных.

Пример спроектированной таблицы Station на основе онтологии		Пример заполненной таблицы Station на основе онтологии		
Station (Энергоисточник)		id_station	name	
PK	id_station	1	Тепловая э/станция	
	name	2	Дизельная э/станция	
		3	Котельная	

**Рис. 5.** Пример спроектированной в БД таблицы «Энергоисточник»

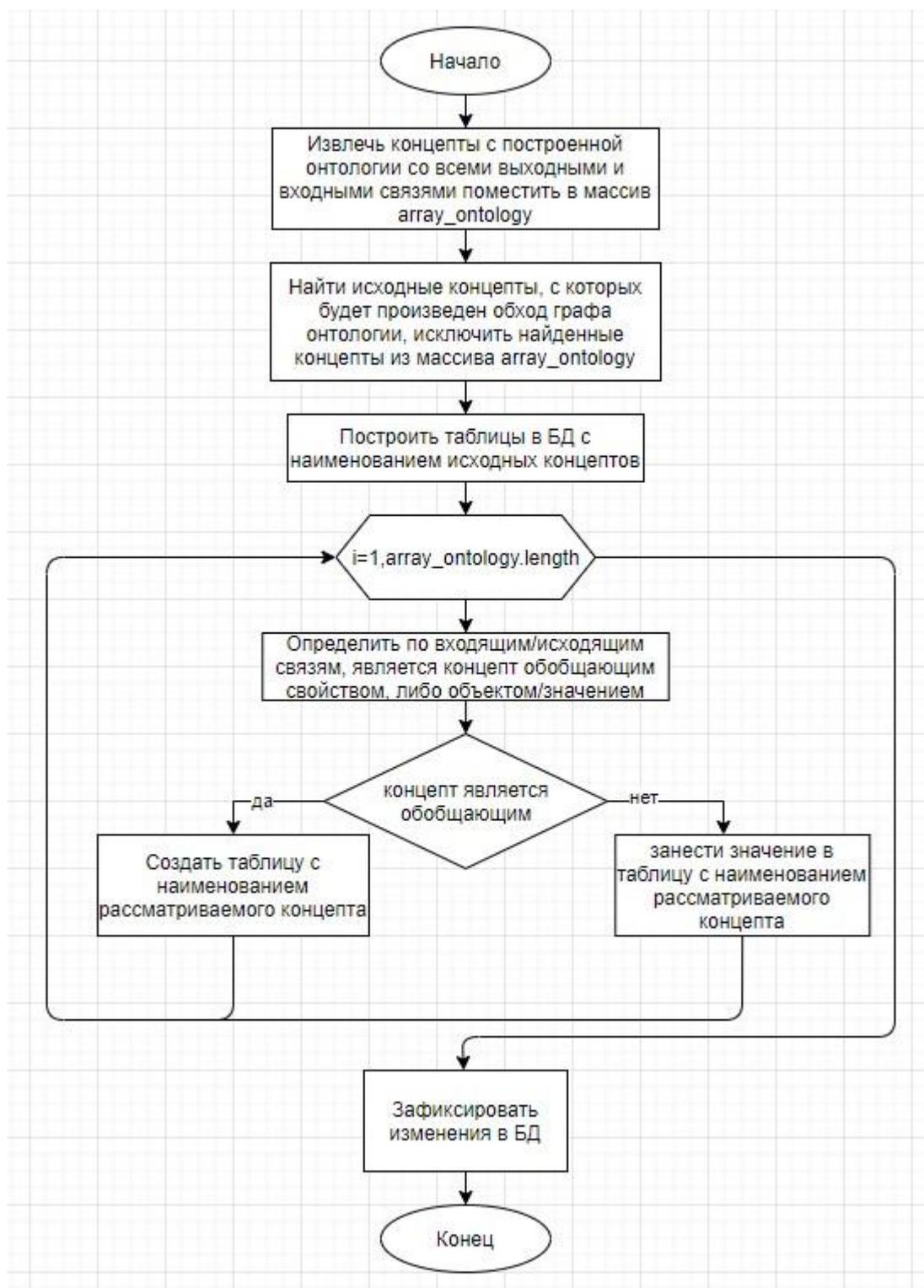
Таким образом, выполнив все этапы проектирования (концептуальное, инфологическое, физическое), получаем готовую базу данных, которая будет использована при реализации интеллектуальной информационной системы.

На основании описанных особенностей проектирования БД авторами построена блок-схема алгоритма преобразования онтологии в модель данных, показанная на рисунке 6.

На рисунке 7 представлен полный вид физической модели данных для хранения информации по объекту энергетики – котельной. Поскольку онтология использует атрибуты естественного языка, то для ее согласования с моделью данных необходимо ограничить набор допустимых связей между концептами («имеет», «является», «измеряется в», «может принимать значение», «состоит из»).

В дальнейшем на основе построенной модели данных будет сформирована информационная система, которая будет выполнять роль информационной поддержки экспертам при принятии решений в энергетике и экологии.

**Заключение.** В работе описан пример применения семантических технологий управления знаниями [18] для выполнения энергетических и социо-экологических исследований. Для этого разрабатывается интеллектуальная информационная система, которая интегрирует сложную методологическую и информационную базу междисциплинарных исследований. Для обеспечения поддержки принятия согласованных решений в области энергетики с учетом ее влияния на экологию используются методы семантического моделирования, включающие онтологическое и когнитивное моделирование. Разрабатывается единое онтологическое пространство знаний, обеспечивающее интеграцию исследований ПО энергетики и экологии.



**Рис. 6.** Алгоритм преобразования онтологии в модель данных

Разрабатываемая система онтологий обеспечивает информационную поддержку исследований путем интеграции необходимых данных, проектирования и разработки баз данных на основе онтологического описания. Онтологии используются для выявления и согласования базовых понятий разных предметных областей, имеющих отношение к совместным исследованиям, установления взаимосвязей между ними, а также для структурного представления знаний.

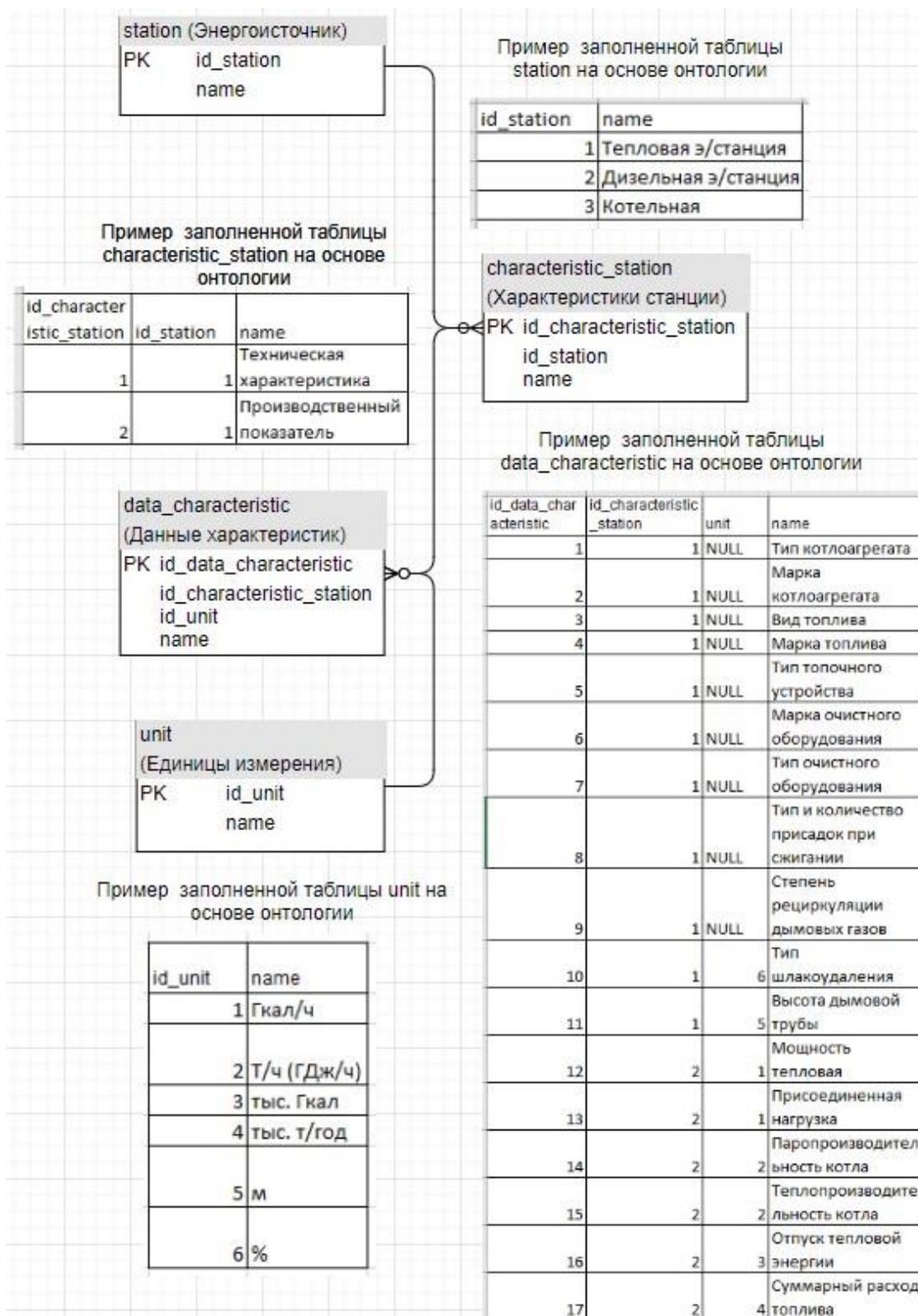


Рис. 7. Физическая модель данных котельной

Когнитивное моделирование применяется для анализа и моделирования ситуаций и принятия согласованных решений в энергетических исследованиях с учетом влияния энергетики на экологию и качество жизни. Когнитивные модели используются для выявления причинно-следственных связей между концептами, их наглядного представления в процессе описания возможных ситуаций и принятия решений.

**Благодарности.** Результаты получены в рамках выполнения проекта по госзаданию ИСЭМ СО РАН АААА-А17-117030310444-2 (проект №349-2016-0005) и при частичной финансовой поддержке грантов РФФИ № 19-07-00351, № 20-07-00195.

#### СПИСОК ЛИТЕРАТУРЫ

1. Авдеева З.К., Горелова Г.В., Коврига С.В. Когнитивные исследования качества жизни населения города // Труды Седьмого Международного форума по когнитивному моделированию. В 3-х частях: Труды Творческой школы, Труды XX Международной конференции, Труды VII Международной конференции. 2019. С. 205–214.
2. Авдеева З.К., Коврига С.В., Макаренко Д.И. Когнитивное моделирование для решения задач управления слабоструктурированными системами (ситуациями) // Управление большими системами. 2007. №16. С. 26–39.
3. Баркер Р. CASE\*Method. Моделирование взаимосвязей между сущностями / М.. 1992. 233 с.
4. Большая российская энциклопедия, Режим доступа: <https://bigenc.ru/vocabulary> (дата обращения 8.10.2020).
5. Ворожцова Т.Н. Онтологическая модель пространства знаний для ситуационного управления в энергетике // Труды XX Байкальской Всероссийской конференции «Информационные и математические технологии в науке и управлении». Т.3 Иркутск: ИСЭМ СО РАН. 2015. С. 85-88.
6. Ворожцова Т.Н., Макагонова Н.Н., Массель Л.В. Онтологический подход к проектированию базы данных для оценки влияния энергетики на окружающую среду // Информационные и математические технологии в науке и управлении. 2019. № 3 (15). С. 31–41. DOI: 10.25729/2413-0133-2019-3-03.
7. Гаврилова Т. А., Кудрявцев Д. В., Муромцев Д. И. Инженерия знаний. Модели и методы. Санкт-Петербург. 324. (2016)
8. Дейт К. Дж. Введение в системы баз данных. 8-е издание. М.: Издательский дом "Вильямс". 2005. 1328 с.
9. Загорулько Ю. А., Боровикова О. И., Загорулько Г. Б. Роль Онтологии в Технологии Построения Тематических Научных Интернет-ресурсов // Всероссийская конференция с международным участием "Знания - Онтологии - Теории" (ЗОНТ-2015). - Новосибирск, Институт математики им. С.Л. Соболева СО РАН. - 2015. - С. 101- 110.
10. Загорулько Ю. А., Загорулько Г. Б., Онтологии и их практическое применение в системах, основанных на знаниях // Всероссийская конференция с международным участием "Знания - Онтологии - Теории" (ЗОНТ-2011). Новосибирск. Институт математики им. С.Л. Соболева СО РАН. 2011. Том 1.
11. Когаловский М.Р., Калиниченко Л.А. Концептуальное и онтологическое моделирование в информационных системах. Программирование. 2009. Т. 35. № 5. С. 3-25.
12. Кузнецов О.П., Кулинич А.А., Марковский А.В. Анализ влияний при управлении слабоструктурированными ситуациями на основе когнитивных карт // Человеческий фактор в управлении. М.: КомКнига. 2006. С. 313-344.

13. Кулинич А.А. Семиотические когнитивные карты. Ч. 1. Когнитивный и семиотический подходы в информатике и управлении // Проблемы управления. 2016. № 1. С. 2–10.
14. Массель А.Г. Когнитивное моделирование в исследованиях проблем энергетической безопасности: применения и перспективы развития. // Труды Международного Конгресса по интеллектуальным системам и информационным технологиям «IS&IT'14». М.: Физматлит. 2014. Т.2. С. 153-158.
15. Массель Л.В., Ворожцова Т.Н., Копайгородский А.Н., Макагонова Н.Н., Скрипкин С.К. Применение онтологий в исследованиях и поддержке принятия решений в энергетике // Всероссийская конференция с международным участием «Знания – Онтологии – Теории» (ЗОНТ-13). Новосибирск: ИМ СО РАН. Том 2. С. 29-38.
16. Массель Л.В. Применение онтологического, когнитивного и событийного моделирования для анализа развития и последствий чрезвычайных ситуаций в энергетике // Проблемы безопасности и чрезвычайных ситуаций. 2010. № 2. С.34–43.
17. Массель Л. В. Фрактальный подход к структурированию знаний и примеры его применения // Онтология проектирования. 2016. №2 (20). С. 149-161.
18. Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии). Томск: Изд-во НТЛ. 2005. 260 с.
19. Черников Б.В. Лексикологический синтез документов в комплексах информационных систем // М: «Форум». 2017. 336 с.
20. Axelrod R. The Structure of Decision: Cognitive Maps of Political Elites. Princeton. University Press. 1976
21. Gruber T.R. Towards principles for the design of ontologies used for knowledge sharing // International Journal of Human-Computer Studies. 1995. V. 43. (5/6). 907 – 928.
22. Groumpos P., Stylios C. Modelling supervisory control systems using fuzzy cognitive maps. Chaos. Solitons & Fractals. 2000. Vol. 11(1-3). Pp.329-336.
23. Guarino N., Formal Ontology and Information Systems // Proceedings of International Conference of Formal Ontology and Information Systems (FOIS'98). Amsterdam: IOS Press. 1998. Pp. 3 – 15.

UDK 004.89

## SEMANTIC TECHNOLOGIES FOR KNOWLEDGE MANAGEMENT TO SUPPORT JOINT ENERGY AND SOCIO-ECOLOGICAL RESEARCH

**Tatiana N.Vorozhtsova**

PhD, Leading engineer of Department “Artificial Intelligence Systems in Energy Sector”

e-mail: [tnn@isem.irk.ru](mailto:tnn@isem.irk.ru)

**Dmitriy.V.Pesterev**

Engineer of Department “Artificial Intelligence Systems in Energy Sector”

e-mail: [pesterev.dmitriy@gmail.com](mailto:pesterev.dmitriy@gmail.com)

**Gleb.V.An**

Engineer of Department “Artificial Intelligence Systems in Energy Sector”

Melentiev Energy Systems Institute SB RAS, Irkutsk, Russia

e-mail: [godblessya95@gmail.com](mailto:godblessya95@gmail.com)

**Abstract.** The article discusses the possibilities of using semantic technologies in the development of an intelligent information system to support energy and socio-environmental research. The use of the proposed knowledge management technology makes it possible to integrate research in different subject areas. The authors refer to the semantic technologies as the traditional ontological representation of knowledge and the cognitive approach, which allows formalizing the processes of making expert decisions. The use of ontologies is proposed to describe the knowledge of subject areas of interdisciplinary energy and socio-ecological research related to assessing the impact of energy on the quality of life. Ontologies make it possible to agree on terminology between researchers and are used to describe and structure knowledge and required information. Cognitive modeling is used to support researchers in the decision-making process, identify and describe cause-and-effect relationships between the main concepts of the subject area, and visualize possible situations. The paper describes the components of the ontological space of knowledge for the implementation of joint energy and socio-ecological studies of the impact of energy on the environment and quality of life. Examples of an ontology, a cognitive map, and a variant of the developed data model for designing a database for research are given.

**Keywords:** semantic modeling, cognitive modeling, ontology, ontological space, cognitive map, quality of life, decision support, database.

### References

1. Avdeeva Z.K., Gorelova G.B., Kovriga S.V. Kognitivnye issledovaniya kachestva zhizni naseleniya goroda [Cognitive studies of the quality of life of the city population] // Trudy Sed'mogo Mezhdunarodnogo foruma po kognitivnomu modelirovaniyu. V 3-h chastyah: Trudy Tvorcheskoj shkoly. Trudy XH Mezhdunarodnoj konferencii. Trudy VII Mezhdunarodnoj konferencii = Proceedings of the Seventh International forum on cognitive modeling. In 3 parts:

- Works of the Creative school. Works of the XX International conference, Works of the VII International conference. 2019. Pp. 205–214. (in Russian)
2. Avdeeva Z.K., Kovriga S.V., Makarenko D.I. Kognitivnoe modelirovanie dlya resheniya zadach upravleniya slabostrukturirovannymi sistemami (situatsiyami) [Cognitive modeling for solving problems of managing weakly structured systems (situations)] // Upravlenie bol'shimi sistemami = Management of large systems. 2007. №16. Pp. 26–39. (in Russian)
  3. Barker R. CASE\*Method. Modelirovanie vzaimosvyazey mezhdru sushchnostyami [CASE\*Method. The modeling of relationships between entities]. M.. 1992. 233 p. (in Russian)
  4. Bol'shaya rossijskaya enciklopediya [Great Russian encyclopedia] Available at: <https://bigenc.ru/vocabulary> (accessed 8.10.2020). (in Russian)
  5. Vorozhzcova T.N. Ontologicheskaya model' prostranstva znaniy dlya situacionnogo upravleniya v energetike [Ontological model of the knowledge space for situational management in the energy sector] // Trudy HKH Bajkal'skoj Vserossijskoj konferencii «Informacionnye i matematicheskie tekhnologii v nauke i upravlenii» = Proceedings of the XX Baikal all-Russian conference "Information and mathematical technologies in science and management». Irkutsk: ISEM SO RAN. 2015. Vol.3. Pp. 85-88. (in Russian)
  6. Vorozhzcova T.N., Makagonova N.N., Massel' L.V. Ontologicheskij podhod k proektirovaniju bazy dannyh dlja ocenki vlijanija jenergetiki na okruzhajushhuyu sredu [Ontological approach to the design of a database for assessing the impact of energy on the environment] // «Informacionnye i matematicheskie tekhnologii v nauke i upravlenii» = Information and mathematical technologies in science and management. 2019. № 3 (15). Pp. 31–41. DOI: 10.25729/2413-0133-2019-3-03. (in Russian)
  7. Gavrilova T. A., Kudryavcev D. V., Muromcev D. I. Inzheneriya znaniy. Modeli i metody [Knowledge engineering. Models and methods]. Sankt-Peterburg. 2016. 324p. (in Russian)
  8. Deyt K. Dzh. Vvedeniye v sistemy baz dannykh [An Introduction to Database Systems] (8th Edition). Izdatel'skiy dom "Vil'yams = M. : Williams Publishing House. 2005. 1328 p.
  9. Zagorul'ko YU. A., Borovikova O. I., Zagorul'ko G. B. Rol' Ontologii v Tekhnologii Postroeniya Tematicheskikh Nauchnyh Internet-resursov [The role of Ontology In the technology Of building Thematic Scientific Internet resources] // Vserossijskaya konferenciya s mezhdunarodnym uchastiem "Znaniya - Ontologii - Teorii" (ZONT-2015) = All-Russian conference with international participation "Knowledge-Ontologies-Theories". Novosibirsk. Institut matematiki im. S.L. Soboleva SO RAN. 2015. Pp. 101- 110 (in Russian)
  10. Zagorul'ko YU. A., Zagorul'ko G. B. Ontologii i ih prakticheskoe primenenie v sistemah, osnovannyh na znaniyah [Ontologies and their practical application in knowledge-based systems] // Vserossijskaya konferenciya s mezhdunarodnym uchastiem "Znaniya - Ontologii - Teorii" (ZONT-2011) = All-Russian conference with international participation "Knowledge-Ontologies-Theories". 2011. Vol. 1. (in Russian)
  11. Kogalovskij M.R., Kalinichenko L.A. Konceptual'noe i ontologicheskoe modelirovanie v informacionnyh sistemah [Conceptual and ontological modeling in information systems] // Programmirovaniye = Programming. 2009. V. 35. № 5. Pp. 3-25 (in Russian).
  12. Kuznecov O.P., Kulinich A.A., Markovskij A.V. Analiz vliyanij pri upravlenii slabostrukturirovannymi situatsiyami na osnove kognitivnyh kart [Analysis of influences in the management of semi-structured situations based on cognitive maps] // CHelovecheskij faktor v upravlenii = Human factor in management. M.: KomKniga. 2006. Pp. 313-344. (in Russian)

13. Kulinich A.A. Semioticheskie kognitivnye karty. CH. 1. Kognitivnyj i semioticheskij podhody v informatike i upravlenii [Semiotic cognitive maps. Part 1. Cognitive and semiotic approaches in computer science and management] // Problemy upravleniya = Management problems. 2016. № 1. Pp. 2–10. (in Russian)
14. Massel' A.G. Kognitivnoe modelirovanie v issledovaniyah problem energeticheskoy bezopasnosti: primeneniya i perspektivy razvitiya. [Cognitive modeling in energy security research: applications and development prospects] // Trudy Mezhdunarodnogo Kongressa po intellektual'nym sistemam i informacionnym tekhnologiyam «IS&IT'14» = Proceedings of The international Congress on intelligent systems and information technologies "IS&IT' 14». M.: Fizmatlit. 2014. V.2. Pp. 153-158. (in Russian)
15. Massel' L.V., Vorozhcova T.N., Kopajgorodskij A.N., Makagonova N.N., Skripkin S.K. Primenenie ontologij v issledovaniyah i podderzhke prinyatiya reshenij v energetike [Application of ontologies in research and decision support in the energy sector] // Vserossijskaya konferenciya s mezhdunarodnym uchastiem «Znaniya – Ontologii – Teorii» (ZONT-13) = All-Russian conference with international participation "Knowledge-Ontologies-Theories". Novosibirsk: IM SO RAN. 2013. V.2. Pp. 29-38. (in Russian)
16. Massel' L.V. Primenenie ontologicheskogo, kognitivnogo i sobytijnogo modelirovaniya dlya analiza razvitiya i posledstvij chrezvychajnyh situacij v energetike [The use of ontological, cognitive and event modeling to analyze the development and consequences of emergency situations in the energy sector] // Problemy bezopasnosti i chrezvychajnyh situacij = Security and emergency issues. 2010. № 2. Pp.34–43. (in Russian)
17. Massel' L. V. Fraktal'nyj podhod k strukturirovaniyu znaniy i primery ego primeneniya [Fractal approach to knowledge structuring and examples of its application] // Ontologiya proektirovaniya = Ontology of design. 2016. №2 (20). Pp. 149-161. (in Russian)
18. Tuzovskij A.F., Chirikov S.V., Jampol'skij V.Z. Sistemy upravleniya znaniyami (metody i tekhnologii) [Knowledge management systems (methods and technologies)]. Tomsk: NTL publ. 2005. 260 p. (in Russian)
19. Chernikov B.V. Leksikologicheskij sintez dokumentov v kompleksah informacionnyh sistem [Lexicological synthesis of documents in information systems complexes] // M: «Forum». 2017. 336 p. (in Russian)
20. Axelrod R. The Structure of Decision: Cognitive Maps of Political Elites. Princeton. University Press. 1976
21. Gruber T.R., Towards principles for the design of ontologies used for knowledge sharing // International Journal of Human-Computer Studies. 1995. V. 43. (5/6). 907 – 928.
22. Groumpos P., Stylios C., Modelling supervisory control systems using fuzzy cognitive maps, Chaos, Solitons & Fractals. 2000. Vol. 11(1-3). 329-336.
23. Guarino N., Formal Ontology and Information Systems // Proceedings of International Conference of Formal Ontology and Information Systems (FOIS'98). Amsterdam: IOS Press. 1998/ Pp. 3 – 15

**АНАЛИЗ БОЛЬШИХ ДАННЫХ  
ДЛЯ ОБОСНОВАНИЯ РЕШЕНИЙ ПО НАУЧНО-ТЕХНОЛОГИЧЕСКОМУ  
РАЗВИТИЮ В ЭНЕРГЕТИКЕ**

**Михеев Алексей Валерьевич**

к.т.н., ведущий специалист

Институт систем энергетики им. Л.А. Мелентьева СО РАН,

664130 г. Иркутск, ул. Лермонтова 130,

e-mail: [mikheev@isem.irk.ru](mailto:mikheev@isem.irk.ru)

**Аннотация.** В статье рассматриваются возможности применения методов анализа больших данных для принятия решений по инновационному развитию в энергетике. Выполнен библиометрический обзор научных исследований по использованию анализа больших данных для задач в сфере энергетики на основе публикаций международной базы Scopus за 2010-2020 гг. Приведены содержательные задачи мониторинга, прогнозирования и оценки перспективности технологических решений в энергетике на основе семантического анализа больших данных.

**Ключевые слова:** анализ больших данных, tech mining, семантический анализ текстов, научно-технологическое прогнозирование, инновационное развитие, инновационные индикаторы.

**Цитирование:** Михеев А.В. Анализ больших данных для обоснования решений по научно-технологическому развитию в энергетике//Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 158-167. DOI: 10.38028/ESI.2020.20.4.014

**Введение.** Одними из самых сложных задач в области принятия решений в энергетике являются стратегические задачи управления развитием на уровне компании, региона или страны через внедрение передовых технологий путем модернизации действующей и создания новой энергетической инфраструктуры. Инновации в энергетике приводят к увеличению эффективности и качества генерации, конверсии, транспорта, хранения и распределения электрической и тепловой энергии, к снижению негативного воздействия на окружающую среду и в целом обеспечивают возможности экономического роста. При этом важно найти и реализовать такие пути трансформации уже существующей технологической и организационной структуры энергетики, которые будут наилучшими с точки зрения критериев технической, экономической, экологической и социально-политической эффективности. Принимаемые решения по развитию энергетических систем носят комплексный характер, отличаются высокой степенью сложности, долгосрочностью планирования и реализации, многокритериальностью выбора, неопределенностью и зависимостью от многочисленных внешних факторов [2]. Поэтому для обеспечения максимальной информированности при обосновании научно-технических рекомендаций, предложений, программ, а также при планировании и внедрении перспективных технологических инноваций в сфере энергетики крайне важно использование всех возможных и доступных данных, включая слабоструктурированные и неструктурированные,

объем которых в мире постоянно растет. Особый интерес для систем поддержки принятия решений представляет применение технологий, лежащих в русле концепции анализа больших данных: Big / Large Data Analysis – BDA или LDA.

Термин «большие данные» уже прочно вошел в оборот исследователей. Технологии анализа больших данных широко используется в качестве информационно-аналитического инструментария для решения задач во многих областях науки и техники, в первую очередь в медицине, социологии, компьютерных и технических науках [6].

Само понятие «большие данные» обычно определяется через 6 ключевых характеристик или 6-V измерений [5]. Большие данные должны обладать:

- 1) значительным объемом (*Volume*), когда величина объема характеризует полноту и репрезентативность данных для целей анализа;
- 2) разнообразием форм и источников данных (*Variety*), что подразумевает вовлечение для анализа не только структурированных, но слабоструктурированных и неструктурированных данных, которые по оценкам составляют до 95% всех полезных данных;
- 3) высокой скоростью (*Velocity*) накопления данных и регулярностью их обработки.

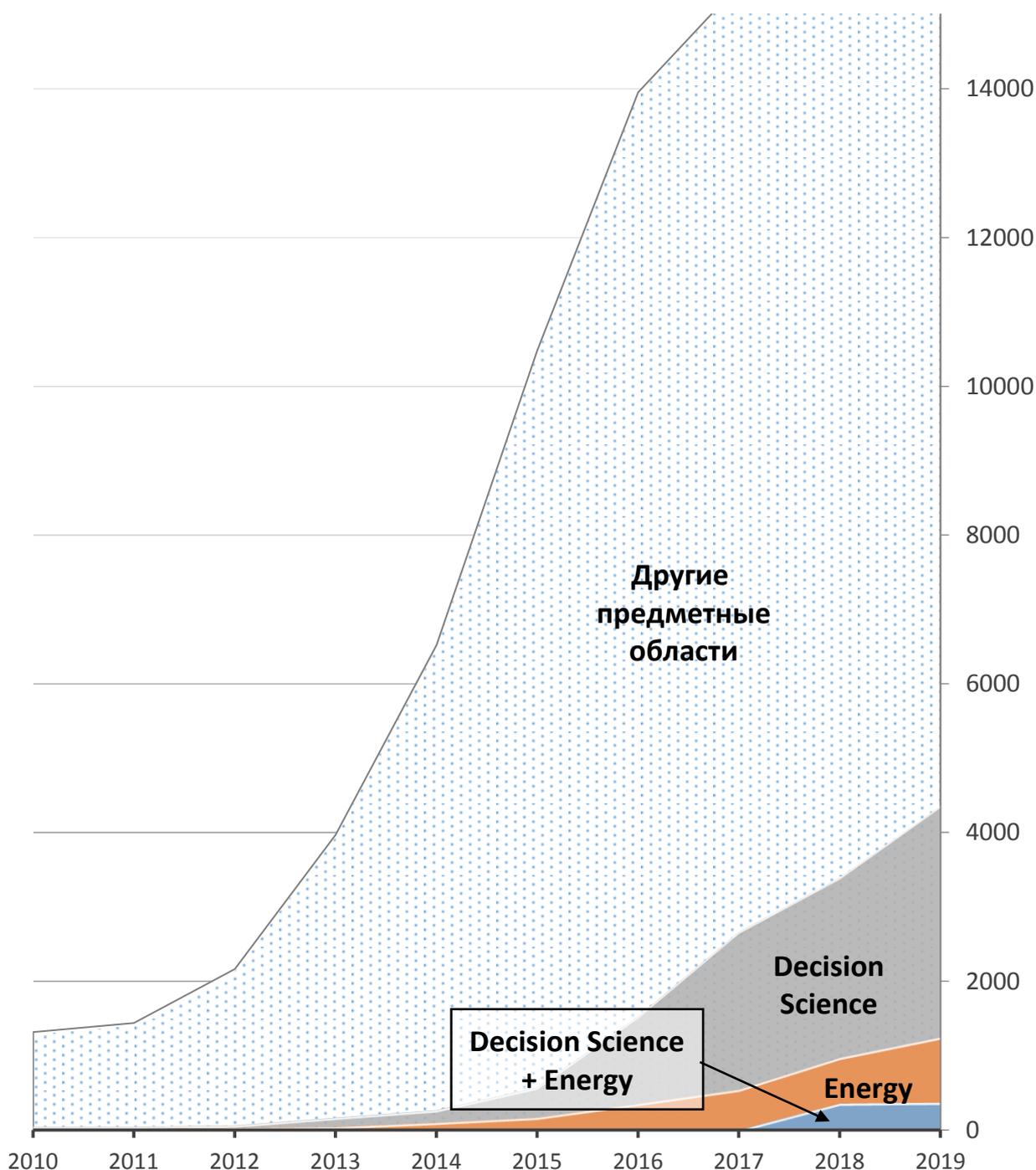
Некоторые исследователи [7] добавляют к этим характеристикам еще такие:

- 4) вариативность (*Variability*) или комплексность данных по времени или иной метрике, выражаемая в непостоянстве информационных потоков и гетерогенности их источников;
- 5) достоверность (*Veracity*): несмотря на ненадежность и неопределенность данных из отдельных источников, только анализ данных по совокупности позволяет получать близкие к достоверным результатам;
- 6) ценность (*Value*): часто данные, полученные в исходном виде или фрагментарно, сами по себе имеют низкую ценность, но высокую ценность можно получить, анализируя большие объемы таких данных.

Используемые методы анализа больших данных – это методы широкого спектра от классических методов математической статистики до методов глубокого обучения и нейронных сетей [8]. Подробное их рассмотрение находится за пределами данной статьи.

О популярности и востребованности аналитики на основе больших данных свидетельствует взрывной рост статей с ключевым выражением «big data» и его аналогами в международных индексах цитирования, начиная со второй декады XXI века (см. рис. 1). Именно к этому моменту в арсенале исследователей и аналитиков появились технологические возможности анализировать большие массивы накопленных и постоянно пополняемых данных, которые обладают максимально возможной информационной полнотой и доступностью. Как показано на диаграмме, с небольшим запозданием концепция больших данных начинает активно проникать в область наук о принятии решений (Decision Science) на рубеже 2014-2015 годов. Примерно в это же время начинается рост количества исследований с использованием анализа больших данных в предметной области энергетика (Energy). Однако, значимое число научных публикаций по методологии принятия решений в энергетике на основе больших данных начинает прослеживаться с достаточно недавнего времени с 2017-2018 гг. (см. область Decision Science + Energy на рис. 1), и этих пересекающихся публикаций относительно немного в общем потоке научных работ, связанных с BDA. Таким образом, текущее состояние, тренды, потенциальные пути

использования и общая полезность концепции больших данных для обоснования решений по развитию энергетических систем актуальны и требуют рассмотрения.



**Рис. 1.** Динамика количества публикаций по объединенному запросу из поисковых фраз «big data», «large data» по всем и по отдельным предметным областям наук - принятие решений (Decision Science), энергетика (Energy) и их пересечение - за 2010-2020 годы в международном индексе цитирования Scopus.

**Краткий обзор применения анализа больших данных для задач энергетики.**

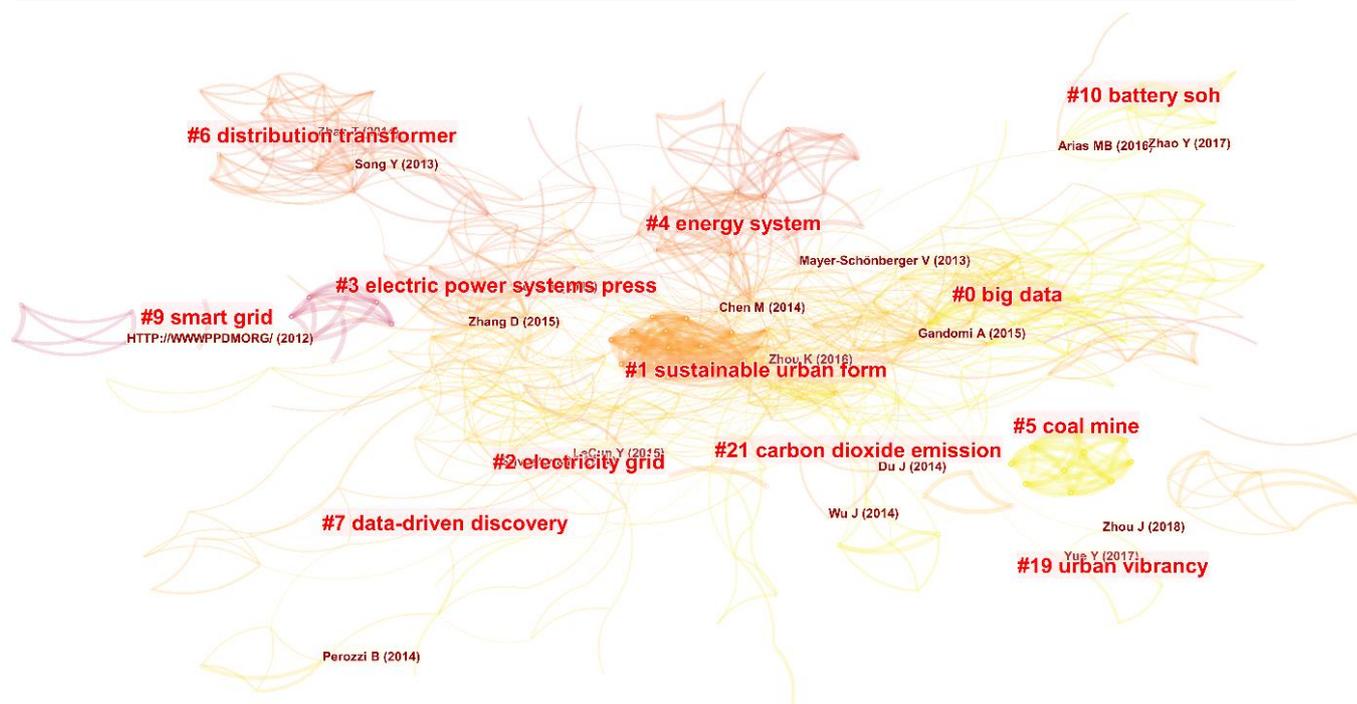
В энергетике существует большая история работы со структурированными большими данными, особенно в электроэнергетике, где давно и успешно внедряются, постоянно совершенствуются автоматизированные системы контроля и учёта энергоресурсов (АСКУЭ). Через глубокую интеграцию электроэнергетических систем с продвинутыми

информационно-телекоммуникационными и интернет-технологиями управления и оснащение современными средствами измерений (устройства Smart Meters, PMU, FACTS и др.) происходит трансформация в интеллектуальные энергосистемы в русле концепции Smart Grid [11]. Энергокомпании используют свои накопленные большие данные для повышения эффективности функционирования систем и управления режимами, например, для прогнозирования и создания предиктивных моделей энергопотребления с использованием регрессии и обученных нейросетевых алгоритмов, мониторинга оборудования электростанций и сетей, и т.п. В перспективе планируется переход к «цифровым двойникам» энергетических систем и объектов, естественным образом погруженных в контекст больших данных [1, 7].

Поиск научных публикаций Big Data в международной базе Scopus по поисковому запросу TITLE-ABS-KEY ( "big data" OR "large data" OR "tech mining" ) AND LIMIT-TO ( SUBJAREA , "ENER" ) за период с 2010 по 2021 годы на начало декабря находит 4108 публикаций (количество публикаций будет увеличиваться по мере поступления новых статей, начиная с 2020 года). С помощью открытого программного средства CiteSpace [4], построив карту цитирований и применив процедуру кластеризации на основе анализа неструктурированных текстов заголовков и абстрактов статей, выделены 10 основных тематических кластеров исследований, как показано на рис. 2. Публикации в кластерах можно обобщить в несколько направлений:

1. Энергетические системы в целом с выделением электроэнергетических систем в русле концепции Smart Grid, систем распределенной генерации и решения проблем безопасности таких систем.
2. Развитие устойчивой экологичной энергетической инфраструктуры и энергоэффективных урбанистических форм будущего для «умных» городов (Smart City) и «умного» управления зданиями.
3. Разработку методов и технологий управления «умными» угольными шахтами.
4. Разработка систем мониторинга и контроля за состоянием электрических батарей как важного компонента гибридных электростанций на основе возобновляемых источников энергии.
5. Общая проблематика применения анализа больших данных для обеспечения устойчивого развития, оптимального управления цепочками продуктовых поставок, решения экологических проблем, разработки методов и алгоритмов, междисциплинарного анализа текстовых источников и т.п.

В целом большинство выделенных содержательных тем BDA лежит в области повышения эффективности работы энергетических систем на базе развивающихся технологий как информационных, так и энергетических, через потоки данных измерений и структурированной научно-технической информации о работе энергетических систем. Причем, порой сложно определить, к какому классу относятся решаемые с помощью методов анализа больших данных задачи – к классу задач управления функционированием энергетических систем или задач управления их развитием.



**Рис. 2.** Тематические кластеры, выделенные на сетевой карте совместного цитирования научных публикаций по Big Data в области энергетики в базе данных Scopus за 2010-2020 годы.

**Анализ больших данных для обоснования решений по научно-технологическому развитию в энергетике.** Практики использования анализа больших массивов данных для обоснования научно-технологического развития появились задолго до широкого распространения этой концепции и даже до эпохи интернета, начиная с традиционного контентного и библиометрического анализа. Одной из форм ВДА безусловно является Tech Mining (ТМ) - анализ текстовых документов из электронных ресурсов научно-технологической направленности и извлечение знаний из них в целях лучшего понимания контекста и трендов развития науки и технологий [10]. В дополнение к стандартным способам (статистические формы, финансовая статистика и т.п.) Tech Mining измеряет разнообразные статистические и семантические параметры прямых результатов исследований и разработок - научных статей, патентов, отчетов, новостных колонок СМИ и других документов. Кратко говоря, Tech Mining отвечает на вопросы «Что?», «Где?» и «Когда?». Подход ТМ не ограничивается статистическим анализом метаданных документов и связанной с ними текстовой информации, но часто включает различные методы и технологии семантического анализа полнотекстовых документов с использованием методов обработки естественного языка (Natural Language Processing) [9].

С другой стороны, в России сложилась комплексная методология системных исследований по обоснованию технологического развития энергетики [3], включающая в себя методы системного анализа, математического моделирования, технико-экономического анализа, анализа рисков и другие применительно к объекту исследования – энергетическим технологиям. Методология позволяет получать воспроизводимые, объясняемые, отражающие реальное положение дел результаты, хотя достаточно сложна в использовании

и требует высокой квалификации ее пользователей. Естественно, вывод результатов любого анализа должен быть подержан качественной экспертизой.

Привнесение концепции анализа больших данных может значительно расширить возможности приложений с использованием данной методологии, а в некоторых случаях упростить ее этапы. В таблице 1 приведен перечень обобщенных задач системных исследований по обоснованию в энергетике, которые можно рассмотреть в русле применения концепции больших данных.

**Таблица 1.** Задачи по обоснованию инновационного развития энергетики, сформированы на основе [3].

Этапы	Основные задачи по обоснованию инновационного развития энергетики
1	Научно-технологический мониторинг в энергетике, выявление научно-технологических тенденций, определение перспективных технологий, глобальных мировых трендов технологического развития
2	Прогнозирование технико-экономических показателей перспективных энергетических технологий, которое связано с определением зрелости технологии, а также скорости обучения. Для энергетических технологий – это как правило, удельная стоимость на единицу установленной мощности и на выработанную единицу энергии, например, USD / кВт, USD / кВт ч.
3	Системная оценка конкурирующих энергетических технологий, сравнительный анализ и выбор приоритетных технологий для развития энергетики, в т.ч. выбор технологий для разработки для последующей реализации собственных проектов НИОКР
4	Оценка системных эффектов (технических, экономических, экологических, социальных) от ожидаемых инноваций: эффекты, влияющие на энергетическую систему в целом, способны кардинально изменить выводы о целесообразности внедрения той или иной технологии.
5	Стратегическое научно-технологическое планирование, разработка дорожных карт развития технологий, формирование политики в сфере инновационного развития
6	Обеспечение выхода технологий на внутренние или внешние рынки, оценка перспектив сотрудничества и трансфера при собственной разработке или импорте технологий, оценка конкурентной среды и т.п.

На рисунке 3 представлена диаграмма взаимосвязей задач из таблицы 1 с содержательными задачами анализа больших данных. К каждой задаче BDA в качестве примеров даны ключевые фразы, кратко описывающие возможные представления полезной информации для последующей оценки экспертами. Диаграмма является предварительным эскизом для функциональной онтологии задач анализа больших данных по обоснованию научно-технологического развития в сфере энергетики.

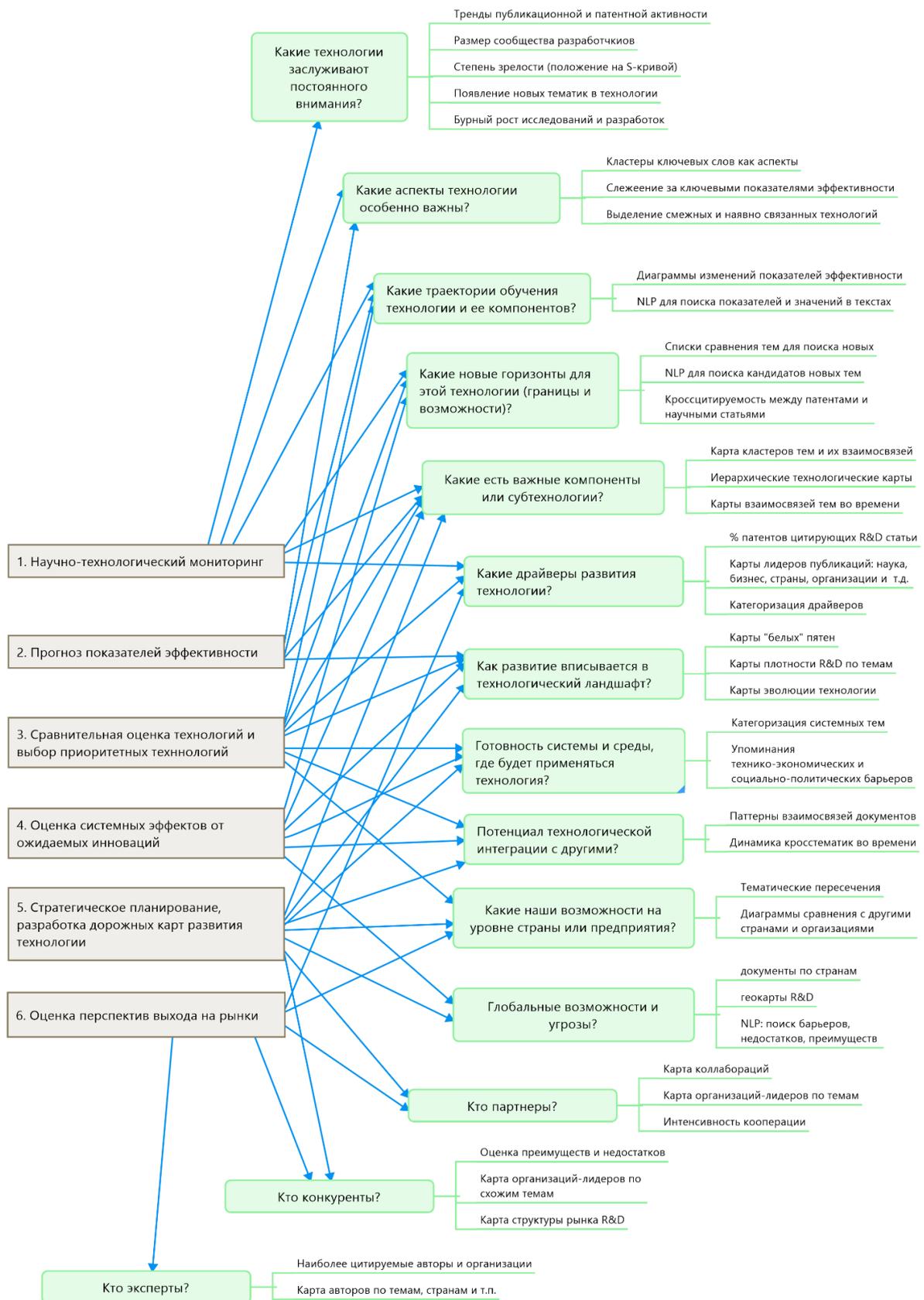


Рис. 3. Диаграмма взаимосвязей задач по обоснованию инновационного развития (серый цвет) с задачами анализа больших данных (зеленый цвет) и примерами результатов.

**Заключение.** Анализ больших данных имеет значительные перспективы использования для задач мониторинга и прогнозирования научно-технологического развития в энергетике. Результаты библиометрического анализа научных публикаций области больших данных показывают существенный рост интереса исследователей к этому направлению в приложениях в энергетике. Выполнено картирование между обобщенными задачами по обоснованию научно-технологического развития в энергетике и задачами семантического анализа больших данных.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (РФФИ), научный проект № 20-07-00994.

#### СПИСОК ЛИТЕРАТУРЫ

1. Массель Л.В., Массель А.Г., Копайгородский А.Н. Эволюция технологий исследований энергетики и применения их результатов: от математических моделей и компьютерных программ к цифровым двойникам и цифровым образам // Информационные и математические технологии в науке и управлении. № 4 (16). 2019. С. 5-19. DOI: 10.25729/2413-0133-2019-4-01
2. Обоснование развития электроэнергетических систем: Методология, модели, методы, их использование / отв. ред. Н.И. Воропай. Новосибирск: "Наука", 2015. 448 с.
3. Филиппов С.П., Дильман М.Д. Системные исследования приоритетов научно-технологического развития энергетики: методологические аспекты // В кн. Системные исследования в энергетике: методология и результаты. Под. ред. Макарова А. А., Воропая Н.И. 2017. С. 63-86.
4. Chen C., Song M. Visualizing a field of research: A methodology of systematic scientometric reviews. PLoS ONE, 14(10). 2019. DOI.org/10.1371/journal.pone.0223994
5. Chen H., R.H.L. Chiang, V.C. Storey Business Intelligence and Analytics: From Big Data To Big Impact. MIS Quarterly. 36(4), 2012. Pp. 1165–1188. <http://www.jstor.org/stable/41703503>
6. Chen M., Mao S., Liu Y. Big data: A survey. *Mobile Networks and Applications*, 19(2), 2014. P. 171–209. DOI.org/10.1007/s11036-013-0489-0
7. Gandomi A., Haider M. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*. 35(2). 2015. Pp. 137–144. DOI.org/10.1016/j.ijinfomgt.2014.10.007
8. Handbook of Big Data Analytics. Eds. Härdle W.K., Lu H. H.-S., Shen X. 2020. 538 pp. DOI.org/10.1007/978-3-319-18284-1
9. Kopyagorodsky A.N. Natural Language Processing for Forecasting Innovative Development of the Energy Infrastructure. E3S Web of Conferences 209. 03015. 2020. DOI.ORG/10.1051/E3SCONF/202020903015
10. Porter A. L., Cunningham S. W. Tech mining: exploiting new technologies for competitive advantage. 2005. 384 p.
11. Voropai N., Efimov D., Kurbatsky V., Tomin N. Smart control in the Russian electric power system // Proc. Int.Conf. of Smart Greens 2012, Porto, Portugal, 2012. Portugal. 19-20 April 2012. P.133-136. ISBN (print): 9789898565.

**BIG DATA ANALYSIS  
FOR INNOVATION DEVELOPMENT DECISION MAKING  
IN ENERGY SECTOR**

**Alexey V. Mikheev**

Ph.D., Leading specialist

Melentiev Energy Systems Institute

Siberian Branch of the Russian Academy of Sciences

130, Lermontov Str., 664033, Irkutsk, Russia, e-mail: [mikheev@isem.irk.ru](mailto:mikheev@isem.irk.ru)

**Annotation.** The article discusses the feasibility and possible applications of big data analysis for making decisions on innovative development in the energy sector. A bibliometric review of scientific research on the use of big data analysis for problems in the energy sector was carried out based on publications of Scopus database for 2010-2020. The substantive tasks of monitoring, forecasting and evaluating the prospects of technological solutions in the energy sector based on semantic analysis of big data are presented.

**Keywords:** big data analysis, tech mining, semantic analysis of texts, scientific and technological forecasting, innovative development.

**References**

1. Massel L.V., Massel A.G., Kopaygorodsky A.N. The evolution of energy research and application of their results: from mathematical models and computer programs to digital twins and digital images // *Informacionnye i matematicheskie tekhnologii v nauke i upravlenii* [Information and Mathematical Technologies in Science and Management]. 4(16). 2019. Pp. 5-19. DOI: 10.25729/2413-0133-2019-4-01 (in Russian)
2. Obosnovanie razvitija jelektrojenergeticheskikh sistem: Metodologija, modeli, metody, ih ispol'zovanie [Feasibility of energy power systems development: Methodology, models, methods and its use] / Eds. Voropai N.I. Novosibirsk: "Nauka". 2015. 448 p. (in Russian).
3. Filippov S.P., Dilman M.D. Sistemnye issledovaniya prioritetov nauchno-tekhnologicheskogo razvitiya energetiki: metodologicheskie aspekty [Systemic studies of the priorities of scientific and technological development of energy: methodological aspects] // In the book. Sistemnye issledovaniya v energetike: metodologiya i rezul'taty [Systems research in energy: methodology and results]. Eds. Makarov, A.A., Voropai. N.I. 2017. Pp. 63-86. (in Russian)
4. Chen C., Song M. Visualizing a field of research: A methodology of systematic scientometric reviews. *PLoS ONE*, 14(10). 2019. DOI.org/10.1371/journal.pone.0223994
5. Chen H., R.H.L. Chiang, V.C. Storey Business Intelligence and Analytics: From Big Data To Big Impact. *MIS Quarterly*. 36(4). 2012. Pp. 1165–1188. <http://www.jstor.org/stable/41703503>
6. Chen M., Mao S., Liu Y. Big data: A survey. *Mobile Networks and Applications*, 19(2), 2014. P. 171–209. DOI.org/10.1007/s11036-013-0489-0
7. Gandomi A., Haider M. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 2015. Pp. 137–144. DOI.org/10.1016/j.ijinfomgt.2014.10.007

8. Handbook of Big Data Analytics. Eds. Härdle W.K., Lu H. H.-S., Shen X. 2020. 538 pp. DOI.org/10.1007/978-3-319-18284-1
9. Kopyaygorodsky A.N. Natural Language Processing for Forecasting Innovative Development of the Energy Infrastructure. E3S Web of Conferences 209. 03015. 2020. DOI.ORG/10.1051/E3SCONF/202020903015
10. Porter A. L., Cunningham S. W. Tech mining: exploiting new technologies for competitive advantage. 2005. 384 p.
11. Voropai N., Efimov D., Kurbatsky V., Tomin N. Smart control in the Russian electric power system // Proc. Int.Conf. of Smart Greens 2012, Porto, Portugal, 2012. Portugal. 19-20 April 2012. P.133-136. ISBN (print): 9789898565.

УДК 004.8:001.8:620.9

## АРХИТЕКТУРА ИНТЕЛЛЕКТУАЛЬНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ ПОДДЕРЖКИ ЭКСПЕРТНЫХ РЕШЕНИЙ ПО СТРАТЕГИЧЕСКОМУ ИННОВАЦИОННОМУ РАЗВИТИЮ ЭНЕРГЕТИКИ

**Копайгородский Алексей Николаевич**

К.т.н., ведущий специалист отдела Систем искусственного интеллекта в энергетике,

e-mail: [kopaygorodsky@isem.irk.ru](mailto:kopaygorodsky@isem.irk.ru)

**Мамедов Тимур Габилевич**

инженер отдела Систем искусственного интеллекта в энергетике,

e-mail: [mamedowtymur@yandex.ru](mailto:mamedowtymur@yandex.ru)

Институт систем энергетики им. Л.А. Мелентьева СО РАН,

664033 г. Иркутск, ул. Лермонтова 130

**Аннотация.** В статье рассмотрены методы построения интеллектуальной информационной системы для поддержки экспертных решений по стратегическому инновационному развитию энергетики. Обоснована необходимость применения методов анализа Больших данных (Big Data). Представлена архитектура интегрированного хранилища интеллектуальной информационной системы, основным компонентом которой является система онтологий, объединяющая данные и знания из различных источников.

**Ключевые слова:** научно-технологическое прогнозирование, семантический анализ, поддержка научных исследований, архитектура информационной системы.

**Цитирование:** Копайгородский А.Н., Мамедов Т.Г. Архитектура интеллектуальной информационной системы для поддержки экспертных решений по стратегическому инновационному развитию энергетики // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 168-176. DOI: 10.38028/ESI.2020.20.4.015

**Введение.** Активное развитие информационно-телекоммуникационных технологий в значительной мере повлияло на сокращение времени выхода новых инновационных разработок на рынок. В первую очередь такой эффект наблюдается благодаря ускорению передачи научных знаний и разработке на их основе новых производственных технологий с последующим выпуском продукции. При этом некоторые вещи, казавшиеся еще не так давно фантастикой, входят в повседневную жизнь, существенно меняя среду обитания человека. Подобным образом новые разработки и технологии влияют и на традиционные отрасли народного хозяйства, к которым относится энергетика, являющаяся базовой и от состояния которой зависит многие отрасли экономики. Научно-технологическое прогнозирование и организация мониторинга инновационных технологических решений в области энергетики требуют разработки методов семантического анализа Big Data для выработки оценок и опережающих рекомендаций, а также создания новых инструментальных средств для поддержки этой деятельности. Необходимость развития методов анализа и обработки Big Data с помощью интеллектуальных информационных систем подчеркивается Национальной технологической инициативой (НТИ) России, а их применение в области энергетики

соотносится с рынком EnergyNet, что находит отражение в «Дорожной карте», одобренной в 2016 году Президиумом Совета при Президенте Российской Федерации по модернизации экономики и инновационному развитию России.

**Управление инновациями и поиск технологических решений.** В последние десятилетия наблюдается активное развитие целого спектра интеллектуальных подходов и методов для поддержки принятия решений по планированию и управлению инновационным развитием как отдельных секторов, так и экономики в целом [9, 12]. С начала 2000-х гг. была сформирована международная рабочая группа из ведущих ученых США, Европы и стран Восточной Азии, которая координирует исследования по перспективному анализу научно-технологического развития (Future Oriented Technology Analysis) [8]. Целью подобных исследований является разработка средств интеллектуальной поддержки систематического процесса обоснования возможных путей развития науки и технологий в различных областях, оценки перспективного влияния новых технологий на общество и окружающий мир, в том числе и на конкретные отраслевые инфраструктуры, а также поддержка выработки «скользящих» стратегических решений по инновационному развитию как отдельных технологических компаний, так и отраслей мировой экономики. Традиционно обоснованные научно-технологические прогнозы и программы инновационного развития являлись одной из важнейших устоявшихся форм регулирования экономики в таких странах, как США и Великобритания. Однако, в связи со стремительным развитием информационных технологий и наступлением «эпохи больших данных» (Big Data Age), в последние несколько лет особенно активно выполняются научные исследования в этой области коллективами из Китая, которые проспонсированы Национальным фондом естественных наук (National Natural Science Foundation of China – NSFC). Значительная часть методов определения новых технологических решений основана на использовании интеллектуальных семантических технологий поиска, извлечения и анализа гетерогенных данных из электронных источников информации в соответствии с концепцией Big Data [4, 5]. Методика Tech Mining применяется, как форма статистического контекстного анализа текстовых документов на базе научно-технических ресурсов для выявления прорывных технологий и разработок, оценки их инновационного потенциала [10]. Эта методика, основанная на подходе Text Mining, позволяет выявлять в фрагментах текстовых данных корреляции, связи и ранее неизвестные, но полезные с практической точки зрения интерпретируемые знания [11, 13, 14], применение которых возможно в различных сферах.

В России для решения задач по поддержке принятия решений в научно-технологическом прогнозировании энергетики используются две группы подходов. Первая группа основана на методах научно-технологического форсайта: все обзоры, оценки, прогнозы и рекомендации формируются на основе консолидированного мнения экспертов [3]. Этот подход часто критикуется за субъективность, слабую обоснованность и отсутствие ответственности экспертов, хотя и является достаточно простым, оперативным и малозатратным. Методология системного анализа относится ко второй группе и является наиболее строго научно обоснованным подходом, поскольку использует в качестве основных инструментов методы системного оценивания и сопоставления технологий и системное технологическое моделирование (математические модели технологий и их структуры). В

математических моделях учитываются не только энергетические технологии по всей совокупности их основных характеристик, но и целый комплекс внешних факторов (потребности, ресурсы, экология, функционирование, институты и т.п.) [1]. Применение указанных инструментов для разработки программ инновационного развития в энергетике требует выполнения полномасштабных научно-исследовательских работ силами высококвалифицированных коллективов, больших финансовых затрат и достаточное количество времени. Результаты таких исследований, как правило, очень высокого качества и предлагают отличные решения для конкретных отраслевых научно-технических задач и задач формирования обобщенных стратегий развития энергетики [2]. Поскольку исследования выполняются значительное время, возникает проблема отсутствия эффективных механизмов оперативной адаптации аппарата системного технологического моделирования к меняющимся условиям технологического развития. Особенно это проявляется в части учета факторов научно-технического прогресса в смежных или слабосвязанных областях, в сложности учета междисциплинарных и инфраструктурных технологических эффектов и в возросшей информационной, технической и организационной комплексности объектов исследования.

**Предлагаемые методы и подходы для поддержки прогнозирования инновационного развития энергетики.** Методы прогнозирования развития энергетической отрасли, основанные на традиционных математических моделях и программных комплексах, не всегда эффективны в условиях неопределенности и отсутствия необходимой достоверной информации для имеющихся моделей. Для решения поставленных задач предлагается использовать интеллектуальные методы семантического анализа, машинного обучения и технологии Big Data для создания инструментария, облегчающего работу экспертных групп, и инструментария, выполняющего предварительную обработку информации, анализируемой экспертами. Источниками информации могут выступать Открытые данные (Open Data) и Большие данные (Big Data). Кроме того, могут эффективно использоваться методы семантического моделирования для «экспресс-анализа» собранной информации [7], разработка которых ведется в отделе Систем искусственного интеллекта в энергетике Института систем энергетики им Л.А. Мелентьева Сибирского отделения РАН (ИСЭМ СО РАН)<sup>1</sup>.

Авторами предлагается использовать методы построения онтологического пространства знаний в энергетике и семантического моделирования: когнитивного (для описания причинно-следственных связей), событийного (для моделирования последствий принимаемых решений) и вероятностного (для оценки перспектив и рисков) моделирования. Предлагается использование паттернов (шаблонов), основанных на экспертных суждениях (прогнозных гипотезах) об инновационных путях развития энергетики и перспективных технологических решениях в энергетике, которые могут использоваться для поиска информации и проверки гипотез. В общем случае гипотезы формируются на основе анализа массива собранных данных. Поскольку данные собираются из открытых источников, они носят явно статистический характер и к ним применимы статистические методы, в том числе методы статистической оценки гипотез и критериев. Ввиду большого объема собираемых

---

<sup>1</sup> <http://isem.irk.ru/institute/departments/>

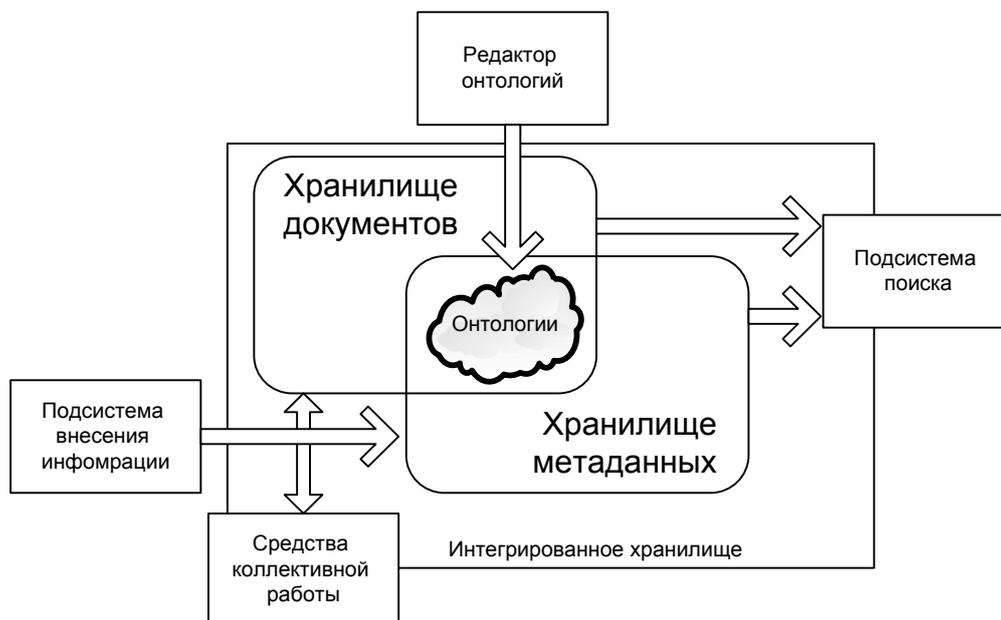
данных целесообразно применение второго уровня статистической значимости (0,01 или 1%), что задает уровень статистической ошибки первого рода не более 1% и достоверность получаемых результатов при проверке гипотез не менее 99%.

Использование в качестве источников данных для анализа и последующего научно-технологического прогнозирования развития энергетики исключительно российских информационных систем и тематических баз данных является ошибочным, поскольку из-за глобализации инновационные разработки и результаты научных исследований не ограничены экономикой одной отдельно взятой страны или макрорегиона. Аналогичным образом обосновывается и необходимость анализа информации не только на русском, но и на английском языке. Наличие множества публикаций на русском языке является одним из индикаторов готовности к практическому внедрению новой технологии в производственные процессы на предприятиях и организациях России.

Кроме того, является целесообразным применение методов и математических моделей исследования направлений развития энергетики, которые неоднократно использовались в ИСЭМ СО РАН при разработке стратегий развития энергетики России. Коллективом отдела Систем искусственного интеллекта в энергетике под руководством д.т.н., проф. Л.В. Массель предложена и реализована двухуровневая технология исследований проблем энергетической безопасности [6], в которой на первом, верхнем уровне экспертами выполняется качественный анализ вариантов развития энергетики с использованием семантических технологий, а на втором (нижнем) уровне для количественного обоснования выбранных вариантов используются экономико-математические модели ТЭК России большой размерности и традиционные программные комплексы для многовариантных расчетов (например, программный комплекс ИНТЭК-М).

**Архитектура интеллектуальной информационной системы.** Одним из основных компонентов интеллектуальной информационной системы для поддержки экспертных решений является интегрированное хранилище (база знаний), включающее семантические модели, а также методы семантического анализа Big Data в виде программных компонентов (агентов). Архитектура интегрированного хранилища интеллектуальной информационной системы и основные его компоненты представлены на рисунке 1. Важным компонентом хранилища является система онтологий, которая включает онтологии отраслей энергетики и отдельных энергетических технологий, онтологию исследований энергетики, онтологии задач, методов исследований и др., используется для интеграции знаний и позволяет унифицировать описания методов анализа, форматов, используемых в них данных, спецификации реализующих эти методы программных компонентов. Подсистема внесения информации инкапсулирует программы-краулеры и адаптеры источников данных, используемые для поиска, построения описания и загрузки потенциально интересной информации в хранилище. Подсистема поиска является точкой получения информации (метаданных, ссылок на внешние ресурсы и документов исследователей, описанных в понятиях системы онтологий), которая после извлечения может быть представлена пользователю или обработана другими программными агентами. Средства коллективной работы, научные прототипы которых были реализованы под руководством Копайгородского А.Н., позволяют разрабатывать, описывать и управлять документами, создаваемыми

исследователями и отражающими некоторые полученные научные результаты. В качестве редактора онтологий предполагается применение инструментального средства OnotMap, которое имеет расширяемую архитектуру и может быть использовано как средство конфигурирования, задающее специфику предметной области и решаемой задачи. Указанная особенность интеллектуальной информационной системы является ключевой и позволяет использовать некоторые ее компоненты для решения задач, например, в области научного обоснования стратегических решений по цифровой трансформации энергетики. Это достигается за счет более позднего наложения ограничений на предметную область, что значительно увеличивает гибкость системы. Кроме того, предлагается применение сервис-ориентированного подхода, упрощающего использование как оригинальных авторских, так и имеющихся в свободном доступе готовых реализаций методов, что позволяет снять ряд проблем, связанных с различием платформ, стандартов, форматов данных, а также с удаленным размещением реализаций этих методов.



**Рис. 1.** Архитектура интегрированного хранилища интеллектуальной информационной системы.

Исходя из опыта реализации других научных проектов, связанных со сбором и обработкой большого объема данных, было принято решение о реализации хранилища метаданных в виде распределенной базы данных (рис. 2). Для улучшения производительности в хранилище метаданных индексы выделяются в отдельные структуры. Доступ к файлам, содержащим текстовые данные, которые хранятся на отдельных носителях, осуществляется через дескрипторы, размещенные в базах метаданных.

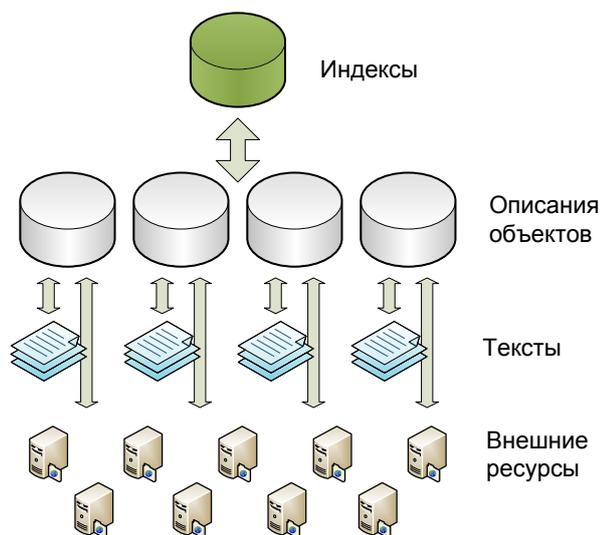


Рис. 2. Распределенная архитектура хранилища метаданных.

На рисунке 3 показана архитектура интеллектуальной информационной системы, включающая средства семантического анализа тактовых данных, онтологического, когнитивного и событийного моделирования, средства проверки гипотез и визуализации результатов поиска. При обработке результатов поисковых запросов во внешних, по отношению к интегрированному хранилищу, программных средствах извлекаемых данных исследователи получают новые знания, которые могут быть представлены в явной форме и загружены в интегрированное хранилище с использованием средств коллективной работы.

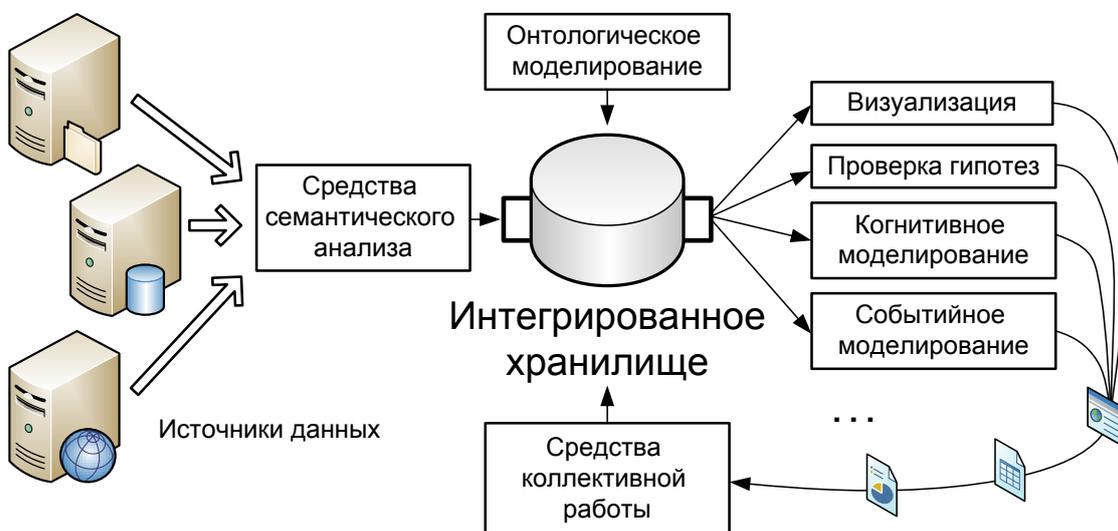


Рис. 3. Архитектура интеллектуальной информационной системы.

Применение для связи с внешними компонентами стандартизированных интерфейсов и сетевых протоколов Интернет, в частности JSON и HTTP, обеспечивает интеграцию методов при решении нетривиальных задач и позволяет получить доступ к информации и функциональным компонентам, размещенным на удаленных серверах.

**Заключение.** Реализация интеллектуальной информационной системы с применением сервис-ориентированного подхода и интеграцией существующих сторонних и авторских компонентов позволит эффективно решать научно-практические задачи, как в

области поддержки экспертных решений по стратегическому инновационному развитию энергетики, так и в области научного обоснования стратегических решений по цифровой трансформации энергетики. Применение предложенных методов и подходов для проектирования и реализации отдельных компонентов интеллектуальной информационной системы обеспечивает достаточный уровень ее гибкости и расширяемости.

Работа выполнена при частичной финансовой поддержке грантов РФФИ №20-07-00994 и № 19-07-00351.

#### СПИСОК ЛИТЕРАТУРЫ

1. Системные исследования в энергетике: ретроспектива научных направлений СЭИ-ИСЭМ / отв. ред. Н.И. Воропай. Новосибирск: Наука. 2010. 686 с.
2. Энергетика XXI века. Условия развития. Технологии. Прогнозы. / Отв. ред. Н.И. Воропай. Новосибирск: Наука. 2010. 386 с.
3. Прогноз научно-технологического развития России: 2030 / под. ред. Л.М. Гохберга, С.П. Филиппова. Москва: Министерство образования и науки Российской Федерации, Национальный исследовательский университет «Высшая школа экономики». 2014. 244 с.
4. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим. М.: Манн, Иванов и Фербер. 2014. 230 с.
5. Н. Марц, Дж. Уоррен. Большие данные. Принципы и практика построения масштабируемых систем обработки данных в реальном времени. М.: Вильямс. 2016. 368 с. ISBN: 978-5-8459-2075-1
6. Массель Л.В. Создание и интеграция интеллектуальных информационных технологий и ресурсов для комплексных исследований в энергетике // Вестник РФФИ. 2012. №4 С. 74-81
7. Массель Л.В., Массель А.Г. Интеллектуальные вычисления в исследованиях направлений развития энергетики // Известия Томского политехнического университета. 2012. Т. 321. № 5. Управление, вычислительная техника и информатика. С. 135-141.
8. Future-Oriented Technology Analysis. Strategic Intelligence for an Innovative Economy. Springer. 2008. 170 p. DOI: 10.1007/978-3-540-68811-2.
9. V. Coates, et al. On the future of technological forecasting // Technol. Forecast. Soc. Change. 67 (1). 2001. Pp. 1 – 17.
10. Cunningham S.W., Porter, A.L., and Newman, N.C. Tech Mining, special issue of Technological Forecasting & Social Change. 73 (8). 2006. Pp. 915-1060.
11. Mirhosseini M.A Clustering Approach using a Combination of Gravitational Search Algorithm and k-Harmonic Means and its Application in Text Document Clustering // Turkish Journal of Electrical Engineering and Computer Sciences. Vol. 25. № 2. 2017. Pp. 1251-1262. DOI: 10.3906/elk-1508-31.
12. Zhang Y., et al. Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research. Technol. Forecast. Soc. Change. 2016. DOI: 10.1016/j.techfore.2016.01.015.
13. Zhang M.L, Zhou Z.H. Multilabel Neural Networks with Applications to Functional Genomics and Text Categorization // IEEE Transactions on Knowledge and Data Engineering. Vol. 18. № 10. 2006. Pp. 1338-1351. DOI:10.1109/TKDE.2006.162.
14. Zheng L., Noroozi V., Yu P.S. Joint Deep Modeling of Users and Items using Reviews for Recommendation. 2017. DOI:10.1145/3018661.3018665.

**UDK 004.8:001.8:620.9**

**ARCHITECTURE OF THE INTELLECTUAL INFORMATION SYSTEM TO SUPPORT  
EXPERT DECISIONS ON STRATEGIC INNOVATIVE ENERGY DEVELOPMENT**

**Alex N. Kopaygorodsky**

Ph.D., Leading specialist, Department of Artificial Intelligence Systems,

e-mail: [kopaygorodsky@isem.irk.ru](mailto:kopaygorodsky@isem.irk.ru),

**Timur G. Mamedov**

Graduate student, Department of Artificial Intelligence Systems,

e-mail: [mamedowtymur@yandex.ru](mailto:mamedowtymur@yandex.ru),

Melentiev Energy Systems Institute

Siberian Branch of the Russian Academy of Sciences

130, Lermontov Str., 664033, Irkutsk, Russia.

**Abstract.** The article discusses methods of building an intelligent information system to support expert decisions on strategic innovative development of the energy sector. The necessity of using Big Data analysis methods has been substantiated. Architecture of an integrated repository of an intelligent information system is presented, in which the main component is a system of ontologies on the basis of which information, data and knowledge from various sources are combined.

**Keywords:** scientific and technological forecasting, semantic analysis, support for scientific research, information system architecture

**References**

1. Sistemnye issledovaniya v jenergetike: retrospektiva nauchnyh napravlenij SEI-ISEM [Systems research in energy: a retrospective of the scientific directions of SEI-ISEM] / Ed. N.I. Voropay. Novosibirsk: Nauka = Science. 2010. 686 p.
2. Energetika XXI veka. Uslovija razvitija. Tehnologii. Prognozy. [Energy of the XXI century. Development conditions. Technologies. Forecasts.] / Ed. N.I. Voropay. Novosibirsk: Nauka = Science. 2010. 386 p.
3. Prognoz nauchno-tehnologicheskogo razvitija Rossii: 2030 [Forecast of scientific and technological development of Russia: 2030] / Ed. L.M. Gohberg, S.P. Filippov. Moscow: Ministerstvo obrazovaniya i nauki Rossijskoj Federacii, Nacional'nyj issledovatel'skij universitet «Vysshaja shkola jekonomiki» = Ministry of Education and Science of the Russian Federation, National Research University Higher School of Economic. 2014. 244 p.
4. Mayer-Shenberger V., Cukier K. Bol'shie dannye. Revoljucija, kotoraja izmenit to, kak my zhivem, rabotaem i myslim [Big Data. A Revolution That Will Transform How We Live, Work and Think]/ Trans. from Eng. I. Gayduk. Moscow: Mani, Ivanov and Ferber. 2014. 230 p.
5. Nathan Marz, James Warren. Bol'shie dannye. Principy i praktika postroenija masshtabiruemyh sistem obrabotki dannyh v real'nom vremeni [Big Data: Principles and best practices of scalable realtime data systems]. Moscow: Wilyams = Moscow.: Williams. 2016. 368 p. ISBN: 978-5-8459-2075-1
6. Massel' L.V. Sozdanie i integracija intellektual'nyh informacionnyh tehnologij i resursov dlja kompleksnyh issledovanij v jenergetike [Creation and integration of intelligent information

- technologies and resources for complex research in energy] // RFBR Bulletin. 2012. №4. Pp. 74-81
7. Massel' L.V., Massel' A.G. *Intellektual'nye vychislenija v issledovanijah napravlenij razvitija jenergetiki [Intelligent computing in studies of energy development directions]* // Bulletin of the Tomsk Polytechnic University. 2012. Vol. 321. № 5. Management, computer technology and informatics. Pp. 135-141.
  8. *Future-Oriented Technology Analysis. Strategic Intelligence for an Innovative Economy.* Springer. 2008. 170 p. DOI: 10.1007/978-3-540-68811-2.
  9. V. Coates, et al. *On the future of technological forecasting* // Technol. Forecast. Soc. Change. 67 (1). 2001. Pp. 1 – 17.
  10. Cunningham S.W., Porter A.L., and Newman, N.C. *Tech Mining*, special issue of *Technological Forecasting & Social Change*. 73 (8). 2006. Pp. 915-1060.
  11. Mirhosseini M.A *Clustering Approach using a Combination of Gravitational Search Algorithm and k-Harmonic Means and its Application in Text Document Clustering* // Turkish Journal of Electrical Engineering and Computer Sciences. Vol. 25. № 2. 2017. Pp. 1251-1262. DOI: 10.3906/elk-1508-31.
  12. Zhang Y., et al. *Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research.* Technol. Forecast. Soc. Change. 2016. DOI: 10.1016/j.techfore.2016.01.015.
  13. Zhang M.L, Zhou Z.H. *Multilabel Neural Networks with Applications to Functional Genomics and Text Categorization* // IEEE Transactions on Knowledge and Data Engineering. Vol. 18. № 10. 2006. Pp. 1338-1351. DOI:10.1109/TKDE.2006.162.
  14. Zheng L., Noroozi V., Yu P.S. *Joint Deep Modeling of Users and Items using Reviews for Recommendation.* 2017. DOI:10.1145/3018661.3018665.