



Российская Академия Наук

ТРУДЫ

Кольского научного центра РАН

10/2018(9)

**ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ**

выпуск 9

Апатиты
2018

Российская Академия Наук

10/2018 (9)

издается с декабря 2010 г.

УДК 004.9

ISSN 2307-5252

ТРУДЫ

Кольского научного центра

Главный редактор —
чл.-корр. РАН, проф. С. В. Кривовичев

Заместители главного редактора:
д. т. н. В. А. Маслобоев

Редакционный совет:
академик РАН Г. Г. Матишов,
академик РАН Н. Н. Мельников,
чл.-корр. РАН В. К. Жиров,
чл.-корр. РАН А. Н. Николаев,
д. э. н. Ф. Д. Ларичкин,
д. т. н. В. А. Путилов,
д. ф.-м. н. Е. Д. Терещенко,
к. г.-м. н. А. Н. Виноградов (отв. секретарь)

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Редколлегия серии
«Информационные технологии»:
профессор, д.т.н. В.А. Путилов (отв. редактор),
д.т.н. А.Г. Олейник (зам. отв. редактора),
профессор, д.т.н., В.А. Марлей,
д.т.н. В.А. Маслобоев

выпуск 9

DOI: 10.25702/KSC.2307-5252.2018.10

Научное издание

Редактор А.Л. Щур
Технический редактор: В.Ю. Жиганов

184209, Мурманская область, г. Апатиты, ул. Ферсмана, 14,
ФИЦ Кольский научный центр РАН
Тел.: (81555) 79393, 79380, факс: (81555) 76425
E-mail: admin@admksk.apatity.ru, <http://www.kolasc.net.ru>

Подписано к печати 15.11.2018. Формат бумаги 70×108 1/16.
Усл. печ. л. 15.6. Заказ № 40. Тираж 300 экз.
Издательство ФГБУН ФИЦ КНЦ РАН
184209, г. Апатиты, Мурманская область, ул. Ферсмана, 14

© Институт информатики и математического моделирования технологических процессов – обособленное подразделение Федерального государственного бюджетного учреждения науки Федерального исследовательского центра «Кольский научный центр Российской академии наук», 2018
© ФГБУН ФИЦ «Кольский научный центр Российской академии наук», 2018

СОДЕРЖАНИЕ

		Стр.
	Введение.....	8
A.B. Маслобоев B.A. Путилов M.B. Игнатъев	Лингво-комбинаторная модель управления региональной безопасностью.....	9
A.A. Зуенко O.B. Фридман	Применение методов программирования в ограничениях в задачах интеллектуального планирования.....	21
C.H. Малыгина B.B. Быстров Д.Н. Халиуллина	Логистика кадрового обеспечения региона: формализация и структура полимодельного комплекса	36
B.B. Быстров Д.Н. Халиуллина	Формализованная концептуальная модель обеспечения экологической безопасности арктических коммуникаций.....	48
E.C. Давидюк M.Г. Шишаев B.B. Быстров	Прогнозное моделирование для мониторинга и управления кадровым обеспечением программ регионального развития.....	61
A.B. Вицентий M.Г. Шишаев П.А. Ломов Г.Г. Гогоберидзе	Разработка онтологии предметной области для формального логического вывода и визуализации в системах информационной поддержки задач морского пространственного планирования.....	77
A.B. Вицентий M.Г. Шишаев Г.Г. Гогоберидзе	Разработка технологии геовизуализации для систем информационной поддержки задач пространственного планирования с учетом интересов и противоречий различных групп природопользователей на основе концептуальной модели конфликта.....	91
A.B. Вицентий M.Г. Шишаев B.B. Диковицкий	Опыт использования визуализации семантических моделей документов для экспресс контент-анализа (на примере нормативно-правовых актов, регламентирующих развитие пространственно-распределенных систем).....	99
B.B. Диковицкий A.A. Менькова Л.М. Шишаева	Модальная логика как средство формализации нормативных документов.....	112

П.А. Ломов	Программная реализация технологии генерации лексико-синтаксических паттернов для поддержки решения задач обучения онтологий.....	120
А.Г. Олейник	Предварительные результаты исследований по проекту «Разработка модели, методов интеграции и системного использования междисциплинарных знаний и данных для поддержки управления комплексным развитием АЗРФ»	130
С.Ю. Яковлев	Особенности развития арктической зоны РФ (по материалам конференции OMR 2018 и конкурса Арктика 2018).....	144
И.О. Датъев М.Г. Шишаев А.М. Федоров	Модель коллективной идентичности для изучения местных сообществ.....	151
А.О. Орлов Ю.Г. Смирнов В.В. Бирюков	Анализ проектных решений подземных комплексов атомных станций малой мощности в условиях Арктики	161
А.С. Шемякин Ю.В. Федотова	Алгоритм визуализации распределения зон геодинамического риска.....	170

Russian Academy of Sciences

10/2018 (9)

UDC 004.9
ISSN 2307-5252

TRANSACTIONS

Kola Science Centre

Editor-in-Chief — S. V. Krivovichev,
Corr. Member of the RAS, Prof.
Deputy Editors-in-Chief:
V. A. Masloboev, Dr. Sc. (Engineering)

Editorial Council:
G. G. Matishov, Acad. of the RAS,
N. N. Melnikov, Acad. of the RAS,
V. K. Zhirov, Cor. Member of the RAS,
A. N. Nikolaev, Cor. Member of the RAS,
F. D. Larichkin, Dr. Sc. (Economics),
V. P. Petrov, Dr. Sc. (Geology and Mineralogy),
V. A. Putilov, Dr. Sc. (Engineering),
E. D. Tereshchenko, Dr. Sc. (Physics and Mathematics),
A. N. Vinogradov, PhD (Geology and Mineralogy) —
Executive Secretary

**INFORMATION
TECHNOLOGIES**
series 9

Editorial Board of Information Technologies Series:
V.A. Putilov, Dr.Sc., Prof. (Editor-in-Chief),
A.G. Oleynik, Dr.Sc. (Eng.) (Vice Editor-in-Chief),
V.A. Marley, Dr.Sc. (Eng.),
V.A. Masloboev, Dr.Sc. (Eng.)

DOI: 10.25702/KSC.2307-5252.2018.10

14, Fersman str., Apatity, Murmansk Region, 184209, Russia
Tel. (81555) 79380. Fax: (81555) 76425
E-mail: admin@admksk.apatity.ru, <http://www.kolasc.net.ru>

© Institute for Informatics and Mathematical Modeling - Subdivision of the Federal Research Centre "Kola Science Centre of the Russian Academy of Sciences", 2018
© Federal Research Centre "Kola Science Centre of the Russian Academy of Sciences", 2018

CONTENTS

	Page
Introduction	8
A.V. Masloboev V.A. Putilov M.B. Ignatyev	Linguo-combinatorial model of regional security control..... 9
A.A. Zuenko O.V. Fridman	Application of constraint programming methods in intellectual planning problems..... 21
S.N. Malygina V.V. Bystrov D.N. Khaliullina	Logistics of personnel provision in the region: formalization and structure of the poly-model complex..... 36
V.V. Bystrov D.N. Khaliullina	Formalized conceptual model of ecological safety of the arctic communications..... 48
E.S.Davidyuk M.G.Shishaev V.V.Bystrov	Forecast simulation for monitoring and management of staffing need for regional development programs..... 61
A.V. Vicentiy M.G. Shishaev P.A. Lomov G.G. Gogoberidze	Development of domain ontology for formal inference and visualization in information support systems for the tasks of maritime spatial planning..... 77
A.V. Vicentiy M.G. Shishaev G.G. Gogoberidze	Development of geovisualization technology for information support systems for spatial planning tasks taking into account interests and contradictions of various groups of nature users based on a conceptual model of conflict..... 91
A.V. Vicentiy M.G. Shishaev V.V. Dikovitskiy	Experience in using the visualization of document semantic models for express content analysis (on the example of legal acts regulating the development of spatially distributed systems)..... 99
V.V. Dikovitskiy A.A. Menkova L.M. Shishaeva	Modal logic as a means of formalization of normative documents..... 112
P.A. Lomov	Software implementation of the technology for generation of lexical-syntactic patterns for support of ontology learning..... 120

A.G. Oleynik	Preliminary results of research on the project «Development of the model, methods of integration and use of interdisciplinary knowledge and data for management support of the Russian Arctic complex development».....	130
S.Yu. Yakovlev	Development features of the russian federation arctic zone (on the materials of the OMR 2018 conference and the Arctic 2018 competition)....	144
I.O. Datyev M.G. Shishaev A.M. Fedorov	The collective identity model for the local communities study.....	151
A. O. Orlov Y. G. Smirnov V.V. Birukov	Analysis of design decisions for underground complexes of small nuclear power plants in Arctic.....	161
A.S. Shemyakin Iu.V. Fedotova	Algorithm for visualization of the distribution of geodynamic risk zones.....	170

INTRODUCTION

The development of the theory and technologies of modeling has given us the opportunity to use them for the ever wider study of the systems and processes that fall into the category of weakly formalized ones. Identification and deep understanding of the systems properties, interconnections and regularities that determine these properties make it possible to work out valid decisions concerning their development and predict the possible reactions of the systems to various actions. At the same time, an important role is given to technologies of transforming the qualitative descriptions of the systems under study into formal structures suitable for computer processing. A number of articles presented in the ninth edition of the “Information Technologies” series of the Works of the Kola Scientific Center of the Russian Academy of Sciences are devoted to such technologies. Specifically, linguo-combinatorial modeling, intelligent planning using constraint programming methods, and a logistic approach to managing complex systems are considered. The application aspect of the developed methods and modeling technologies is determined by the features of the Arctic zone of the Russian Federation (AZRF). The subjects of research include regional safety in a broad sense, workforce for the development of the Russian Arctic, marine economic complexes.

The main source of information about weakly formalized systems are natural language texts. This determines the urgency of the methods development for automated extraction, formalization and systematization of knowledge received from texts. This issue includes several articles on researches in this area, such as development of technologies for semantic analysis and visualization of semantic models of text documents, as well as the use of modal logic in the analysis of statements in normative documents to identify possible contradictions and incompleteness of the legal field. A technology that uses ontological design pattern is proposed to generate, expand and fill ontologies based on the analysis of thematic natural language texts. The technology that uses ontological design pattern is proposed to generate, expand and fill ontologies based on the analysis of thematic natural language texts. The integration of ontological modeling technologies with the methods for solve problems of satisfying constraints makes it possible to reduce the time of information systems response to a database query by optimizing the execution of ontology-based data access queries.

The current issue also includes an overview of the development features of the Russian Arctic, based on materials from the Offshore Marintec Russia 2018 conference and a competition of scientific, technical and innovative ideas aimed at the development of the Arctic and the continental shelf (Arctic 2018)

The issue is addressed to experts in the field of creation and practical use of information systems and technologies in various areas of management and production activities, teachers and students of universities of relevant specialties.

ВВЕДЕНИЕ

Развитие теории и технологий моделирования обеспечило возможность их все более широкого применения для исследования систем и процессов, которые относятся к категории слабо формализованных. Выявление и глубокое понимание свойств систем, определяющих эти свойства взаимосвязей и закономерностей, дает возможность вырабатывать обоснованные решения по управлению их развитием и прогнозировать возможные реакции систем на различные воздействия. При этом важная роль отводится технологиям преобразования качественных описаний исследуемых систем в формальные структуры, пригодные для компьютерной обработки. Именно таким технологиям посвящен ряд работ, вошедших в девятый выпуск серии «Информационные технологии» Трудов Кольского научного центра РАН. В частности, рассматриваются вопросы лингво-комбинаторного моделирования, интеллектуальное планирование с применением методов программирования в ограничениях, применение логистического подхода к управлению сложными системами. Прикладной аспект применения разрабатываемых методов и технологий моделирования определяется особенностями Арктической зоны Российской Федерации (АЗРФ). В качестве предметов исследования фигурируют региональная безопасность в широком смысле, кадровое обеспечение развития АЗРФ, морехозяйственные комплексы.

Основным источником информации о слабо формализованных системах являются тексты на естественном языке. Это определяет актуальность разработки и развития методов автоматизированного извлечения, формализации и систематизации знаний, получаемых из текстов. В сборник включены несколько работ по этому направлению. Представлены разработка технологий семантического анализа и визуализации семантических моделей текстовых документов, а также использование модальной логики в задачах анализа утверждений в нормативных документах для выявления возможных противоречий и неполноты правового поля. Для генерации, расширения и наполнения онтологий на основе анализа тематических естественно-языковых текстов предлагается технология использования онтологических паттернов проектирования. Интеграция технологий онтологического моделирования и методов решения задач удовлетворения ограничений позволяет сократить время отклика информационных систем на запрос к базе данных за счет оптимизации выполнения запросов доступа к данным на основе онтологий.

В настоящий выпуск включен обзор особенностей развития АЗРФ, выполненный на основе материалов конференции Offshore Marintec Russia 2018 и конкурса научных, научно-технических и инновационных разработок, направленных на развитие и освоение Арктики и континентального шельфа (Арктика 2018)

Сборник адресован специалистам в области создания и практического использования информационных систем и технологий в различных сферах управленческой и производственной деятельности, преподавателям и студентам вузов соответствующих специальностей.

УДК 004.94, 338.24

DOI: 10.25702/KSC.2307-5252.2018.10.9-20

А.В. Маслобоев¹, В.А. Путилов¹, М.Б. Игнатъев²

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² Санкт-Петербургский государственный университет аэрокосмического приборостроения

ЛИНГВО-КОМБИНАТОРНАЯ МОДЕЛЬ УПРАВЛЕНИЯ РЕГИОНАЛЬНОЙ БЕЗОПАСНОСТЬЮ*

Аннотация

В работе рассматриваются вопросы лингво-комбинаторного моделирования слабо формализованных региональных социально-экономических систем, для которых существует лишь описание на естественном языке. Разработана лингво-комбинаторная модель поддержки управления безопасностью региона. Модель состоит из трех групп переменных: характеристик основных понятий предметной области, изменения этих характеристик и структурированной неопределенности в эквивалентных уравнениях, которая может быть использована для адаптации и управления социально-экономической системой. На практике модель обеспечивает анализ поведения системы и реализацию различных механизмов удержания ее в зоне адаптационного максимума для обеспечения безопасности функционирования и развития системы.

Ключевые слова:

лингво-комбинаторное моделирование, управление, социально-экономическая система, региональная безопасность, поддержка принятия решений, адаптационный максимум

A.V. Masloboev, V.A. Putilov, M.B. Ignatyev

LINGUO-COMBINATORIAL MODEL OF REGIONAL SECURITY CONTROL

Abstract

The paper considers linguo-combinatorial simulation problems of loosely-formalized regional socio-economic system which are described only in natural language. A linguo-combinatorial model for regional security management support has been developed. The model consists of three class variables: knowledge domain basic characteristics, dynamics of these characteristics and structural uncertainty represented in equivalent equations. That structural uncertainty is used for control and adaptation of socio-economic system. In practical applications a proposed model provides complex system analysis and different mechanism implementation for system retention in the adaptation maximum area to security support of system functioning and development.

Keywords:

linguo-combinatorial simulation, control, socio-economic system, regional security, decision-making support, adaptation maximum

Введение

В современных условиях повышаются требования к технологиям и средствам обеспечения безопасности во всех сферах развития социально-экономических систем. В связи с этим, разработка средств информационной поддержки принятия решений по управлению безопасностью крупномасштабных

* Работа выполнена при поддержке Министерства науки и высшего образования РФ и Российского фонда фундаментальных исследований - проект № 18-29-03022-мк.

систем в условиях кризисных ситуаций социально-экономического характера, несмотря на достаточно большое количество работ в данной области [1-4], является актуальным и перспективным направлением исследований. Принципиально важно совершенствовать это направление на современном этапе развития нашей страны.

Понятие безопасности предполагает представление об опасностях на уровне гипотез с неопределенностью. Различная природа, скрытый характер, неоднородность, отложенность результатов воздействия и медленное нарастание потенциальных угроз и опасностей в социально-экономической сфере обуславливают то, что для кризисных ситуаций не может быть все заранее учтено и расписано. Поэтому существующие модели и методы управления безопасностью требуют адаптации для приложения к социально-экономическим системам.

Наиболее остро проблемы безопасности проявляются на региональном уровне, что может привести к дестабилизации социально-экономических систем более высокого уровня – федерального, международного, мирового. Поэтому важной и актуальной задачей является совершенствование существующей системы и механизмов организационного управления безопасностью региональных социально-экономических систем. В настоящее время эта задача еще далека от эффективного решения. Особенно она актуальна для арктических регионов России [9].

Работа посвящена развитию методологии информационной поддержки управления безопасностью сложных социально-экономических объектов и инфраструктур различной природы и масштаба на основе применения метода лингво-комбинаторного моделирования, предложенного в работах профессора М.Б. Игнатьева [5-7] для исследования сложных плохо формализованных систем. Этот метод развивался в различных приложениях: от задач синтеза атомно-молекулярных систем и физико-химических реакций до моделирования атмосферы и других биологических, социально-экономических и технических систем. Исходная посылка метода лингво-комбинаторного моделирования – решение проблем координации и управления сложными системами путем удержания этих систем в зоне так называемого адаптационного максимума [5].

В связи с этим, целесообразным представляется анализ применимости метода лингво-комбинаторного моделирования для решения задач управления безопасностью сложных плохо формализованных систем, к которым относятся региональные социально-экономические системы. Это обеспечит развитие теории управления безопасностью сложных систем в части структурирования, формализации и компьютерного моделирования задач управления в этой сфере.

1. Подход к лингво-комбинаторному моделированию сложных систем

Для того чтобы управлять региональной безопасностью, необходимо создавать новые модели социально-экономических систем и процессов, протекающих в этих системах. При этом в сами модели должна быть заложена возможность информационного управления. В этом сущность процесса информационной поддержки принятия управленческих решений. Под информационной поддержкой понимается процесс внедрения новых информационных и телекоммуникационных технологий в различные сферы социально-экономической жизни, но этого не всегда достаточно для принятия

эффективных решений в процессе управления безопасностью сложных систем. ЛПР в основном используют инструментарий, основанный на формальных моделях прошлого века и их современных модификациях.

Другой недостаток современной теории безопасности сложных систем заключается в том, что факт признается достоверным, если он воспроизводим. В социально-экономических системах обеспечить повторяемость практически невозможно. Можно лишь утверждать, что повторяющихся явлений в социально-экономических системах крайне мало, как число похожих людей в обществе – каждый человек характеризуется своей индивидуальностью. Для управления безопасностью таких систем важна не только воспроизводимая информация, но и невоспроизводимая.

Любая предметная область опирается на модели реальных процессов. При этом в каждой области эти модели имеют разную полноту формализации, но все они используют естественный язык. Естественный язык – это универсальная моделирующая система.

Теория моделирования должна помогать решать еще нерешенные задачи, одной из которых является управление безопасностью социально-экономических систем. Эти системы относятся к классу слабо формализованных систем. Поэтому, чтобы создать новые и адаптировать существующие модели и методы информационной поддержки и превратить их в сквозную технологию управления социально-экономической безопасностью, необходимо осуществить большую работу по формализации этих систем. Для этого необходимо:

- 1) сформулировать экспертные знания об исследуемой предметной области или объекте на естественном языке;
- 2) описать ситуации и задачи в данной области на естественном языке;
- 3) построить математическую модель системы или объекта управления, сформулировать задачи на языке основных соотношений;
- 4) перевести математические формулировки на тот или иной язык программирования;
- 5) реализовать программу на компьютере на языке конкретной машины;
- 6) получить результат решения задачи на языке результата в виде таблиц, графиков, текстов, анимаций и т.д.

Центральная проблема – как перейти от описания на естественном языке к описанию на языке основных соотношений. Для решения этой проблемы предлагается использовать лингво-комбинаторное моделирование плохо формализованных систем [5], базирующееся на использовании ключевых слов, основных понятий, сложившихся в предметной области. Модель состоит из трех групп переменных: характеристик основных понятий, изменения этих характеристик и структурированной неопределенности в эквивалентных уравнениях, которая может быть использована для адаптации и управления.

При исследовании сложных систем важным вопросом является выбор языка описания систем. В качестве языка описания в лингво-комбинаторных моделях используется естественный язык – универсальная знаковая система, позволяющая рассматривать широкий класс систем вплоть до слабо формализованных. Главной особенностью естественного языка является обозначение слов и подразумевание смыслов. Эта трудность преодолевается путем формального введения понятия смысла слов как множителей этих слов в исходной фразе и путем перехода к лингвистическому уравнению, что позволяет

построить исчисление смыслов. Такой подход можно применить к анализу всего корпуса текстов на естественном языке. Это трудоемкая задача по извлечению смыслов для современных компьютеров. Данный подход также можно использовать, опираясь на ключевые слова в конкретной предметной области, что позволит получить новые модели для этой области знаний. В этом случае лингво-комбинаторное моделирование заключается в том, что в конкретной предметной области выделяются ключевые слова, которые объединяются во фразы, на основе которых строятся эквивалентные системы уравнений с произвольными коэффициентами. В частном случае они могут быть дифференциальными уравнениями, и при их исследовании может быть использован известный математический аппарат. Метод лингво-комбинаторного моделирования на практике является эффективным эвристическим приемом при анализе слабо формализованных систем.

Лингво-комбинаторные модели обеспечивают возможность построения исчисления смыслов, которое хорошо реализуемо на компьютерах. Определение смысла включает две важные характеристики – контекстуальность (смыслы вычисляются исходя из контекста) и интенциональность (произвольные коэффициенты позволяют задавать те или иные устремления). В результате задача анализа сложных систем с применением метода лингво-комбинаторного моделирования сводится к исследованию эквивалентных уравнений с произвольными коэффициентами. В работе [7] доказано, число этих произвольных коэффициентов равно числу сочетаний из n по $m + 1$, где n – число переменных, различных слов в исходных фразах, m – число ограничений, наложенных на переменные, описывающих систему, число различных фраз. Из анализа этой формулы вытекают важные выводы. Во-первых, для многомерных систем с числом переменных больше шести имеется максимум в числе произвольных коэффициентов в структуре эквивалентных уравнений. Произвольные коэффициенты используются для управления системой и ее адаптации к окружающей среде. Этот максимум называется феноменом адаптационного максимума [8]. Тогда принцип управления сложными системами заключается в том, что система должна управляться так, чтобы удерживать ее в зоне адаптационного максимума в динамически изменяющихся условиях. Во-вторых, для удержания системы в зоне адаптационного максимума необходимо наложение и снятие ограничений, а также объединение систем в коллектив. Предложенный в 1963 году профессором М.Б. Игнатьевым феномен адаптационного максимума применяется для управления и координации в биологических, социально-экономических и технических системах, а также для других актуальных приложений.

Региональные социально-экономические системы обладают собственной спецификой и относятся к классу слабоструктурированных сложных динамических систем. С точки зрения системного подхода сложность управления региональной безопасностью состоит в том, что регион, как объект управления, представляет собой совокупность тесно взаимосвязанных различных по масштабам и структуре естественных и искусственных подсистем, объединяющих в единое целое экономическую инфраструктуру, социальную сферу, производство, окружающую среду, сферу перераспределения, различные социальные группы и большие коллективы людей [9]. Чем выше степень гармонизации развития этих элементов, тем выше уровень региональной

безопасности, то есть защищенность региона от влияния внутренних и внешних угроз различной природы.

Таким образом, в настоящей работе предлагается расширить область применения метода лингво-комбинаторного моделирования на класс социально-экономических систем для исследования системных связей, закономерностей и тенденций развития региона, определяющих региональную безопасность в условиях неполноты информации для принятия управленческих решений и высокой динамики социально-экономической среды.

2. Структура метода построения лингво-комбинаторных моделей

Математические модели имеются лишь для небольшого числа реальных систем. Системы описываются, прежде всего, с помощью естественного языка. Предлагается способ перехода от описания на естественном языке к математическим уравнениям. Пусть имеется фраза:

$$Word_1 + Word_2 + Word_3 . \quad (1)$$

В этой фразе обозначаются слова и подразумевается только их смысл. В сложившейся структуре естественного языка смысл не обозначается. Ввести понятие смысла можно в следующей форме:

$$Word_1 \cdot Sense_1 + Word_2 \cdot Sense_2 + Word_3 \cdot Sense_3 . \quad (2)$$

Обозначим слова как A_i (от англ. *Appearance*), а смыслы – как E_i (от англ. *Essence*). Тогда уравнение (2) может быть представлено как:

$$\sum_{i=1}^3 A_i E_i = 0 . \quad (3)$$

Уравнения (2) и (3) являются моделями фразы (1). Если взять математическое уравнение $F(x_1, x_2, x_3) = 0$, то получим форму (3) путем дифференцирования этого уравнения. Тогда A_i будут частными производными, а E_i – производными по времени от переменных.

Эта модель является алгебраическим кольцом и можно разрешить уравнение (3) относительно A_i , либо относительно E_i , путем введения третьей группы переменных – произвольных коэффициентов U_s :

$$\begin{aligned} A_1 &= U_1 E_2 + U_2 E_3 \\ A_2 &= -U_1 E_1 + U_3 E_3 \\ A_3 &= -U_2 E_1 + U_3 E_2 \end{aligned} \quad (4)$$

или

$$\begin{aligned} E_1 &= U_1 A_2 + U_2 A_3 \\ E_2 &= -U_1 A_1 + U_3 A_3 , \\ E_3 &= -U_2 A_1 + U_3 A_2 \end{aligned} \quad (5)$$

где U_1, U_2, U_3 – произвольные коэффициенты, которые можно использовать для решения различных задач на многообразии (3). Например, если хотим достигнуть максимум по переменной x_3 , то можем назначить произвольные коэффициенты $U_2 = -bA_1, U_3 = -bA_2$ и тогда получим:

$$\begin{aligned}\frac{dx_1}{dt} &= U_1 A_2 - b A_1 A_3 \\ \frac{dx_2}{dt} &= -U_1 A_1 - b A_2 A_3 . \\ \frac{dx_3}{dt} &= b(A_1 A_1 + A_2 A_2)\end{aligned}\tag{6}$$

Если $b > 0$, то переменная x_3 устойчиво стремится к максимуму, а для манипуляции траекторией остается коэффициент U_1 .

В общем случае, при n переменных и m многообразий, ограничений, число произвольных коэффициентов S будет равно числу сочетаний из n по $m+1$ (таблица):

$$S = C_n^{m+1}, n > m.\tag{7}$$

Число произвольных коэффициентов является мерой неопределенности и адаптивности.

Таблица. Определение числа произвольных коэффициентов S

$n \setminus m$	1	2	3	4	5	6	7	8
2	1							
3	3	1						
4	6	4	1					
5	10	10	5	1				
6	15	20	15	6	1			
7	21	35	35	21	7	1		
8	28	56	70	56	28	8	1	
9	36	84	126	126	84	36	9	1

Структурная стабильность, совокупность устойчивых связей объекта, обеспечивающих его целостность и тождественность самому себе, то есть сохранение основных свойств при различных внешних и внутренних воздействиях, обеспечивается адаптационными возможностями элементами этих систем [10]. В представленных лингво-комбинаторных моделях адаптационные возможности систем определяются числом произвольных коэффициентов в структуре эквивалентных уравнений и наибольшая структурная стабильность достигается в зоне адаптационного максимума, который обнаруживается у различных систем с числом переменных больше шести [5]. Для удержания систем в зоне адаптационного максимума используются различные методы – рост числа переменных, наложение и снятие ограничений, объединение систем в коллективы.

Пусть имеется две системы:

$$\begin{aligned}S_1 &= C_{n_1}^{m_1+1}, \\ S_2 &= C_{n_2}^{m_2+1},\end{aligned}\tag{8}$$

где n_1 и n_2 – число переменных в системах; m_1 и m_2 – число

ограничивающих многообразий, наложенных на переменные системы; S_1 и S_2 соответственно – число произвольных коэффициентов в структурах эквивалентных уравнений.

Тогда путем наложения m_{col} общих ограничивающих многообразий получим объединенную систему – коллектив, который будет характеризоваться формулой:

$$S_{col} = C_{n_1+n_2}^{m_1+m_2+m_{col}+1}. \quad (9)$$

При этом в зависимости от конкретных параметров возможны два случая с точки зрения роста адаптационных возможностей:

- 1) $S_{col} > S_1 + S_2$, когда объединение в коллектив целесообразно и приводит к росту адаптационных возможностей всей системы и исходных систем;
- 2) $S_{col} < S_1 + S_2$, когда адаптационные возможности меньше суммы адаптационных возможностей исходных систем (объединение в коллектив нецелесообразно).

Лингво-комбинаторное моделирование является эффективным инструментом для анализа и синтеза многоуровневых распределенных социально-экономических систем. Такой прием для моделирования распространяется на тексты с большим числом слов и фраз. Построенная система моделей нуждается в исследовании на предмет адекватности объектам моделирования, но в ряде случаев она представляет собой эффективный эвристический способ перехода к компьютерному моделированию [6].

3. Феномен адаптационного максимума

Решение проблем управления безопасностью сложных систем во многом упирается в изучение свойств этих систем и моделирование их поведения. Безопасное функционирование социально-экономических систем в динамически изменяющихся условиях возможно только в зоне так называемого адаптационного максимума.

Если рассматривать систему из n переменных, на которые наложено m ограничивающих многообразий, то в структуре эквивалентных уравнений будет содержаться S произвольных коэффициентов $S = C_n^{m+1}$, $n > m$, которые можно использовать как для управления, так и для приспособления системы к изменениям окружающей среды, с которой она взаимодействует. Чем больше этих коэффициентов, тем выше адаптационные возможности системы. При наложении дополнительных связей число произвольных коэффициентов может как уменьшаться, так и увеличиваться: для $n > 6$ при наложении новых ограничивающих многообразий число произвольных коэффициентов будет сначала возрастать, достигнет максимума, а потом будет убывать. Поэтому наложение ограничивающих многообразий можно рассматривать как способ приспособления системы к изменению окружающей среды. Манипуляции с неопределенными коэффициентами представляют собой настройку (или самонастройку) системы.

Система – целостная совокупность элементов, в которой все элементы настолько тесно связаны между собой, что она выступает по отношению к другим системам и окружающей среде как нечто единое. Сложная система – это система,

в которой проявляется феномен адаптационного максимума, то есть система с числом переменных больше шести.

На рис. 1 представлена схема, где система взаимодействует со средой. Переменные системы x_1, \dots, x_k взаимодействуют с переменными среды y_1, \dots, y_k , а сигналы рассогласования передаются в блок управления. При этом система использует два механизма адаптации:

а) настройка или самонастройка системы с помощью манипуляции произвольными коэффициентами в структуре эквивалентных уравнений системы;

б) обучение или самообучение системы, которые заключаются в наложении новых ограничений на переменные системы.

Кроме этих механизмов адаптации возможны и другие, такие как рост числа переменных системы, размножение, эффективное забывание, ограничение контактов со средой, объединение систем в коллектив и т.д.

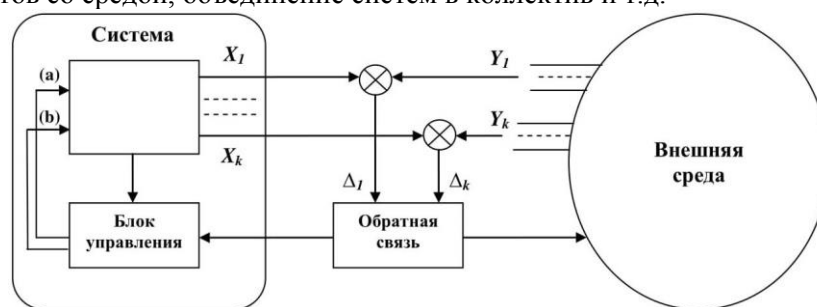


Рис. 1. Модель «среда – сложная система»

Применительно к социально-экономическим системам, которые находятся во взаимодействии с изменяющейся средой, цель управления этими системами заключается в удержании их в зоне адаптационного максимума – если требуется обеспечить безопасность развития системы, и наоборот, вывод из зоны адаптационного максимума – если требуется их дестабилизировать.

Ниже в качестве примера рассмотрим лингво-комбинаторную модель для управления региональной безопасностью.

4. Лингво-комбинаторная модель региона

При построении лингво-комбинаторных моделей региональных социально-экономических систем будем исходить из базовых положений, которые уже сложились в этой области [3, 9].

Если в качестве ключевых слов взять «Население», «Производство», «Финансы», «Наука, образование, инновации», «Внешние связи, пассионарность», «Территория, ресурсная база, логистика», «Окружающая среда и безопасность», то в соответствии с изложенной методикой уравнение региона будет представлено следующим образом:

$$\sum_{i=1}^7 A_i E_i = 0, \quad (10)$$

а эквивалентные уравнения будут иметь вид

$$\begin{aligned}
 E_1 &= U_1 A_2 + U_2 A_3 + U_3 A_4 + U_4 A_5 + U_5 A_6 + U_6 A_7 \\
 E_2 &= U_1 A_1 + U_7 A_3 + U_8 A_4 + U_9 A_5 + U_{10} A_6 + U_{11} A_7 \\
 E_3 &= -U_2 A_1 - U_7 A_2 + U_{12} A_4 + U_{13} A_5 + U_{14} A_6 + U_{15} A_7 \\
 E_4 &= -U_3 A_1 - U_8 A_2 - U_{12} A_3 + U_{16} A_5 + U_{17} A_6 + U_{18} A_7, \\
 E_5 &= -U_4 A_1 - U_9 A_2 - U_{13} A_3 - U_{16} A_4 + U_{19} A_6 + U_{20} A_7 \\
 E_6 &= -U_5 A_1 - U_{10} A_2 - U_{14} A_3 - U_{17} A_4 - U_{19} A_5 + U_{21} A_7 \\
 E_7 &= -U_6 A_1 - U_{11} A_2 - U_{15} A_3 - U_{18} A_4 - U_{20} A_5 - U_{21} A_6
 \end{aligned} \tag{11}$$

где A_1 – характеристика населения, которая включает в себя характеристики качества жизни, демографии, здоровья, рынка труда, образования и др.; E_1 – изменение этой характеристики; A_2 – характеристика производства, включающая оценки промышленного потенциала региона и различных видов деятельности хозяйствующих субъектов по отраслям региональной экономики (научной, производственной, транспортной, торговой и др.); E_2 – изменение этой характеристики; A_3 – характеристика финансов, регионального бюджета, финансовых потоков (доходы и расходы населения и предприятий), ВРП, инвестиционного климата в регионе и др.; E_3 – изменение этой характеристики; A_4 – характеристика научно-образовательного комплекса, научно-инновационного потенциала региона, кадровых потребностей отраслей региональной экономики и др.; E_4 – изменение этой характеристики; A_5 – характеристика внешних связей региона и пассионарности населения и хозяйствующих субъектов с учетом географического положения и региональной специфики, включая миграционный прирост, устремления групп населения (люди обладают свободой выбора при принятии решений и этот выбор является важным, что оценивается путем социологического анализа), оценки социетальной безопасности, а также входящих и выходящих потоков энергии, материалов, информации, финансов; E_5 – изменение этой характеристики; A_6 – характеристика территории и логистики, включая ресурсную базу, транспортную инфраструктуру, наземные и подземные постройки и т.д.; E_6 – изменение этой характеристики; A_7 – характеристика окружающей среды и безопасности, включая оценки экологических рисков, степени гармонизации развития региональных подсистем (устойчивого развития региона) и уровня региональной безопасности в целом; E_7 – изменение этой характеристики; U_1, U_2, \dots, U_{21} – произвольные коэффициенты, которые могут быть использованы для управления и решения различных задач на многообразии (10).

Модель (10)-(11) может быть использована в составе системы поддержки принятия решений по управлению региональной безопасностью в условиях кризисных ситуаций социально-экономического характера [11].

Число блоков в лингво-комбинаторной модели региона может быть различным. С точки зрения точности моделирования чем больше блоков задействуется, тем лучше, но при этом ухудшается наглядность модели, ее восприятие субъектами управления – лицами, принимающими решения. Например, если блок «Окружающая среда и безопасность» поделить на три блока – «экономическая безопасность», «социальная безопасность» и «экологическая безопасность», то число переменных возрастет до девяти. Тогда уравнение региона будет содержать девять переменных

$$\sum_{i=1}^9 A_i E_i = 0. \quad (12)$$

На рис. 2 представлена структура поддержки принятия решений в сфере управления региональной безопасностью с применением комплексной лингво-комбинаторной модели региона.

При моделировании региона важно рассматривать всю иерархию систем, из которых этот регион состоит. Для этого может быть использован принцип декомпозиции, основанный на применении «семиблочной модели», предложенный в [6, 7]. При этом структура блоков для каждой из региональных подсистем будет однотипной, а изменяться будет содержание только отдельных блоков. Однотипность подмоделей позволяет гибко производить анализ и синтез такой сложной системы как регион.

В статистических материалах по городам и регионам и по странам в целом имеются почти все данные, необходимые для запуска лингво-комбинаторной модели региона средствами современной вычислительной техники. Другие данные для оценки, например, пассионарности, можно получить из социологических опросов.



Рис. 2. Поддержка принятия управленческих решений в сфере региональной безопасности

Стремительное развитие информационных технологий позволяет поставить вопрос об обязательном предварительном моделировании последствий от принимаемых управленческих решений в сфере обеспечения безопасности региональных социально-экономических систем. Это позволит избежать многих ошибочных решений и обеспечить превентивное управление объектами критической инфраструктуры региона в условиях региональных кризисных ситуаций.

Заключение

Лингво-комбинаторные модели – это новый класс моделей слабо формализованных систем. Развитие формального аппарата построения лингво-комбинаторных моделей систем различной природы и его приложения позволили выявить новое эмерджентное свойство сложных систем – феномен адаптационного максимума. Наличие феномена адаптационного максимума в жизненном цикле социально-экономических систем позволяет объяснить природу кризисных ситуаций, которые периодически происходят в этих системах. Отсюда вытекает принцип управления безопасностью социально-экономических систем в кризисных ситуациях – системы должны управляться так, чтобы удерживать их в зоне адаптационного максимума в динамически меняющихся условиях обстановки. Глубина кризисных ситуаций определяется отклонением от зоны адаптационного максимума. Эту глубину можно существенно уменьшить, осуществляя непрерывный проблемный мониторинг состояния социально-экономических систем и принимая соответствующие антикризисные меры. Наличие произвольных коэффициентов и возможность расширения модели позволяют гибко настраивать модель для моделирования сложных социально-экономических объектов.

В ходе исследований получены следующие основные результаты:

1. Показана применимость метода лингво-комбинаторного моделирования для формализации и компьютерного моделирования социально-экономических систем и процессов.

2. Разработана лингво-комбинаторная модель поддержки управления региональной безопасностью. Разработка осуществлялась при выполнении государственного задания в рамках темы НИР ИИММ КНЦ РАН «Модели и методы конфигурирования адаптивных многоуровневых сетевых систем управления региональной безопасностью в Арктической зоне Российской Федерации» (№ 0226-iiimm-2017-14-08).

3. Предложено развитие теории ситуационного управления безопасностью и расширена область применения формального аппарата лингво-комбинаторного моделирования на класс социально-экономических систем для исследования системных связей, закономерностей и тенденций развития региона, определяющих его безопасность и кризисные ситуации, в условиях неполноты информации для принятия управленческих решений и высокой динамики социально-экономической среды. Результат получен в рамках реализации проекта РФФИ № 18-29-03022-мк.

В условиях децентрализованного управления безопасностью социально-экономических систем при каждом конкретном применении метода лингво-комбинаторного моделирования необходимо осуществлять верификацию модели, проверять ее на соответствие поведению реального объекта, проводить анализ реализуемого в модели алгоритма управления на сходимость, оценивать ошибку рассогласования связующих входов при синтезе структуры управления, определять оптимальную конфигурацию параметров модели для решения задачи стабилизации объекта в зоне адаптационного максимума и задачи координации взаимодействия моделируемых подсистем на разных уровнях управления. Решение этих задач является предметом дальнейших исследований.

Литература

1. Цыгичко, В.Н. Безопасность критических инфраструктур / В.Н. Цыгичко, Д.С. Черешкин, Г.Л. Смолин. – М.: Красанд, 2018. – 200 с.
2. Цыгичко, В.Н. Управление рисками в организационных системах / В.Н. Цыгичко, Д.С. Черешкин, Г.Л. Смолин. – М.: LAP Lambert, 2018. – 112 с.
3. Шульц, В.Л. Сценарный анализа в управлении геополитическим информационным противоборством / В.Л. Шульц, В.В. Кульба, А.Б. Шелков, И.В. Чернов. - М.: Наука, 2015.- 542 с.
4. Шумов, В.В. Государственная и общественная безопасность: Моделирование и прогнозирование / В.В. Шумов. - М.: ЛЕНАНД, 2016. - 144 с.
5. Ignatyev, M. Linguo-combinatorial simulation of complex systems / M. Ignatyev // Journal of Mathematics and System Science. – 2012. - vol.2. – no. 1. - P. 58-66.
6. Игнатъев, М.Б. Просто Кибернетика / М.Б. Игнатъев. – СПб: «Страта», 2016. - 248 с.
7. Игнатъев, М.Б. Кибернетическая картина мира. Сложные киберфизические системы. 3-е изд. / М.Б. Игнатъев. – СПб.: ГУАП, 2014. - 472 с.
8. Игнатъев, М.Б. Голономные автоматические системы / М.Б. Игнатъев. – М.-Л.: АН СССР, 1963. – 204 с.
9. Маслобоев, А.В. Информационное измерение региональной безопасности в Арктике / А.В. Маслобоев, В.А. Путилов.- Апатиты: КНЦ РАН, 2016.- 222 с.
10. Бейдер, Р. Атомы в молекулах / Р. Бейдер. – М.: Мир, 2001. – 465 с.
11. Маслобоев, А.В. Система информационно-аналитической поддержки сетцентрического управления региональной безопасностью / А.В. Маслобоев // Информационные ресурсы России.- 2016.- №3(151).- С. 25-31.

Сведения об авторах

Маслобоев Андрей Владимирович – д.т.н., доцент, ведущий научный сотрудник
e-mail: masloboev@iimm.ru

Andrey V. Masloboev – Dr. of Tech. Sc., associate professor, leading research fellow

Путилов Владимир Александрович – заслуженный деятель науки РФ, д.т.н., профессор,
научный руководитель
e-mail: putilov@iimm.ru

Vladimir A. Putilov – honoured science worker of the Russian Federation, Dr. of Tech. Sc.,
professor, research supervisor

Игнатъев Михаил Борисович – доктор технических наук, профессор, профессор
кафедры вычислительных систем и сетей
e-mail: kira@robotek.ru

Mikhail B. Ignatyev – Dr. of Tech. Sc., professor, professor of Computer system and network
department

УДК 004.832

DOI: 10.25702/KSC.2307-5252.2018.10.21-35

А.А. Зуенко¹, О.В. Фридман¹

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

ПРИМЕНЕНИЕ МЕТОДОВ ПРОГРАММИРОВАНИЯ В ОГРАНИЧЕНИЯХ В ЗАДАЧАХ ИНТЕЛЛЕКТУАЛЬНОГО ПЛАНИРОВАНИЯ*

Аннотация

В статье рассмотрены особенности существующих подходов к решению задач интеллектуального планирования с применением методов программирования в ограничениях. Проведен анализ современных методов решения задач классического планирования, составления расписаний, интеллектуального планирования.

Ключевые слова:

программирование в ограничениях, планирование, составление расписаний, интеллектуальное планирование

A.A. Zuenko, O.V. Fridman

APPLICATION OF CONSTRAINT PROGRAMMING METHODS IN INTELLECTUAL PLANNING PROBLEMS

Abstract

The article discusses the features of existing approaches to solving problems of intelligent planning using constraint programming techniques. The analysis of modern methods for solving problems of classical planning, scheduling, intelligent planning has been carried out.

Keywords:

constraint programming, planning, scheduling, intelligent planning

Введение

Использование подходов и методов искусственного интеллекта (ИИ) позволяет решать многие прикладные задачи, такие как задачи теории расписаний, задачи проектирования экспертных систем и систем поддержки принятия решений, доказательство теорем, задачи тестирования электронных схем, обработки изображений.

Теория удовлетворения ограничений предлагает удобный аппарат и простую формальную схему для представления и решения комбинаторных задач искусственного интеллекта. Целью решения задачи удовлетворения ограничений является нахождение значений переменных, удовлетворяющих определенным ограничениям. Процесс рассуждений на ограничениях сводится к поэтапному усечению изначально заданных областей определения переменных. Любой метод удовлетворения ограничений должен проектироваться особым образом и состоять из двух основных частей: части, реализующей поиск, и части, реализующей вывод на ограничениях. Как правило, под выводом на ограничениях понимается процесс сокращения размерности пространства поиска,

* Работа выполнена при финансовой поддержке РФФИ (проекты №№ 16-07-00377а, 18-07-00615а).

обеспечивающий «сужение» доменов переменных, упрощение ограничений и т.п. и, при этом, имеющий низкую вычислительную сложность (оценивается полиномом низкой степени). Алгоритмы, реализующие вывод на ограничениях, называются алгоритмами-распространителями или алгоритмами-пропагаторами. Методы, реализующие стратегии интеллектуального поиска, в отличие от методов вывода, имеют существенную вычислительную сложность, поскольку сопряжены с перебором вариантов гипотез о возможных значениях тех или иных переменных. В отличие от методов динамического программирования, например, метода перемножения числовых матриц (matrix multiplication method), в качестве стратегий поиска в рамках парадигмы программирования в ограничениях, в основном, используются различные варианты информированного поиска в глубину с возвратами, а также методы локального поиска. Особенности архитектуры систем программирования в ограничениях создают предпосылки для унификации совместной обработки разнородной (количественной и качественной) информации (ограничений). Кроме того, технология программирования в ограничениях позволяет организовать процедуры вывода с недоопределенными параметрами, что дает возможность моделировать и аспекты неопределенности, и недетерминированность действий исполнителя, ограниченную конечным перечнем возможных альтернатив.

Теперь перейдем к описанию задач, при решении которых применяются (или могут быть применены) методы программирования в ограничениях.

Интеллектуальное планирование – это область искусственного интеллекта, которая в настоящее время вызывает значительный интерес [1]. В планировании объединяются два основных направления развития искусственного интеллекта – поиск и логика. Перекрестное обогащение идеями, взятыми из этих двух областей, привело не только к повышению производительности, но и к расширению использования планировщиков в производственных приложениях.

Следует различать два направления исследований: интеллектуальное планирование (AI-planning) и составление расписаний (scheduling). В рамках настоящей статьи будем придерживаться точки зрения, принятой в AI-community (сообществе искусственного интеллекта), что задача составления расписаний является одной из разновидностей задач интеллектуального планирования.

Составление расписаний (scheduling) – это проблема назначения множества задач множеству рабочих ресурсов, на которые (как на задачи, так и на ресурсы) накладывается ряд ограничений. В качестве примеров ограничений можно привести предельный срок выполнения задачи (deadline), количество рабочих ресурсов, ограничения предшествования на множестве задач, приоритеты задач и т.п. Таким образом, составление расписаний – это распределение (атомарных) задач по исполнителям. Расписание определяет, кто чем занимается (как правило, с привязкой ко времени). То есть здесь действия (задачи), которые нужно выполнить, заданы заранее. Нужно отыскать лишь их порядок и распределение по рабочим ресурсам.

В отличие от составления расписаний, планирование заключается в отыскании действий, необходимых для решения задачи, и выстраивании этих действий в последовательность, приводящую к решению. То есть нужно найти не только порядок, но и сами действия, которые позволят решить задачу в целом.

В значительной степени, работа по интеллектуальному планированию сосредоточена на задачах, которые включают каскадные уровни выбора действия

со сложными логическими взаимосвязями между действиями. Фактически, планирование – это нахождение способа достижения цели при заданных возможностях. Задача планирования заключается в поиске последовательности действий, применение которой в начальном состоянии модели мира, приведет к такому состоянию, в котором достигается заранее заданная цель. При решении задач планирования приходится держать под контролем комбинаторный взрыв. Если в проблемной области имеется p примитивных высказываний, то количество состояний становится равным 2^p . В сложных проблемных областях величина p может стать весьма значительной.

Выполнение всякого плана окружено контекстом, в котором этот план выполняется. Модель контекста, в котором выполняются действия, называется моделью мира (world model). Модель мира может включать статическую и динамическую составляющие. К статическим элементам относят модели объектов мира и отношений между ними. Динамическим элементом является модель изменения мира по своим внутренним законам, не зависящим от воздействия со стороны субъекта, исполняющего план. Будем называть динамическую составляющую динамикой мира (world dynamics).

Действия плана и динамика мира осуществляют преобразование модели мира (как правило, только статической составляющей). Про изменение модели мира говорят, что модель мира переходит в новое состояние (state). Состояние модели мира в момент, непосредственно предшествующий моменту начала выполнения плана, называется начальным состоянием (initial state) задачи планирования. Большинство современных планировщиков при построении плана оперируют именно понятием состояние, а не модель мира. Формально, состояние может быть представлено различными способами. Конструктивным элементом для процесса планирования является действие. Действие (action) определяет, какие изменения в модели мира произойдут, если действие будет выполнено. В качестве модели действия в интеллектуальном планировании, обычно, используется сущность, описывающая вид деятельности, условия, когда это действие выполнимо, и эффекты, которые это действие производит.

В зависимости от условий, в которых происходит процесс планирования, выделяют несколько разновидностей сред планирования (planning environment). Они определяются рядом свойств исполнителя и мира, в котором он оперирует. Вот эти свойства:

- Наблюдаемость мира. Если это так, то исполнитель знает все о свойствах (но не законах) мира в каждый момент времени. То есть он знает все об объектах, имеющихся в мире, и отношениях между ними. Кроме того, это свойство говорит о том, что мир конечен.
- Статичность мира. Это свойство отражает возможность (или невозможность) мира меняться независимо от действий исполнителя. Или, иными словами, свойство показывает, существует ли динамика мира.
- Детерминированность действий исполнителя. Если действия детерминированы, то исполнитель всегда точно знает, каков будет результат выполнения каждого действия.
- Дискретность или протяженность действий во времени. Если действия дискретны, то они не имеют длительности, и свойства мира во время выполнения действий не определены. О свойствах мира можно говорить, только если ни какое из действий не выполняется.

- Возможность параллельного исполнения действий. Если параллельность допустима, то каков коэффициент параллелизма, т.е. какое максимальное количество действий может выполняться одновременно. Правда, с коэффициентом параллелизма, отличным от бесконечности, задача ставится редко.

Все вышесказанное позволяет утверждать, что применение методов программирования в ограничениях в задачах интеллектуального планирования позволит решать подобные задачи более эффективно за счет повышения эффективности процедур вывода на ограничениях. В рамках технологии программирования в ограничениях имеются предпосылки унификации совместной обработки числовых и нечисловых ограничений. Технология программирования в ограничениях позволяет организовать процедуры вывода с недоопределенными параметрами, что дает возможность моделировать и аспекты неопределенности, и недетерминированность действий исполнителя, ограниченную конечным перечнем возможных альтернатив.

Далее в настоящей работе рассматриваются применение методов программирования в ограничениях в задачах классического планирования, составления расписаний, попытки их развития и сопоставление с методами искусственного интеллекта.

Методы классического планирования и попытки их развития

Значительная часть работ по интеллектуальному планированию, выполненных за последние 30 лет, слабо укладывается в парадигму, которую можно назвать классическим планированием. Тем не менее, чтобы понять современные тенденции в интеллектуальном планировании, остановимся на особенностях классического планирования более подробно.

В классическом планировании результатом процесса поиска является достижение заданного набора целей, обычно выражаемых в виде набора положительных и отрицательных литералов исчисления высказываний. Начальное состояние мира, называемое начальными условиями, также выражается как набор литералов. Возможные действия характеризуются с помощью так называемых STRIPS-операторов (STRIPS (Stanford Research Institute Problem Solver) – это автоматический планировщик, разработанный Ричардом Файксом и Нильсом Нилсоном в 1971. В последующем слово STRIPS стало также использоваться для обозначения формального языка, описывающего входные данные этого планировщика). STRIPS-оператор представляет собой параметризованный шаблон, содержащий набор предварительных условий, которые должны быть истинными прежде, чем действие может быть выполнено, а также содержит набор изменений или эффектов, которые произойдут в мире в результате выполнения действия.

В языке представления классических планировщиков STRIPS [2] применяются описания действий в терминах их предусловий и результатов, а также описания начальных и целевых состояний в виде конъюнкций положительных литералов. В языке ADL (ADL – Action Description Language – язык описания действий) [3] некоторые ограничения языка STRIPS ослаблены и допускается использование дизъюнкции, отрицания и кванторов.

Тем не менее, существует ряд более серьезных ограничений в представлениях STRIPS и ADL. Во-первых, это атомарное время. В указанных представлениях нет явной модели времени. Нельзя указать продолжительность действия или указать временные ограничения для целей или действий. Фактически, действия моделируются так, как если бы они были мгновенные и бесперебойные, поэтому нет никаких условий для одновременных действий или представления внешних или экзогенных событий. Во-вторых, нет возможности указывать потребности в ресурсах или потребление ресурсов. В-третьих, нет возможности моделировать неопределенность. Первоначальное состояние мира, как и результаты действий, должны быть достоверно известны. Наконец, единственными типами целей являются цели достижения. Невозможно указать цель, которая была бы связана с поддержанием состояния или достижением состояния к крайнему сроку. Отчасти это происходит из-за того, что нет явной модели времени. Также нет возможности указывать цель, которая предполагает оптимизацию. В классическом планировании оптимизация обычно считается несущественной – достаточно просто найти допустимый план.

Для устранения некоторых из перечисленных недостатков представления STRIPS и/или ADL были расширены. В частности, в работе [4] были разработаны более богатые модели времени в рамках представления STRIPS и освещены аспекты представления и использования различных типов ресурсов, а в работе [5] рассматривались возможности учета различных форм неопределенности. В [6] введены более интересные типы целей. Все эти усовершенствования в языках представления задач планирования требуют методов, расширяющих возможности классических методов планирования. Однако эти расширения, как правило, оказывают значительное негативное влияние на производительность.

Для решения задач классического планирования был разработан ряд различных методов. Здесь мы приведем краткий обзор наиболее важных подходов, перечислим их преимущества и недостатки.

До недавнего времени большинство классических работ по планированию сосредотачивалось на построении планов путем обратного поиска или поиска от цели (Goal-directed Planning). Однако, если действия упорядочены только частично, необходимы дополнительные средства, чтобы убедиться, что неупорядоченные действия не мешают друг другу. Для планирования с учетом отношения частичного порядка на множестве действий были разработаны POCL (Partial Order Causal Link) – планировщики [1]. Несмотря на все усилия в области целенаправленного классического планирования (в частности, планирование POCL), эти планировщики не имели большого практического успеха. Хотя коэффициент ветвления обычно ниже при поиске в обратном направлении от целей, он по-прежнему достаточно велик. Как следствие, успех целенаправленного планирования сильно зависит от используемых эвристик.

В 1995 году Блум и Фурст представили систему планирования, названную Graphplan [7], которая использует совсем другой подход для поиска планов. Основная идея состоит в том, чтобы выполнить анализ достижимости, чтобы исключить многие комбинации и последовательности действий, которые несовместимы друг с другом. Как и при планировании POCL, технология Graphplan была расширена для поддержки операторов с кванторами и других особенностей языка ADL [8]. Были предприняты попытки расширить Graphplan для работы с неопределенностью [9, 10], но эти усилия до сих пор не доказали

свою эффективность на практике. Также известны работы, направленные на преодоление описанных ограничений, связанных с обработкой времени [11] и вещественных величин в Graphplan [12]. Более подробное введение в Graphplan и расширения метода Graphplan можно найти в [13].

Изложенное показывает, что методы интеллектуального планирования напрямую зависят от возможностей языков представления задач планирования. Помимо представления задач планирования на языках типа STRIPS и ADL планирование может осуществляться по принципу доказательства некоторой теоремы в рамках ситуационного исчисления. В подобной теореме утверждается, что при наличии начального состояния и аксиом состояния-преемника, которые описывают результаты действий, цель будет истинной в ситуации, которая является результатом некоторой последовательности действий. В ранний период развития искусственного интеллекта данный подход считался слишком неэффективным для того, чтобы с его помощью можно было находить интересные планы. Проведенные в последнее время разработки в области эффективных алгоритмов формирования рассуждений для пропозициональной логики привели к возрождению интереса к планированию с помощью логических рассуждений.

Близкий подход основан на проверке выполнимости логического высказывания, а не на доказательстве теорем [13]. Планировщики, основанные на проверке выполнимости, способны обрабатывать крупные задачи планирования.

В алгоритме SATplan [14] задача планирования преобразуется в пропозициональные аксиомы и после этого к ним применяется алгоритм проверки выполнимости для поиска модели, соответствующей действительному плану.

Основная идея планирования как выполнимости – угадать длину плана, перевести задачу планирования в набор пропозициональных формул и попытаться решить возникающую проблему выполнимости (SAT). Если формулы не удовлетворяются, то длина плана увеличивается, и процесс повторяется. На сегодняшний день изучено множество различных схем кодирования, но основная идея в большинстве этих схем состоит в том, чтобы иметь пропозициональную переменную: а) для каждого возможного действия на каждом этапе; б) для каждого возможного утверждения на каждом шаге.

Каждая переменная действия указывает на наличие или отсутствие действия на определенном этапе плана. Каждая переменная утверждения указывает, истинно ли это утверждение на этом этапе плана. Лучшие планировщики SAT и планировщики на основе Graphplan имеют очень схожие характеристики. Они значительно превосходят POCL-планировщики на большинстве тестовых примеров. Подобно планировщикам POCL и Graphplan, планировщики SAT могут быть расширены, для поддержки операторов с кванторами и т.п. Фактически это расширение влияет только на процесс перевода задачи, а не на механизм решения. Вещественные величины, представляют более серьезные трудности. Вольфман [15] обрабатывает вещественные величины, используя линейное программирование в сочетании с методами планирования SAT.

Перечислим наиболее серьезные недостатки подхода представления задач планирования в виде задачи SAT. Во-первых, количество переменных и утверждений может быть очень большим, потому что все возможные действия и

утверждения явно представлены для каждой дискретной временной точки. В результате SAT-планировщики часто требуют огромных объемов памяти (гигабайт) для представления задач сравнительно небольшой размерности. Во-вторых, описанная выше кодировка ограничивается дискретным временем и, следовательно, не поддерживает представление действий, которые имеют разную продолжительность или связаны временными ограничениями.

Зачастую описанные направления интеллектуального планирования (на основе графа планирования и на основе логических методов) интегрируются, образуя новые группы методов. Например, метод, используемый в системе Blackbox, состоит в преобразовании графа планирования в выражение КНФ, а затем извлечении плана с использованием решателя задач SAT. Такой подход обладает более высокой производительностью, чем SATplan, и причиной этого, скорее всего, является то, что в графе планирования уже устранены многие невозможные состояния и действия из рассматриваемой задачи. Кроме того, такой подход действует лучше по сравнению с алгоритмом Graphplan, по-видимому, из-за того, что поиск условий выполнимости, подобный алгоритму WalkSAT, характеризуется гораздо большей гибкостью, чем ограниченный поиск с возвратами, используемый в алгоритме Graphplan.

Таким образом, большинство классических методов планирования не может представлять или рассуждать о ресурсах, количественных (вещественнозначных) величинах, непрерывном времени. Большинство методов также игнорирует оптимизацию. Были предприняты попытки расширить методы классического планирования для обработки ресурсов [16] и вещественнозначных величин, а также обеспечить поддержку критериев оптимизации [12, 15]. Предпринимались попытки расширить методы планирования для обработки временных ограничений [17].

Выход за рамки классического планирования

Далее кратко опишем некоторые исследования, которые нацелены на преодоление перечисленных недостатков классического планирования.

Отметим такое направление исследований в области интеллектуального планирования как планирование иерархической сети задач (HTN Planning). В HTN-планировании первоначальный план, который описывает задачу, рассматривается как описание на очень высоком уровне абстракции. Планы уточняются путем декомпозиции действий. Идея использовать подобное структурированное описание для повышения эффективности процедур планирования оказалась очень продуктивной. Нельзя отрицать, что в большинстве практических систем планирования использовались методы HTN [18]. Например, система O-Plan [1], в которой планирование сочеталось с составлением расписаний, использовалась для разработки производственных планов в компании Hitachi.

Тем не менее, многие исследователи недовольны HTN-планированием, потому что оно ближе к «программированию» конкретного приложения, чем к предоставлению декларативного описания доступных действий и использованию общих методов для планирования. Авторам неизвестна всеобъемлющая обзорная статья, в которой описываются различные системы и методы HTN-планирования. Введение в HTN-планирование можно найти в [19].

Теперь перейдем к описанию одной из технологий планирования, предназначенной для работы с неопределенностью. На протяжении многих лет исследователи в области исследований операций и принятия решений моделировали последовательные решения, используя Марковские процессы принятия решений (MDP). В принципе, MDP – это пространство состояний, в котором переходы между состояниями являются вероятностными по своей природе. В частности, они оказались полезными в задачах навигации робота, где есть неопределенность в расположении робота и ориентации после перемещения [20]. Размер пространства состояний по-прежнему является значительным препятствием для более широкого применения MDP-методов. Справедливости ради следует отметить, что альтернативные подходы (расширения методов POCL, Graphplan и SAT) для планирования с учетом неопределенности также не слишком хорошо зарекомендовали себя на практике. Перечислим некоторые другие недостатки MDP-методов. Во-первых, MDP предполагает, что после выполнения действия с неопределенным результатом агент может наблюдать полученное состояние. Во-вторых, в представлении MDP нет явной модели времени. Действия моделируются так, как если бы они были дискретными, мгновенными и бесперебойными.

Как уже упоминалось, наличие вещественных величин и ограничений может вызвать большие трудности при планировании. В общем случае вещественные ограничения на действия могут быть нелинейными или могут включать производные. Простейший подход к работе с такими величинами при планировании состоит в том, что их следует игнорировать до тех пор, пока значения вещественных переменных не будут точно известны. Этот пассивный подход использовался в нескольких системах планирования [21], но он довольно слабый, поскольку обнаруживает трудности в конце процесса планирования и не дает никаких указаний при фактическом выборе соответствующего действия.

Было построено несколько планировщиков, которые проверяют согласованность наборов вещественных ограничений еще до того, как все переменные известны. Несколько планировщиков использовали методы линейного программирования (LP) для управления этими ограничениями [4, 15].

Также известен планировщик LPSAT [15]. LPSAT использует методы LP совместно с методами планирования SAT. Задача планирования кодируется как проблема SAT, за исключением того, что вещественнозначные предусловия и результаты заменяются булевыми триггерными переменными в кодировке SAT. В ходе планирования, если триггерная переменная становится истинной, соответствующее ограничение передается инкрементному Simplex-решателю. Если Simplex-решатель сообщает, что набор ограничений несовместим, SAT отступает чтобы изменить одну или несколько триггерных переменных. Производительность LPSAT весьма перспективна, но метод имеет те же недостатки, что и основной SAT-подход.

Другим подходом к обработке вещественнозначных величин является представление задачи планирования как смешанной целочисленной линейной задачи (ILP). В работах [22] обсуждаются методы перевода задач планирования в проблемы ILP. До сих пор методы ILP еще не конкурировали с методами планирования SAT, поскольку этап релаксации LP гораздо более дорогостоящий. Основным преимуществом подхода ILP является то, что все аксиомы, включая вещественные ограничения, переводятся в равенства и неравенства.

После того, как были освещены некоторые технологии планирования, нацеленные на поддержку неопределенности, вещественных величин, кратко затронем подход, предназначенный для работы с непрерывным временем. Рассмотрим системы [23, 24], которые объединили идеи планирования POCL и используют интервальное представление для действий и предложений. Данные системы опираются на методы удовлетворения ограничений для представления и управления отношениями между интервалами. Мы будем называть этот подход – основанный на ограничениях интервальный подход (CBI approach).

Аллен представил набор из семи основных интервальных отношений (и их обратных), которые могут быть использованы для описания отношений между интервалами. Аллен описывает мир, утверждая, что предложения занимают интервалы времени. Аналогичным образом, действия и события описываются как некоторые интервалы. Ограничения между интервалами описывают отношения между действиями (или событиями).

Планировщик работает в обратном направлении от целей, добавляя к плану новые действия, которые, в свою очередь, вводят новые подцели в силу интервальных ограничений. Ограничения на интервалах затем переводятся в простые равенства и неравенства ограничения между конечными точками. Проверка вывода и согласованности в этой сети ограничений часто может быть выполнена с использованием быстрых алгоритмов дуговой совместности [25]. Планировщики CBI можно рассматривать как динамические механизмы удовлетворения ограничений – планировщик поочередно добавляет новые ограничения для сети, затем использует методы удовлетворения ограничений для распространения эффектов этих ограничений и для проверки согласованности.

Теперь кратко опишем исследования в области составления расписаний.

Составление расписаний и методы искусственного интеллекта

Составление расписаний не получало серьезного внимания в сообществе искусственного интеллекта до начала 1980-х годов, пока не начались разработки систем составления расписаний, управляемых ограничениями [26]. С того времени все большее число исследователей ИИ работает в этой области. Общая концепция составления расписаний в ИИ заключается в том, что это особый случай планирования, в котором действия уже выбраны, а требуется определить их допустимый порядок. Два хорошо известных учебника по исследованию операций [27, 28] определяют проблему составления расписания как проблему назначения ограниченных ресурсов заданиям с учетом времени для оптимизации одной или нескольких целей. В этом определении есть три важных момента:

- В основе задач составления расписаний лежат рассуждения о времени и ресурсах.
- Проблема составления расписания – это почти всегда проблема оптимизации.
- Проблема составления расписания также связана с выбором. Часто он не просто ограничивается выбором порядка заданий, но включает выбор того, какие ресурсы использовать для каждого задания. Для данного задания могут быть доступны несколько альтернативных ресурсов с разными затратами и / или продолжительностью.

В рамках теории расписаний принято выделять следующие разделы:

- сетевое планирование или построение расписания для проекта, Project scheduling (PS);
- календарное планирование или построение расписания для приборов, Machine scheduling (MS);
- составление временных таблиц (Time Tabling);
- доставка товаров в магазины (Shop-Floor Scheduling);
- составление расписаний движения транспортных средств (Transport Scheduling);
- циклические расписания для транспортных средств (Vehicle Routing)

Остановимся на двух основных классах задач: построение расписания для проекта и построение расписания для приборов.

В задаче построения расписания для проекта (Resource-Constrained Project Scheduling Problem, RCPSP – построения расписания выполнения работ проекта с учетом отношений предшествования и ограничения на ресурсы) под проектом понимают совокупность взаимосвязанных действий, направленных на достижение конкретных целей.

В задаче RCPSP необходимо построить оптимальное расписание проекта (выполнения работ проекта) с учетом сетевого графика (отношений предшествования между работами) и с учетом необходимых/доступных ресурсов, при котором будет оптимизирована некоторая целевая функция [29].

В задачах построения расписания для приборов (в задачах календарного планирования) выделяются несколько подтипов задач: задачи для одного прибора, задачи для параллельных приборов, задачи цеха (Shop Scheduling), Job-shop, Flow-shop, Open-shop. Отличие задач построения расписания для проекта и расписания для приборов заключается в том, что при построении расписания для проекта подразумевается одновременное участие нескольких исполнителей. Календарное планирование – это процесс составления и корректировки расписания, в котором работы, выполняемые различными исполнителями, связываются между собой во времени и с возможностями их обеспечения различными видами материально-технических и трудовых ресурсов. При календарном планировании обязательно должно учитываться соблюдение заданных ограничений (продолжительность работ, лимиты ресурсов) и оптимальное распределение ресурсов.

В теории расписаний для решения перечисленных выше задач используются следующие методы: эвристические алгоритмы, метаэвристические методы, метод динамического программирования, графический метод, метод ветвей и границ.

Эвристические алгоритмы – алгоритмы, основанные на правдоподобных, но не обоснованных математически предположениях о свойствах оптимального решения задачи. Фактически в эвристическом алгоритме учитывается одно или несколько свойств оптимального решения, на основе которых производится сокращение перебора возможных решений.

Метаэвристические методы. Примерами таких метаэвристических методов являются генетический алгоритм и метод муравьиных колоний [30, 31].

Метод динамического программирования получил большое распространение при решении некоторых задач дискретной оптимизации [29].

Графический метод – это модификация метода динамического программирования [29]. Преимущества графического метода заключаются в следующем:

- с помощью графического метода можно решать примеры с нецелочисленными параметрами;
- известно, что для некоторых задач графический метод имеет полиномиальную трудоемкость, в то время как исходный алгоритм метода динамического программирования – псевдополиномиальную.

Метод ветвей и границ (B&B) широко используется для нахождения точного (оптимального) решения задач дискретной оптимизации. Чтобы построить алгоритм, основанный на методе Ветвей и Границ, необходимо определить способ ветвления и способы вычисления нижних и верхних оценок.

В отличие от задач планирования задачи составления расписаний включают только небольшой фиксированный набор вариантов, в то время как проблемы планирования часто связаны с каскадными множества вариантов, которые взаимодействуют сложными способами. В задаче составления расписаний априорно имеется набор заданий, хотя некоторые из них могут быть необязательными. В проблеме планирования обычно неизвестно, сколько заданий или действий потребуется для достижения цели.

Существует обширная литература по теории расписаний в рамках теории исследования операций [27, 28]. Отличие большинства работ по интеллектуальному планированию от работ в области исследования операций заключается в том, что работы в рамках ИИ имеют тенденцию сосредотачиваться на общих представлениях и методах, которые охватывают ряд различных типов задач планирования. Напротив, работы по исследованию операций часто фокусируются на разработке оптимизированных методов для конкретных классов задач составления расписаний, таких как open-shop, flow-shop, job-shop, release dates и т.п.. В качестве примера рассмотрим [32]. В этой работе задачи составления расписаний классифицируются в соответствии с примерно 10 функциями, каждая из которых имеет от 2 до 10 значений. Конкретные подходы к решению многих из этих классов задач часто основаны на специализированных представлениях и алгоритмах.

В ИИ наиболее распространенным подходом к решению задачи составления расписаний является представление ее как задачи удовлетворения ограничений и использование общих методов удовлетворения ограничений [26, 27]. Для моделирования задач составления расписаний в виде CSP применяются два основных подхода, отличающиеся способом кодирования исходной информации и требованиями к решениям. В рамках первого подхода требуется назначить время начала каждому заданию, чтобы все заданные временные и ресурсные ограничения были удовлетворены. Второй подход предполагает, что в процессе решения между заданиями будут установлены отношения частичного порядка таким образом, чтобы все имеющиеся временные и ресурсные ограничения удовлетворялись.

Большая часть первоначальных работ, где составление расписаний рассматривалось в качестве задачи CSP, была выполнена в рамках первого подхода, и для некоторых приложений он по-прежнему является предпочтительным представлением. Однако существуют также недостатки, возникающие в результате того, что появляется множество вариантов, зависящих

от количества временных меток. Во-первых, набор возможных вариантов излишне велик, поскольку реальное количество вариантов значительно меньше, чем набор всех возможных присваиваний. Это затрудняет поиск решений из-за огромного размера пространства поиска. Во-вторых, подход зависит от дискретизации времени, что делает необходимым определять атомарные временные интервалы, прежде чем задача может быть решена. Хуже того, размер представления зависит от способа дискретизации времени – измеряется ли время в часах, минутах, секундах.

Второй подход к представлению задач составления расписаний в виде задач CSP основан на введении для каждой пары задач двух переменных. Если обеим этим переменным присваивается значение «ложь», то две задачи могут перекрываться во времени; присвоение «истины» для обеих переменных запрещено. Данный подход к представлению задач CSP имеет большое значение для устранения изложенных выше недостатков. В частности, результирующее пространство поиска в этом случае значительно меньше, поиск осуществляется на булевых переменных. Данное представление задачи не зависит от способа дискретизации времени, так как существуют алгоритмы, которые могут отслеживать диапазон времен начала заданий без использования дискретизированных временных точек (см., например, [25]). Следовательно, минимальная длина графика может быть определена непосредственно на основе заданного частичного порядка.

В последние годы многие исследователи ИИ изучали задачи выполнимости в качестве альтернативы общей формулировке задачи удовлетворения ограничений. Задача выполнимости является задачей удовлетворения ограничений, где каждая переменная является булевской переменной и каждое ограничение имеет форму дизъюнкции литералов, каждый из которых представляет переменную или отрицание переменной.

Работа в этой области привела к разработке различных быстрых и эффективных алгоритмов выполнимости [33]. Однако, было выявлено, что затраты на представление временных ограничений в рамках этого подхода нивелируют любые преимущества, обеспечиваемые более быстрыми алгоритмами [34].

До этого задача составления расписаний рассматривалась нами как задача удовлетворения ограничений. Однако задачи составления расписаний часто являются проблемами оптимизации. Задача удовлетворения ограничений легко может быть расширена до задачи оптимизации ограничений (COP). Все, что требуется, – это функция, которая отображает решение с помощью вещественнозначной функции стоимости или функции оценки [35]. Задачи оптимизации ограничений не изучались так широко, как обычные CSP. Тем не менее, имеется ряд полезных методов для решения COP. Многие из этих методов являются адаптацией методов удовлетворения ограничений, а другие используют более ранние методы оптимизации, такие как имитация отжига.

Сведение задачи составления расписаний к задаче CSP позволяет использовать широкий арсенал средств CSP: алгоритмы распространения ограничений, поиск в глубину с возвратами, структурные алгоритмы, алгоритмы локального поиска, алгоритмы обработки глобальных ограничений и т.п. [36, 37].

С учетом того, что задачи составления расписаний требуют анализа временных и ресурсных ограничений, были разработаны специальные методы

для успешного решения подобных задач, такие как: стратегия распространения – edge-finding [38], стратегия поиска в глубину с возвратами – Limited discrepancy search (LDS) [39], стратегия упорядочивания переменных – Slack-based heuristics [40] и др.

Заключение

Подводя итог, можно утверждать, что многие реальные задачи планирования характеризуются тем, что они выходят за тесные рамки классического планирования, где предполагается наличие полной и правильной информации, а также детерминированной, полностью наблюдаемой среды. Это предположение является недействительным во многих проблемных областях.

Несмотря на то, что предпринимаются попытки преодолеть такие ограничения классического планирования, как сложность работы с вещественнозначными величинами, трудности при моделировании неопределенности, требования к полной наблюдаемости среды планирования и детерминированности действий исполнителя, пока нельзя с уверенностью сказать, что с большинством проблем удалось справиться, и все они достаточно хорошо проработаны [41].

Из проведенного обзора следует, что программирование в ограничениях как технология решения задач комбинаторного поиска и комбинаторной оптимизации находит широкое применение в задачах составления расписаний, однако в более общих задачах интеллектуального планирования ее применение все еще весьма ограничено.

По мнению авторов, данная технология в области интеллектуального планирования не получила должного внимания, а многие методы планирования используют методы SAT, а не CSP. При разложении задачи CSP и представлении ее в форме задачи SAT теряется часть информации о структуре задачи, которую можно использовать для ускорения процедур вывода.

Существенным недостатком современных методов интеллектуального планирования является невозможность специфицировать и решать задачи относительно переменных, являющихся сложно структурированными, семантически согласованными данными, в частности, объектно-ориентированными.

При разработке методов интеллектуального планирования возникает дополнительная сложность в случае, когда требуется обеспечить поддержку открытой модели предметной области. Процесс создания такой модели осуществляется поэтапно. На различных этапах в нее могут включаться типы зависимостей, критерии оптимизации, которые не могли быть априорно учтены. Если в некоторой постановке задачи планирования решения не существует, то должна быть обеспечена возможность «ослабить» задачу. В этих условиях применение традиционных методов теории исследования операций представляется проблематичным.

Литература

1. Russel S., Norvig P. Artificial Intelligence: A Modern Approach. 3rd edition. Prentice Hall, 2010. 1132 p.
2. Fikes R. and Nilsson N. STRIPS: A new approach to the application of theorem proving to problem solving. Artificial Intelligence, 1971, pp. 189-208.

3. Pednault E. ADL: Exploring the middle ground between STRIPS and the situation calculus. Proc. 1st Int. Conf. on Principles of Knowledge Representation and Reasoning, 1989. pp. 324-332.
4. Penberthy J. and Weld D. Temporal planning with continuous change. Proc. 12th National Conf. on AI, 1994. pp.1010-1015.
5. Peot M. and Smith D. Conditional nonlinear planning, Proc. 1st Int. Conf. on AI Planning Systems, 1992. pp. 189-197.
6. Williamson M and Hanks S. Optimal planning with a goal-directed utility model. Proc. 2nd Int. Conf. on AI Planning Systems, 1994. pp.176-180.
7. Blum A and Furst M. Fast planning through planning graph analysis. Proc. 14th Int. Joint Conf. on AI, 1995. pp. 1636-1642.
8. Anderson C, Smith D. and Weld D. Conditional effects in Graphplan. Proc. 4th Int. Conf. on AI Planning Systems. 1998. pp. 44-53.
9. Blum A and Langeford J. Probabilistic planning in the Graphplan framework. Proc. AIPS-98 Workshop on Planning as Combinatorial Search, 1998. pp. 8-12.
10. Smith D and Weld D. Conformant Graphplan. Proc. 15th National Conf. on AI, 1998. pp. 889-896.
11. Smith D and Weld D. Temporal planning with mutual exclusion reasoning. Proc. 16th Int. Joint Conf. On AI, 1999. pp. 326-333.
12. Koehler J. Planning under resource constraints. Proc. 13th European Conf. on AI. 1998. pp. 489-493.
13. Weld D. Recent advances in AI planning. AI Magazine. 1999, 20(2), pp. 93-123.
14. Rintanen J. Planning and SAT. In A. Biere, H. van Maaren, M. Heule and Toby Walsh, Eds., Handbook of Satisfiability, IOS Press. 2009, pp. 483-504.
15. Wolfman S. and Weld D. Combining linear programming and satisfiability solving for resource planning. The Knowledge Engineering Review. 1999. 15(1).
16. Drabble B. and Tate A. The use of optimistic and pessimistic resource profiles to inform search in an activity based planner. Proc. 2nd Int. Conf. on AI Planning Systems, 1994. pp. 243-248.
17. Allen J. and Koomen J. Planning using a temporal world model. Proc. 8th Int. Joint Conf. on AI, 1983. pp. 741-747.
18. Wilkins D. Can AI planners solve practical problems? / Computational Intelligence 6(4), 1990.
19. Erol K., Hendler J., and Nau D. Semantics for Hierarchical Task Network Planning. Technical report CSTR-3239, Dept. of Computer Science, U. Maryland. 1994.
20. Dean T., Kaelbling L., Kirman J., and Nicholson A. Planning under time constraints in stochastic domains. AI, 1995. 76, pp. 35-74.
21. Drabble B. Excalibur: a program for planning and reasoning with processes. Artificial Intelligence 1993. 62, pp.1-40.
22. Kautz H and Walser J. State-space planning by integer optimization. The Knowledge Engineering Review. 1999. 15(1).
23. Muscettola N. HSTS: integrating planning and scheduling. Proc. 12th National Conf. on AI, 1994, pp. 169-212.
24. Ferguson G., Allen J. and Miller B. TRAINS-95: towards a mixed-initiative planning assistant. Proc. 3rd Int. Conf. on AI Planning Systems, 1996. pp. 70-77.
25. Dechter R., Meiri I. and Pearl J. Temporal constraint networks. Artificial Intelligence, 1991, 49, pp. 61-95.

26. Zweben M. and Fox M. Intelligent Scheduling. ACM SIGART Bulletin, 1995, 6(2), pp. 49-51.
27. Pinedo M. Scheduling Theory, Algorithms and Systems. Prentice Hall. NY, 1995, 671 p.
28. Baker K. Elements of Sequencing and Scheduling. Dartmouth College, Hanover, NH, 1998.
29. Лазарев, А.А. Теория расписаний. Задачи и алгоритмы. / А.А.Лазарев, Е.Р. Гафаров; – М.: Московский государственный университет им. М.В. Ломоносова (МГУ), 2011, 222 с.
30. Colomi, M. Dorigo et V. Maniezzo. Distributed Optimization by Ant Colonies, actes de la premiere conference europeenne sur la vie artificielle, Paris, France, Elsevier Publishing. 1991. pp. 134–142.
31. M. Dorigo. Optimization, Learning and Natural Algorithms, PhDthesis, Politecnico di Milano, Italie, 1992.
32. Brucker P. Scheduling Algorithms, second edition. Springer-Verlag. 1998.
33. Kautz H. and Selman B. Pushing the envelope: planning, propositional logic, and stochastic search. Proc. 13th Nat. Conf. on AI. 1996, pp. 1194-1201.
34. Jónsson A., Morris P., Muscettola N., Rajan K. Next generation remote agent planner. Proc. 5th Int. Symposium on AI, Robotics and Automation in Space. ESTEC, Noordwijk, Netherlands, 1999, pp. 363-370.
35. Plaunt C., Jónsson A., and Frank J. Satellite tele-communications scheduling as dynamic constraint satisfaction. Proc. 5th Int. Symposium on AI, Robotics and Automation in Space, 1999, pp. 277-284.
36. Chu G., Stuckey P. Learning value heuristics for constraint programming. In Integration of AI and OR techniques in constraint programming. Springer LNCS, 2015, 9075. pp. 108-123.
37. Stefan Kreter, Andreas Schutt, Peter J. Stuckey Using constraint programming for solving RCPSP/max-cal. Constraints, 2017, Volume 22, Issue 3, pp. 432-462.
38. Carlier J. and Pinson E. An algorithm for solving the job shop scheduling problem. Management Science, 1989. 35(2). pp. 164-176.
39. Harvey W, and Ginsberg M. Limited discrepancy search. Proc. 14th Int. Joint Conf. on AI, 1993. pp. 607–613.
40. Smith S. and Cheng C. Slack-based heuristics for constraint satisfaction scheduling. Proc. 11th Nat. Conf. on AI, 1993. pp. 139–144.
41. David E. Smith, Jeremy Frank, and Ari K. Jónsson. Bridging the Gap Between Planning and Scheduling. Knowledge Engineering Review, 15(1), 2000. pp. 47-83.

Сведения об авторах

Зуенко Александр Анатольевич – к.т.н, старший научный сотрудник

e-mail: zuenko@iimm.ru

Alexander A. Zouenko – Ph.D. (Tech. Sci.), senior researcher

Фридман Ольга Владимировна – к.т.н, старший научный сотрудник

e-mail: ofridman @iimm.ru

Olga V. Fridman – Ph.D. (Tech. Sci.), senior researcher

УДК 004.94

DOI: 10.25702/KSC.2307-5252.2018.10.36-47

С.Н. Малыгина^{1,2}, В.В. Быстров¹, Д.Н. Халиуллина¹

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² Филиал ФГБОУ ВО «МАГУ» в г. Апатиты

ЛОГИСТИКА КАДРОВОГО ОБЕСПЕЧЕНИЯ РЕГИОНА: ФОРМАЛИЗАЦИЯ И СТРУКТУРА ПОЛИМОДЕЛЬНОГО КОМПЛЕКСА*

Аннотация

Статья посвящена теоретическим аспектам разработки информационно-аналитических средств поддержки управления региональными социально-экономическими системами. Авторы предлагают формализованную концептуальную модель кадрового обеспечения, построенную на основе логистического подхода к управлению сложными системами. Формальное описание задач управления кадровой логистикой региона позволяет структурировать знания о предметной области и использовать их для автоматизации процессов разработки полимодельного комплекса. В работе предлагается структура полимодельного комплекса логистики кадрового обеспечения в рамках региона.

Ключевые слова:

Логистика, кадровое обеспечение, полимодельный комплекс, концептуальное моделирование, имитационное моделирование

S.N. Malygina, V.V. Bystrov, D.N. Khaliullina

LOGISTICS OF PERSONNEL PROVISION IN THE REGION: FORMALIZATION AND STRUCTURE OF THE POLY-MODEL COMPLEX

Abstract

The article is devoted to the theoretical aspects of the development of information and analytical tools to support the management of regional socio-economic systems. The authors propose a formalized conceptual model of personnel provision, based on the logistics approach to the management of complex systems. A formal description of the tasks of personnel logistics management in the region allows to structure knowledge about the subject area and use them to automate the development of the poly-model complex. The paper proposes the structure of the poly-model complex of logistics staffing within the region.

Keywords:

logistics, personnel provision, poly-model complex, conceptual modeling, simulation

Введение

В любом регионе обеспечение стабильной социально-экономической ситуации и ее благоприятное развитие зависят от множества разнообразных факторов. Одним из таких факторов является формирование трудового потенциала, под которым обычно понимается наличие необходимых трудовых ресурсов для покрытия всех кадровых потребностей региональной экономики. Одной из проблем формирования трудового потенциала региона является

* Работа выполнена в рамках темы госзадания "Модели и методы конфигурирования адаптивных многоуровневых сетевых систем управления региональной безопасностью в Арктической зоне Российской Федерации" (№ 0226-iiimm-2017-14-08).

высокая подвижность кадров. Она не позволяет получать количественные и качественные оценки характеризующихся высокой степенью достоверности явлений, таких как: наличие или отсутствие свободных рабочих мест, количество доступных трудовых ресурсов, возможность их трудоустройства, а также полное или частичное покрытие вакансий. В области регулирования динамичных процессов, связанных с обеспечением необходимыми ресурсами социально-экономических систем, зарекомендовали себя на практике методы логистического управления. Таким образом, в рамках проводимой научно-исследовательской работы предлагается применить логистический подход для кадровых потоков региональных социально-экономических систем.

Анализируя открытые источники информации, можно сделать вывод, что в большинстве случаев понятие «кадровая логистика» используется в рамках управления отдельным предприятием, где применяется термин «микрологистика». В то же время встречаются публикации, в которых данное понятие расширяется до уровня экономической отрасли, и вводится термин «макрологистика». Само понятие «логистика» включает в себя планирование, выполнение и контроль движения и размещения людей и/или товаров, а также поддерживающие действия, связанные с таким движением и размещением, в пределах экономической системы, созданной для достижения своих специфических целей [1]. В свою очередь, кадровая логистика – это раздел логистики, в котором изучается оптимизация потоков трудовых ресурсов предприятий и отрасли в целом. При этом основной задачей кадровой логистики является выработка управляющих воздействий, направленных на обеспечение оптимального баланса между входными и выходными кадровыми потоками социально-экономической системы.

Если рассматривать кадровую логистику в рамках предприятия, принято выделять четыре основных ее направления: процессы организации входных кадровых потоков, процессы использования кадров, процессы развития кадров и высвобождение кадров [2]. При таком рассмотрении входные кадровые потоки представляют собой процессы, связанные с организацией приема сотрудников на работу. В зависимости от сферы деятельности и требований организации, данный поток может быть сформирован за счет выпускников различных учебных заведений (школ, вузов, ссузов), трудовых ресурсов, стоящих на бирже труда (или иных служб занятости), а также потенциальных работников, которые уже занимают рабочее место в другой организации.

Наличие выходных кадровых потоков обусловлено процессами высвобождения рабочей силы. В качестве основных причин появления таких процессов выступают следующие факторы: увольнение по различным причинам (собственному желанию, сокращению рабочих мест, нарушение работником трудовой дисциплины), выход на пенсию, смерть.

Очень важным вопросом при организации кадровой логистики является проблема развития кадров. В первую очередь она связана с организацией условий и управлением профессиональным ростом работника. Одним из направлений, в данном случае, является обучение и переобучение трудовых ресурсов, которое может происходить в специализированных учреждениях – организациях дополнительного образования, – либо напрямую на предприятии в виде так называемого профессионального обучения. Профессиональный рост работников в большинстве случаев приводит к изменению кадровых потоков организации.

При переходе от рассмотрения отдельного предприятия к изучению отрасли (переход к макрологистике), на первый план выходит вопрос оптимизации потоков кадров на стадии их обучения. При этом особое внимание уделяется подготовке квалифицированных кадров, взаимосвязи учебных заведений с отраслью, притоку и адаптации специалистов в отрасли, отраслевой системе переподготовки и повышения квалификации кадров и т.д. [3].

В кадровой логистике движение кадровых потоков может быть как вертикальным, так и горизонтальным (рис. 1).

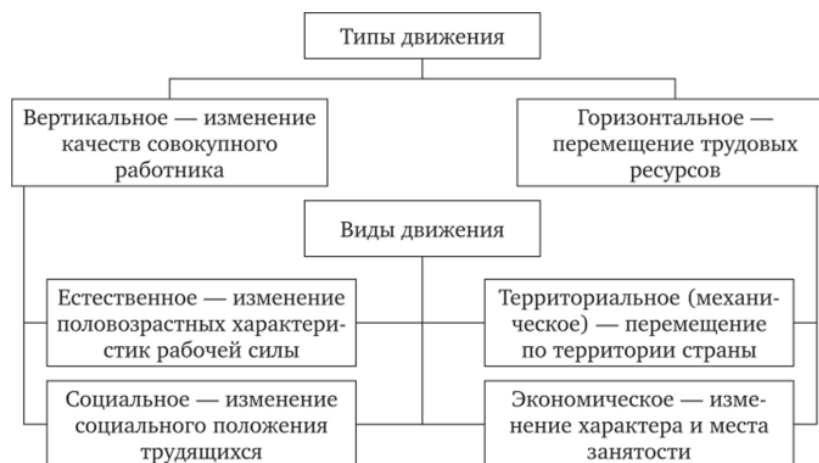


Рис. 1. Классификация движения трудовых ресурсов [4]

При «вертикальном движении» происходит изменение (развитие) качеств каждого отдельного индивидуума (изменение с течением времени возраста, расширение и обновление общих и специальных знаний, трудовых навыков, накопление производственного опыта и т.д.), что приводит к соответствующим сдвигам в составе трудовых ресурсов отрасли в целом. «Горизонтальное движение» связано с перемещением трудовых ресурсов по территории региона, из одних сфер деятельности в другие.

Можно провести деление видов движения трудовых ресурсов и по другим критериям. Например, условная классификация видов движения кадров относительно среды движения:

1. Естественное движение, которое связано с изменением половозрастной структуры трудовых ресурсов, выбытием людей из состава трудовых ресурсов вследствие естественного старения их организма, инвалидности или смерти.

2. Социальное движение обусловлено изменением социального положения членов общества, их принадлежности к тому или иному классу, общественному слою или социальной группе.

3. Территориальное движение населения и трудовых ресурсов возникает в результате перемещения людей между странами, регионами, городской и сельской местностью, различными населенными пунктами.

4. Экономическое движение рабочей силы объединяет различные формы перемещения работников в совокупном трудовом процессе. К нему относятся

перемещения, связанные с приемом на работу и увольнением с нее, смена места работы внутри организации или за ее пределами в связи с изменением профессии, ростом квалификации, поиском лучших условий реализации трудового потенциала работника и по другим причинам [4].

По мнению И.С. Масловой трудовая миграция населения является объективным процессом перемещения кадров, который неразрывно связан с развитием производительных сил и производственных отношений [6]. Одной из разновидностей трудовой миграции является текучесть кадров (текучесть рабочей силы) – неорганизованные перемещения работников между предприятиями (организациями) в рамках одного города или региона. К основным причинам ухода персонала с прежнего места работы можно отнести следующие: недовольство оплатой труда, графиком работы, условиями труда, отсутствие карьерного роста, социальная напряженность и т.п. [5]. Экономическая функция миграции связана с обеспечением определенного уровня мобильности рабочей силы, ее территориальным перераспределением и обеспечением количественного и качественного соответствия между спросом и предложением рабочей силы различного направления подготовки и квалификации в разных районах страны и населенных пунктах [4].

Обобщая выше написанное, можно сделать вывод, что независимо от точки зрения на кадровую логистику региона процессы, связанные с организацией и управлением кадровыми потоками, так или иначе оперируют большими объемами разнородной информации. Такая информация содержит данные о многих ключевых факторах и их количественных и качественных показателях, формирующих компоненты регионального рынка труда (трудовая миграция, система подготовки и переподготовки кадров, государственная и негосударственные службы занятости, и др.). Для адекватной оценки текущей ситуации в сфере кадрового обеспечения региональной экономики и рационального управления кадровой политикой можно применять информационно-аналитические средства, автоматизирующие процессы рутинной обработки данных и выработки рекомендаций для лиц, принимающих решения. В основу подобного класса систем может быть положено прогнозное моделирование, позволяющие с помощью вычислительного эксперимента строить и исследовать различные сценарии развития ситуации в логистике кадрового обеспечения региональных социально-экономических систем.

В рамках данной статьи приводится описание некоторых результатов научно-практических исследований, направленных на создание новых методов, технологий и средств информационной поддержки управления логистикой кадрового обеспечения региональной экономики.

Формализованная концептуальная модель кадровой логистики

В работе рассматривается кадровая логистика региональных социально-экономических систем, которые в рамках предложенного формального описания были названы экономическим сектором. В зависимости от парадигмы организации региональной экономики секторам могут соответствовать отдельные отрасли экономики (отраслевой подход), либо кластеры экономики (кластерный подход). Кадровую логистику региона представим в терминах концептуального моделирования как систему управления движением кадров и формально опишем следующим образом:

$$PL = \{ES, LM, Ed, SC\},$$

где $ES = \{S\}$ – множество секторов региональной экономики,

LM – рынок труда, который представляет собой множество нетрудоустроенных людей (потенциальные трудовые ресурсы), проживающих в регионе,

Ed – система подготовки и переподготовки кадров региона,

$SC = \{CO, AM, CM, M, R^{Res}\}$ – система управления движением кадров.

Здесь CO – множество объектов управления, включающих в себя представителей каждой составляющей кадровой логистики. Так, например, в рамках экономического сектора можно рассматривать в качестве объекта управления штатное расписание, в системе подготовки кадров – перечень направлений подготовки и количество мест, на рынке труда – количество вакансий.

AM – множество субъектов управления, в качестве которых могут выступать региональные и муниципальные власти, администрация предприятий секторов экономики, руководство образовательных учреждений,

CM – множество средств управления. Средствами управления, например, могут выступать законодательные акты, регулируемые экономические и социальные показатели/параметры (заработная плата, размер пособия по безработице, стипендий, состав социального пакета на предприятиях и т.д.)

M – множество мероприятий, под которыми подразумеваются проекты, направленные на изменение региональной кадровой логистики. Например, реализация крупного инвестиционного проекта в масштабах всего региона (так из стратегии социально-экономического развития Мурманской области – создание и развитие мурманского транспортно-логистического центра).

R^{Res} – отношение «ответственности», которое позволяет задать зависимость между субъектом управления и управляемым им объектами управления.

В предлагаемой концептуальной модели кадровой логистики используется понятие сектор экономики. В общем случае данное понятие представляет собой совокупность взаимосвязанных хозяйствующих субъектов, объединенных по однотипному виду экономической деятельности, который они осуществляют в регионе. В практическом смысле сектор экономики может интерпретироваться как отрасль экономики или как экономический кластер в зависимости от выбранного способа представления региональной экономики.

Каждый сектор региональной экономики S состоит из обобщенного штатного расписания всех предприятий отрасли St (Staff), трудоустроенных работников сектора экономики W , нетрудоустроенных людей в данном секторе экономики Um , прибывших In и убывших Out работников в/из сектор(-а) экономики. Движение трудовых ресурсов может образовывать как внутренние кадровые потоки (внутренняя миграция), так и внешние (внешняя миграция). Таким образом, кадровую логистику экономического сектора с формальной точки зрения можно описать в виде следующего кортежа:

$$S = \langle St, W, Um, In, Out, R^{ls}, R^{In}, R^{Out} \rangle,$$

где $St = \{st\}$ – обобщенное штатное расписание сектора, представляющее собой множество всех должностей предприятий сектора. Каждый элемент множества имеет следующие атрибуты:

$$st = \langle name, level, salary, qualReq, organization, flag \rangle,$$

где *name* – название должности,

level – категория должности, позволяющая производить разделение множества должностей по определенному признаку. Например, разделение по функциональному принципу: управленцы, рабочие, служащие и др., или по разрядной сетке, или по квалификационному уровню профессиональных стандартов.

salary – заработная плата,

qual_req – квалификационные требования (уровень образования, направление образования, стаж, соответствие профстандарту).

organization – организация, к штатному расписанию которой относится должность,

$$flag = \begin{cases} 1, & \text{если должность занята;} \\ 0, & \text{если должность свободна.} \end{cases}$$

$W = \{w\}$ – множество работников сектора, каждый элемент имеет следующие атрибуты:

$w = \langle persData, qualLevel, organization \rangle$,

где *persData* – персональные данные (пол, возраст),

qualLevel – квалификационный уровень (уровень образования, специализация, стаж),

organization – организация, в которой работает работник.

Отметим, что в предлагаемой концептуальной модели под специализацией понимается профиль или направление подготовки образовательных программ среднего профессионального и высшего образования (например, в рамках перечня укрупненных групп специальностей и направлений подготовки Министерства образования РФ).

$Um = \{um\}, Um \subset LM$ – множество нетрудоустроенных сектора. Каждый элемент имеет следующие атрибуты:

$um = \langle persData, qualLevel \rangle$,

где *persData* – персональные данные,

qualLevel – квалификационный уровень (уровень образования, специализация, стаж).

$In = \{Imm, Ar, Gr\}$ – множество прибывших в сектор, состоящее из трех непересекающихся подмножеств:

Imm – множество прибывших из других регионов,

Ar – множество перешедших из другого сектора,

Gr – множество выпускников по соответствующей специализации.

Элементы всех этих множеств имеют следующие атрибуты:

$\langle persData, qualLevel \rangle$,

где *persData* – персональные данные,

qualLevel – квалификационный уровень (уровень образования, специализация, стаж).

$Out = \{Ret, Emm, Inc, Dead\}$ – множество выбывших из сектора, состоит из следующих непересекающихся подмножеств:

Ret – множество мигрировавших в другой кластер,

Emm – множество мигрировавших (уехавших) из региона,

Inc – множество перешедших в категорию нетрудоспособных (в результате выхода на пенсию и по инвалидности),

Dead – множество умерших.

Элементы всех этих множеств имеют следующие атрибуты:

$\langle persData, qualLevel, organization \rangle$,

где *persData* – персональные данные,

qualLevel – квалификационный уровень (уровень образования, специализация, стаж),

organization – организация, к которой работник принадлежит.

На вышеперечисленных множествах заданы следующие отношения.

$R^{ls}: W \rightarrow St$ – вспомогательное отношение для установления соответствия трудоустроенного работника сектора с должностью в штатном расписании сектора («отношение трудоустройства»):

$R^{ls} = \{(w, st): \forall w \in W \exists st \in St, qualLevel_w = qualReq_{st}\}$,

где $qualLevel_w$ – квалификационный уровень работника *w*, $qualReq_{st}$ – квалификационные требования к должности *st*.

Движение кадров внутри экономического сектора можно разделить на горизонтальную миграцию, под которой подразумевается движение кадров между предприятиями одного сектора в рамках региона, и вертикальную миграцию, под которой подразумевается рост кадров по карьерной лестнице и понижение в должности. Таким образом, внутреннюю миграцию можно представить набором следующих отношений:

$R^{In} = \{R^{Hor}, R^{Ver}\}$,

где R^{Hor} – горизонтальной миграции, R^{Ver} – вертикальная миграция.

$R^{Hor}: W_i \rightarrow St_j, i \neq j, W_i \subset W, St_j \subset St, i, j$ – номер организации.

Отношение R^{Hor} показывает переход работника *i*-ой организации на должность в *j*-ой организации одного сектора экономики в пределах региона.

Вертикальную миграцию можно представить набором следующих отношений:

$R^{Ver} = \{R^{Up}, R^{Down}\}$,

где $R^{Up}: W_i \rightarrow St_i$.

Отношение R^{Up} показывает рост по карьерной лестнице. Это отношение можно задать правилом:

Если $w \in W_i, qualLevel_w(t) > qualLevel_w(t-1) \exists st \in St_i : flag_{st} = 0, level_{st} > level_{st(w)}, qualLevel_w = qualReq_{st}$,

То $flag_{st} = 1, flag_{st(w)} = 0, R^{ls}(w)$.

Здесь $qualLevel_w(t)$ – текущий уровень квалификации работника, $qualLevel_w(t-1)$ – предыдущий уровень квалификации работника, $level_{st}$ – категория свободной должности, $level_{st(w)}$ категория должности, занимаемой работником.

В случае несоответствия работника занимаемой должности или сокращения штата, применяется отношение понижения в должности:

$R^{Down}: W_i \rightarrow St_i$.

Это отношение можно описать следующим образом:

Если $w \in W_i, (qualLevel_w(t) \neq qualReq_{st(w)}) \vee (St_i = St_i \setminus st(w)) \exists st \in St_i : flag_{st} = 0, level_{st} \leq level_{st(w)}, qualLevel_w = qualReq_{st}$

то $flag_{st} = 1, flag_{st(w)} = 0, R^{ls}(w)$.

Внешняя миграция представляет собой входные кадровые потоки и высвобождение кадров. Формально ее можно задать набором из двух типов отношений:

$$R^{Out} = \{R^E, R^I\}.$$

где R^E – отношение высвобождения кадров, R^I – отношение входного потока кадров. Каждое из них в свою очередь состоит из множества отношений.

Высвобождение кадров может произойти по различным причинам: переход на предприятия другого экономического сектора, переезд в другой регион, выход на пенсию, увольнение по различным другим причинам. Исходя из этого, отношение высвобождения кадров можно задать в виде следующего множества:

$$R^E = \{R^{Ret}, R^{Emm}, R^{Inc}, R^D, R^{Um}\},$$

где R^{Ret} – переход в другой сектор экономики региона, R^{Emm} – переезд в другой регион, R^{Inc} – переход в категорию нетрудоспособных, R^D – переход в категорию умерших, R^{Um} – переход в категорию нетрудоустроенных.

Переход в другой сектор экономики задается отображением вида

$$R^{Ret}: W \rightarrow Ret$$

Переезд в другой регион задается отображением следующего вида

$$R^{Emm}: W \rightarrow Emm$$

Переход в категорию нетрудоспособных может произойти в связи выходом работников на пенсию или получения инвалидности, и задается отношением вида

$$R^{Inc}: W \rightarrow Inc$$

Переход в категорию умерших задается отображением множества работников на множество умерших

$$R^D: W \rightarrow Dead$$

Переход в категорию нетрудоустроенных может произойти в случае сокращения штата, увольнения работника из-за невыполнения им трудовых функций, и задается отношением вида

$$R^{Um}: W \rightarrow Um.$$

Входной поток кадров формируется за счет приехавших из другого региона, перешедших из другого сектора региональной экономики, за счет новых кадров (выпускников вузов и ссузов) и трудоустройства неработающих на свободные должности в штатном расписании. Следовательно, отношение входного потока можно задать набором из четырех видов отношений:

$R^I = \{R^{Imm}, R^{Ar}, R^{Gr}, R^{Um}\}$, где R^{Imm} – отношение приезда из другого региона, R^{Ar} – отношение перехода из другого экономического сектора, R^{Gr} – отношение появления молодых кадров, R^{Um} – отношение трудоустройства нетрудоустроенных на свободные рабочие места в штатном расписании.

Приезд из другого региона может быть вызван приглашением специалистов на определенную должность, или самостоятельным приездом в поисках лучших условий. В первом случае приехавшие сразу переходят в категорию работников отрасли, во втором случае могут перейти либо в категорию работников при наличии вакансии нужного классификационного уровня, либо в категорию нетрудоустроенных, если такой вакансии нет. Данное отношение задается отображением вида:

$$R^{Imm}: Imm \rightarrow (W \times St) \cup Um.$$

Переход из другой отрасли происходит в случае наличия свободной ставки нужного квалификационного уровня и задается отношением вида:

$$R^{Ar}: Ar \rightarrow W \times St.$$

Выпускники вузов и ссузов могут перейти в категорию работников при наличии свободной ставки или в категорию нетрудоустроенных при ее отсутствии:

$$R^{Gr}: Gr \rightarrow (W \times St) \cup Um.$$

При наличии свободных ставок и нетрудоустроенных людей требуемой квалификации применяется отношение трудоустройства на свободные ставки штатного расписания экономического сектора:

$$R^{Um}: Um \rightarrow W \times St.$$

Построение формализованной концептуальной модели позволяет систематизировать знания и данные о предметной области (в нашем случае, о региональной кадровой логистике). Такая систематизация понятий и терминов дает возможность оперировать ими в автоматизированном режиме с помощью вычислительных средств. Для этого необходимо отобразить формализованную концептуальную модель в одном из форматов представления знаний. Для решения этой задачи предлагается использовать зарекомендовавший себя на практике способ, а именно, построение онтологического описания.

В рамках работы концептуальная модель кадровой логистики представлена в виде прикладной онтологии формата OWL. Данная прикладная онтология используется для автоматизированной генерации набора имитационных моделей. Имитационные модели позволяют осуществлять прогнозирование изменения ситуации в региональной кадровой логистике в различных сценарных условиях.

Структура полимодельного комплекса кадровой логистики

На основе предложенных формальной концептуальной модели и прикладной онтологии кадровой логистики разработан полимодельный комплекс. Комплекс представляет собой совокупность взаимосвязанных компьютерных моделей, каждая из которых описывает одну из составляющих кадровой логистики. Для реализации моделей предлагается использовать метод системной динамики и агентное моделирование.

Общая структура полимодельного комплекса кадровой логистики представлена на рисунке 2.

В соответствии с предложенной формальной моделью кадровой логистики региона полимодельный комплекс состоит из четырех блоков: ES – сектора экономики, LM – рынок труда, Ed – система подготовки и переподготовки кадров, SC – система управления движением кадров. Между блоками существует два вида взаимодействия:

- Движение трудовых ресурсов между блоками «сектора экономики», «рынок труда», «система подготовки и переподготовки кадров». Данный вид взаимодействия на рисунке показан в виде прозрачных стрелок.
- От «системы управления» на остальные блоки поступают управляющие воздействия, которые вырабатываются на основе информации о текущей ситуации в блоках ES, LM и Ed. Эти взаимодействия изображены в виде пунктирных стрелок.

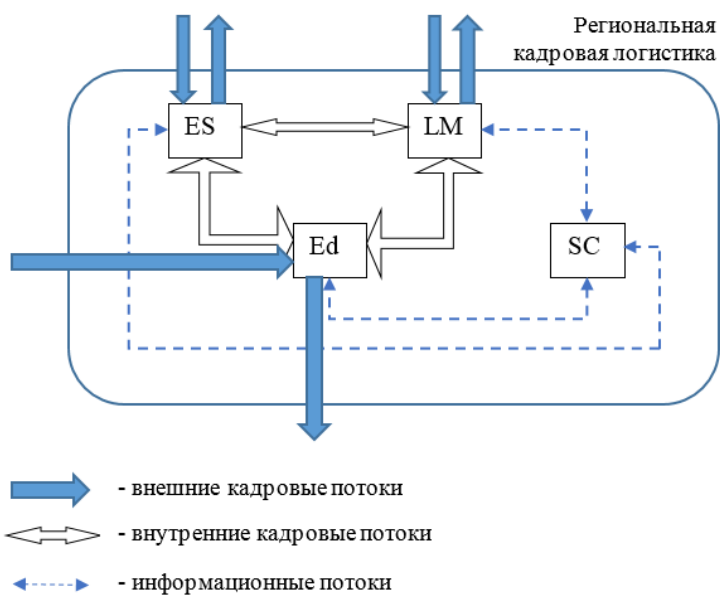


Рис. 2. Общая структура полимодельного комплекса региональной кадровой логистики

На блоки «сектора экономики», «рынок труда», «система подготовки и переподготовки кадров» влияют внешние воздействия, которые отражают движение кадров в/из региона (толстые стрелки).

Каждый из блоков представляет собой набор имитационных моделей. Так, например, блок ES представляет собой совокупность моделей, имитирующих работу региональных экономических секторов (рис.3). Движение трудовых ресурсов происходит как между секторами, так и между блоками, что на схеме отображено двойными (двунаправленными) стрелками.

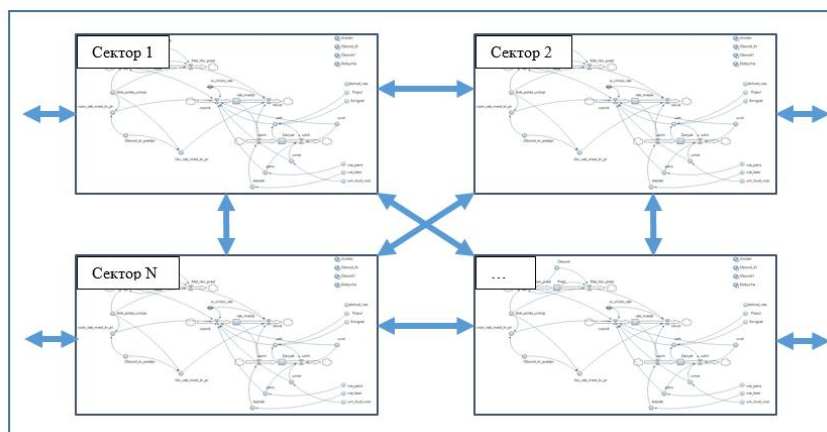


Рис. 3. Структура блока «Сектора экономики»

Блок «Система подготовки и переподготовки кадров» представляет собой набор моделей, имитирующих процесс подготовки кадров вузами и ссузами

региона по различным направлениям образовательных программ, а также системы дополнительного и профобразования. Выходными данными этого блока являются количество выпускников по каждому направлению, которые являются входными данными для блоков «трудовые ресурсы» и «сектора экономики».

В статье предлагается структура полимодельного комплекса логистики кадрового обеспечения региональных социально-экономических систем. Комплекс будет разрабатываться с применением методов имитационного моделирования, таких как метод системной динамики, агентное моделирование. А также осуществляться генерация структуры имитационных моделей с помощью авторских методик автоматизированного синтеза компьютерных моделей на основе когнитивных карт, концептуальных и онтологических описаний, и ранее разработанной библиотекой модельных шаблонов поддержки управления региональной безопасностью [7].

Заключение

Разработка методов и средств информационной поддержки управления сложными социально-экономическими системами является одной из приоритетных задач развития современного информационного общества. К таким сложным системам можно отнести и кадровую логистику региона, включающую в себя потоки трудовых ресурсов, вызванные разными процессами внутри регионального рынка труда. К этим процессам относятся как территориальное движение кадров, обусловленное перемещением работников между предприятиями в рамках хозяйствующего субъекта, региона, страны, так и внутреннее движение, связанное с изменением структуры кадрового обеспечения, перечня квалификационных требований профессий и профессиональной компетентности самих трудовых ресурсов.

Для систематизации информации о кадровой логистике региональных социально-экономических систем можно использовать концептуальное моделирование и методы системного анализа. Это позволяет построить формализованное концептуальное описание системы управления логистикой кадрового обеспечения региона, выделив в нем основные объекты, их атрибуты и отношения между ними. Формализация позволяет структурировать представления об исследуемой системе и процессах, протекающих в ней, что, в свою очередь, может быть взято за основу при разработке компонентов информационных технологий управления данной системой. В частности, в данной работе построенная концептуальная модель послужила фундаментом для разработки полимодельного комплекса логистики кадрового обеспечения региональных социально-экономических систем.

На текущем этапе исследования при построении формализованной концептуальной модели кадровой логистики сделано одно допущение. Оно связано с тем, что в модели пока не рассматривается возможность совмещения одним работником нескольких должностей или трудовая деятельность по совместительству в другой организации. Но данное допущение сильно не влияет на адекватность предложенного формального описания, т.к. внешнее совместительство свойственно для небольшого ограниченного множества секторов экономики, а внутреннее совместительство можно реализовать на этапе представления соответствующих концептов в структуры баз данных.

Литература

1. ELA Certification for Logistics Professionals. Standards 040805. – Brussels: European Certification Body for Logistics (Перевод д.э.н., проф. В.И. Сергеева), 2004. - 15р.
2. Есенькин, Б.С. Крылова, М.Д. Логистика в книжном деле: Учебник / Б.С. Есенькин, М.Д. Крылова – Моск. гос. ун-т печати М.: Изд-во МГУП, 2002. - Гл. 12 - 335 с.
3. Кормин, Н.Г. Особенности кадровой логистики современной организации / Н.Г. Кормин // Бизнес, менеджмент и право: «Национальные проекты в Российской Федерации» – 2007 – №1(13) – С. 13-15.
4. Экономика труда : учебник для академического бакалавриата / М. В. Симонова [и др.] ; под общ. ред. М. В. Симоновой. — М. : Издательство Юрайт, 2018. – 259 с. – (Серия : Бакалавр. Академический курс). – ISBN 978-5-534-05423-1.
5. Пластинина, Н. Текучесть кадров: причины и следствие, пути выхода из ситуации/ Н. Пластинина //Кадровик – 2011 – №4 – С. 44-51.
6. Маслова, И.С. Экономические функции и факторы миграции населения. / И.С. Маслова – М., 2003.
7. Быстров, В.В., Маслобоев, А.В., Малыгина, С.Н. Библиотека имитационных моделей поддержки управления региональной безопасностью // Свидетельство об официальной регистрации Базы данных ФГАНУ «ЦИТиС». – гос. рег. № АААА-Г18-618041090032-0 от 10.04.2018 г. (Объединенный фонд ФГАНУ «ЦИТиС»).

Сведения об авторах

Малыгина Светлана Николаевна – к.т.н., научный сотрудник

e-mail: malygina@iimm.ru

Svetlana N. Malygina – PhD, researcher

Быстров Виталий Викторович – к.т.н., старший научный сотрудник

e-mail: bystrov@iimm.ru

Vitaliy V. Bystrov – PhD, senior researcher

Халиуллина Дарья Николаевна – к.т.н., научный сотрудник

e-mail: khaliullina@iimm.ru

Darya N. Khaliullina – PhD, researcher

УДК 004.94

DOI: 10.25702/KSC.2307-5252.2018.10.48-60

В.В. Быстров, Д.Н. Халиуллина

Институт информатики и математического моделирования ФИЦ КНЦ РАН

ФОРМАЛИЗОВАННАЯ КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ ОБЕСПЕЧЕНИЯ ЭКОЛОГИЧЕСКОЙ БЕЗОПАСНОСТИ АРКТИЧЕСКИХ КОММУНИКАЦИЙ*

Аннотация

В статье рассматриваются вопросы создания методов и средств информационной поддержки сетецентрического управления экологической безопасностью арктических коммуникаций. Внимание уделяется постановке задачи исследования и представлению формализованного интегрированного описания предметной области и виртуальной среды управления. Предлагается концептуальная модель обеспечения экологической безопасности, сформированная в терминах теоретико-множественных операций. Рассматриваются особенности представления концептуальной модели в виде прикладной онтологии и ее практического применения.

Ключевые слова:

экологическая безопасность, арктические коммуникации, концептуальное моделирование, прикладная онтология, имитационное моделирование

V.V. Bystrov, D.N. Khaliullina

FORMALIZED CONCEPTUAL MODEL OF ECOLOGICAL SAFETY OF THE ARCTIC COMMUNICATIONS

Abstract

The article deals with the creation of methods and means of information support for network-centric management of ecological safety of Arctic communications. Attention is paid to the formulation of the research problem and the presentation of a formalized integrated description of the subject area and the virtual management environment. The conceptual model of ecological safety, formed in terms of set-theoretic operations, is proposed. The features of the conceptual model representation in the form of applied ontology and its practical application are considered.

Keywords:

ecological safety, Arctic communications, conceptual modeling, applied ontology, simulation

Введение

В последние годы одним из стратегических направлений внешней политики РФ является развитие Арктических территорий, что обусловлено совокупностью разных причин: крупные запасы природных ресурсов (подземные полезные ископаемые, пресная вода, биоресурсы), важное геополитическое местоположение, большой экономический потенциал в использовании различного вида коммуникаций. В частности, развитие арктических коммуникаций является одной из приоритетных задач, зафиксированной в стратегии освоения Российской Арктики. Большинство представителей исполнительной власти и экспертов ассоциируют деятельность в рамках данного направления с созданием эффективной инфраструктуры для использования Северного морского пути (рис. 1).

* работа выполнена при финансовой поддержке РФФИ проект №№ 18-07-00167 «Методы и средства сетецентрического управления экологической безопасностью Северного морского пути».



Рис. 1. Северный морской путь в Арктической зоне[4].

Приведем некоторые данные об экономической привлекательности Арктических регионов и проектов, направленных на развитие их транспортной инфраструктуры. В настоящее время через северные порты западного региона России (Мурманск, Архангельск, Витино) и нефтяные терминалы Обской губы, Варандейского месторождения осуществляется перевалка около 40 млн. тонн грузов в год. Около половины этого объема (19 млн. тонн) составляют нефтяные грузы, из которых 16 млн. тонн экспортируется [2].

Развитие Северного морского пути для Российской Федерации позволит значительно сократить время доставки грузов, продвинуться на новые рынки сбыта, проводить военную подготовку в экстремальных условиях Арктики. Но суровый арктический климат (вероятность штормов и низкие температуры в зимний период, быстрое оледенение, продолжительный ледовый период) ограничивает возможности судоходства [4].

Помимо положительного эффекта развития Арктики, наблюдаются и отрицательные тенденции:

1. Интенсивное загрязнение окружающей среды (трансграничный перенос загрязняющих веществ водными и атмосферными потоками, нефтяное, химическое и радиоактивное загрязнения) и ухудшение качества поверхностных и подземных вод на прибрежных территориях арктической зоны РФ.

2. Изменение биоразнообразия и снижение запасов биоресурсов.

3. Ухудшение среды обитания коренного населения, проживающего в районах Арктической зоны, и нарушение условий традиционного природопользования коренных малочисленных народов.

4. Негативные последствия и угрозы глобальных изменений климата.

5. Деградация земель и нарушение условий землепользования.

Для решения перечисленных проблем (сохранение и защита природной среды, ликвидация экологических последствий хозяйственной деятельности и др.) на государственном уровне необходимо закрепить основные положения

обеспечения экологической безопасности. Кроме того в рамках решения приведенных выше проблем должны активно использовать информационные технологии для выявления потенциальных угроз и выработки решений по их предотвращению.

Разработка информационных средств и технологий для управления экологической безопасностью позволит не только оценить текущую ситуацию, но и спрогнозировать динамику показателей экологической и связанной с ней экономической безопасности арктических коммуникаций. При этом методология прогнозирования и исследования динамики различных параметров безопасности должна предусматривать системный анализ экономических и социальных аспектов, экологической ситуации и антропогенного загрязнения.

Одним из эффективных решений в данной области является применение компьютерного моделирования для исследования проблем обеспечения экологической и экономической безопасности. Использование имитационного моделирования, в свою очередь, позволяет изучать факторы, влияющие на выделенные аспекты безопасности.

Формализованная концептуальная модель

В рамках научно-исследовательской работы предложено формализованное концептуальное описание экологической безопасности арктических коммуникаций. Оно основано на использовании методов системного анализа и концептуального моделирования, успешно зарекомендовавших себя в исследовании сложных систем. Такая формализация позволяет создать фундамент для единого описания предметной области путем выделения основных концептуальных единиц и отношений между ними. Разработанное формализованное концептуальное описание является логическим развитием научных результатов работы «Исследование и разработка моделей и методов информационной поддержки управления региональной безопасностью (на примере Мурманской области)»[3], в рамках которой рассматривалась концептуальная модель мультиагентной информационно-аналитической среды региональной безопасности. Данная модель была адаптирована для решения задач управления экологической безопасностью арктических коммуникаций за счет реструктуризации основных компонентов и отношений между ними, а также введение новых объектов и их атрибутов.

Интегрированная концептуальная модель экологической безопасности арктических коммуникаций представлена в виде теоретико-множественных конструкций, описывающих как концепты предметной области – экологической безопасности, так и компоненты информационно-аналитической среды управления экологической безопасностью (ЭБ).

$$KM_{уЭБ} = \{O, S, R_s, OS, T, Pro, Cs, IAE, R\}, \quad (1)$$

где

O – множество объектов обеспечения экологической безопасности (objects of ecological safety);

S – множество субъектов обеспечения экологической безопасности (subjects of ecological safety);

R_s – множество ресурсов обеспечения экологической безопасности (resources of ecological safety);

OS – множество организационных структур управления экологической безопасностью (organization structures of ecological safety);

T – множество задач обеспечения экологической безопасности (tasks of ecological safety);

Pro – множество процессов обеспечения экологической безопасности (processes of ecological safety);

Cs – множество кризисных ситуаций при обеспечении экологической безопасности (crisis situations of ecological safety);

IAE – множество компонентов информационно-аналитической среды управления экологической безопасностью арктических коммуникаций (information and analytic environment);

R – множество отношений между элементами концептуальной модели (relations).

В качестве объектов обеспечения экологической безопасности **O** рассматриваются материальные единицы, к которым могут быть напрямую или опосредовано применены управляющие воздействия. В терминах теоретико-множественного описания данное множество представляется в виде:

$$O = \{O_i\},$$

где $O_i = \langle STR^O, Fn^O, Rs^O, S^O, Type^O \rangle$, (2)

STR^O – структура объекта обеспечения ЭБ (structure of object), состоит из элементов и взаимосвязей между ними;

Fn^O – набор функций, который выполняет объект обеспечения ЭБ (functions of object);

Rs^O – множество ресурсов, которые необходимы для функционирования объекта обеспечения ЭБ (resources of object), при чем $Rs^O \subseteq R$;

S^O – множество возможных состояний объекта обеспечения ЭБ (states of object);

$Type^O$ – тип объекта (type of object), например, морской, прибрежный, наземный, воздушный и другие.

Множество субъектов **S** представляет собой набор различных представителей управляющих структур, в сферу деятельности которых входит решение задач по обеспечению экологической безопасности в рамках функционирования систем коммуникаций в арктических регионах. В предлагаемой концептуальной модели экологической безопасности указанное множество представлено следующим образом:

$$S = \{S_j\},$$

$S_j = \langle STR^S, Tr^S, G^S, Rs^S, M^S \rangle$, (3)

где STR^S – структура субъекта обеспечения ЭБ (structure of subject), рассматриваемая с точки зрения системного подхода, а значит состоящая из элементов и связей между ними;

Tr^S – тип субъекта обеспечения ЭБ (type of subject), содержит набор всех возможных видов субъектов, например, ЛПР (лицо, принимающее решение), координатор, организация и другие;

G^S – целевые установки субъекта обеспечения ЭБ (goals of subject), содержат информацию о целевом состоянии объектов управления, за которые ответственен конкретный субъект управления;

Rs^S – множество ресурсов доступных субъекту обеспечения ЭБ (resources of subject), при чем $Rs^S \subseteq R$;

M^S – множество средств (механизмов) управления субъекта обеспечения ЭБ (means of subject), посредством которых конкретный субъект оказывает управляющее воздействие на определенный объект.

Ресурсная обеспеченность задач управления экологической безопасностью арктических коммуникаций представлена в концептуальной модели с помощью введения специального множества ресурсов R . Элементы этого множества описывают атрибуты разнообразных ресурсов, которые используются в функционировании различных систем Арктической зоны, например, информационных, материальных, временных, человеческих и т.д. Множество ресурсов задается следующим образом:

$$R_s = \{R_{s_k}\},$$

$$\text{где } R_{s_k} = \langle Nm^{RS}, Tr^{RS}, Ow^{RS}, Val^{RS}, Lim^{RS}, UM^{RS} \rangle, \quad (4)$$

Nm^{RS} – название ресурса (name of resource);

Tr^{RS} – тип ресурса (type of resource), определяющий происхождение данного ресурса (например, информационный, материальный, временной, человеческий, финансовый и т.д.);

Ow^{RS} – условный владелец ресурса (owner of resource). В рамках концептуальной модели могут быть случаи, когда данный владелец представлен не единственным субъектом управления или когда вообще не имеет потенциального владельца.

Val^{RS} – объем ресурса (value of resource), обладающий количественной характеристикой для оценки его значения;

Lim^{RS} – ограничения на ресурс (limits of resource), в основном представленные в виде граничных условий на объем/количество ресурса;

UM^{RS} – единицы измерения ресурса (unit of measurement).

Реализацию концепции сетецентрического управления экологической безопасностью арктических коммуникаций предлагается реализовывать через создание специальных организационных структур управления, ориентированных на решение определенного класса однотипных задач. Например, на мониторинг и прогнозирование ледовой обстановки в акватории Арктической зоны. Предполагается, что рассматриваемые организационные структуры включают в себя все заинтересованные субъекты управления, которые либо непосредственно имеют возможность осуществлять управление, либо в сфере их интересов функционирования лежит необходимость решения определенного класса задач. В терминах теоретико-множественных описаний данные структуры представлены в концептуальной модели следующим образом:

$$OS = \{OS_i, R^{OS}\},$$

$$\text{где } OS_i = \langle O^{OS}, S^{OS}, G^{OS}, Pro^{OS}, T^{OS} \rangle \quad (5)$$

O^{OS} – множество объектов (objects of organization structure), которыми может управлять организационная структура - $O^{OS} \subseteq O$;

S^{OS} – множество субъектов (subjects of organization structure), образующих организационную структуру, ориентированную на решение определенного класса задач управления – $S^{OS} \subseteq S$;

G^{OS} – множество общих целей организационной структуры (goals of organization structure). Данные цели должны частично или полностью разделять (стремится к их достижению) все субъекты управления, входящие в состав одной организационной структуры.

Pro^{OS} – множество процессов обеспечения ЭБ (processes of organization structure), которыми управляет данная организационная структура – $Pro^{OS} \subseteq Pr$;

T^{OS} – набор задач по обеспечению экологической безопасности (tasks of organization structure), который закреплен за данной организационной структурой – $T^{OS} \subseteq T$.

Стоит отдельно отметить, что сами элементы множества организационных структур могут быть взаимосвязаны между собой разными способами. Способ взаимодействия между организационными структурами обеспечения экологической безопасности в концептуальной модели определяет отношение R^{OS} , которое может задавать такие виды как отношение иерархии, отношение сетевой структуры, или смешанные типы, получаемые в результате сочетания двух предыдущих.

Задачи обеспечения экологической безопасности арктических коммуникаций связаны, в первую очередь, с выявлением и устранением кризисных ситуаций в функционировании социально-экономических, природно-промышленных, транспортных и других видов систем, работающих на территории Арктической зоны. Например, задача планирования судоходной деятельности на акватории Северного морского пути на определенный временной промежуток, или устранение разлива нефтепродуктов в результате транспортировки и т.д. С точки зрения организации управления предлагается рассматривать задачи на трех уровнях: стратегическом, тактическом и оперативном (операционном). Уровень управления влияет на степень детализации описания задачи и определение организационных структур, ответственных за решение задачи. В предлагаемой концептуальной модели задачи T представляются следующим образом:

$$T = \{T_z\},$$

$$\text{где } T_z = \langle Nm^T, L^T, Tm^T, RE^T, IE^T, R^T \rangle, \quad (6)$$

Nm^T – название задачи обеспечения ЭБ (name of task);

L^T – уровень управления задачи обеспечения ЭБ (level of task);

Tm^T – временные характеристики задачи обеспечения ЭБ (times of task), определяющие продолжительность выполнения задачи и временные рамки ее реализации;

RE^T – множество ответственных субъектов управления за выполнение задачи обеспечения ЭБ (responsible entity of task), при чем $RE^T \subseteq S$;

IE^T – множество привлекаемых субъектов к выполнению задачи по обеспечению ЭБ (involved entities to perform the task), при чем $IE^T \subseteq S$;

R^T – множество ресурсов, необходимых для успешной реализации задачи по обеспечению ЭБ (resource of task).

В работе предполагается, что функционирование всех систем, связанных с организацией и использованием арктических коммуникаций, реализуется в виде отдельного набора процессов. В состав набора входят процессы, протекание которых влияет на состояние экологической безопасности исследуемого региона и связанные с реализацией конкретных задач по ее обеспечению. Формальное описание процесса основывается на принципах процессного подхода и представлено в концептуальной модели так:

$$Pro = \{Pro_m\},$$

$$\text{где } Pro_m = \langle Nm^{PR}, IN^{PR}, OUT^{PR}, Ex^{PR}, Rg^{PR}, Prog^{PR} \rangle \quad (7)$$

Nm^{PR} – название процесса (name of process);
 IN^{PR} – множество входных атрибутов процесса (inputs of process), являющихся подмножеством ресурсов ($IN^{PR} \subseteq R$), поступающих на «вход» процесса;

OUT^{PR} – множество выходных атрибутов процесса (outputs of process), являющихся результатами выполнения процесса, для которых справедливо выражение - $OUT^{PR} \subseteq R$;

Ex^{PR} – множество исполнителей процесса (process executors), представляющих собой организационные структуры обеспечения экологической безопасности ($Ex^{PR} \subseteq OS$);

Rg^{PR} – множество регулирующих правил выполнения процесса (regulations of process);

$Prog^{PR}$ – программа мероприятий по реализации процесса (program of process).

Программа мероприятий процесса представляет собой совокупность активностей, выполняющихся в параллельно-последовательном режиме и направленных на получение определенных результатов процесса. В рамках научно-исследовательской работы предлагается использовать проектный подход к организации и управлению отдельными мероприятиями, представленными в программе процесса. За основу была взята концептуальная модель жизненного цикла управления проектами в сфере обеспечения региональной безопасности [1].

В общепринятом представлении проект рассматривается как набор (последовательность) фаз (стадий), которые проходит проект в ходе своего жизненного цикла, а также совокупность свойств, которые характеризует его как относительно обособленный объект исследования:

$$Pr = \langle LF, G^{Pr}, Type, Rs, Rk, T_{LF} \rangle, \quad (8)$$

где

LF – жизненный цикл проекта;

G^{Pr} – цель и задачи проекта;

$Type$ – тип проекта (инвестиционный, некоммерческий, оперативный, стратегический и др.);

Rs – ресурсы, необходимые для реализации проекта;

Rk – риски проекта (для региональной безопасности - угрозы);

T_{LF} – продолжительность жизненного цикла проекта.

$LF = \langle Ph, Lk \rangle$, где Ph – множество фаз жизненного цикла, Lk – множество связей между фазами жизненного цикла.

Такая формальная конструкция жизненного цикла проекта соответствует математическим моделям, зарекомендовавшим себя в теории управления проектами, и позволяет представить его в виде ориентированного взвешенного графа. При такой постановке задачи Ph определяет множество вершин орграфа, а Lk задает множество дуг и их характеристики (направление и веса).

Согласно международному стандарту ISO 21500:2014 по управлению проектами выделяется пять основных фаз проекта: инициализация, планирование, исполнение, управление, завершение.

Сетевая модель представления жизненного цикла проекта позволяет формулировать и решать разнообразные задачи управления. Основной задачей управления является формирование последовательности выполнения фаз проекта

в зависимости от различных параметров проекта и последующая коррекция хода выполнения проекта в зависимости от результатов реализации конкретной стадии.

Формальное описание фазы проекта представляется в виде следующего кортежа:

$$Ph_l = \langle Op^l, Gr^l, St^l, G^l, RS^l, T_0^l, T_g^l, T_{LF}^l, Res^l \rangle, \quad (9)$$

где

$Op^l = \{o_j^l\}, j = \overline{1, N_{Op}^l}$ – множество операций, которые можно/нужно

выполнить на l-ой фазе проекта,

$\exists de N_{Op}^l$ – количество операций на l-ой фазе проекта;

Gr^l – календарный график выполнения l-ой фазы проекта;

St^l – множество стейкхолдеров (заинтересованных лиц) на l-ой фазе проекта;

G^l – цели и задачи l-ой фазы проекта;

RS^l – множество ресурсов, требуемых для выполнения l-фазы;

T_0^l и T_g^l – начальное и конечное время выполнения l-ой фазы проекта;

T_{LF}^l – продолжительность l-ой фазы проекта

Res^l – результаты выполнения l-ой фазы проекта.

Каждая операция описывается с точки зрения процессного подхода, выделяя основной процесс, исходные материалы и информацию, исполнителей, регламенты и процедуры выполнения процесса, а так же его результаты. Календарный график реализации фазы проекта предназначен для описания последовательности выполнения определенных операций и представлен в виде сетевой структуры.

Традиционно в качестве стейкхолдеров рассматривают следующих субъектов: инвестор, заказчик, генконтрактор, руководитель проекта, исполнитель проекта, поставщик, официальный регулятор. При российскими исследователями уже разработаны сторонние математические модели, описывающие поведение каждой из заинтересованных сторон в зависимости от их целевых установок и критериев оценки результатов.

Цели и задачи конкретной фазы выполнения проекта определяются в результате декомпозиции цели и задач самого проекта. Для этого можно использовать методы и алгоритмы функционально-целевого подхода исследования сложных систем. В качестве ресурсов рассматриваются как материальные, так и нематериальные объекты, которые необходимы для успешной реализации фазы проекта. Внутренняя структура множества ресурсов определяется путем выделения в нем отдельных подмножеств потребностей и лимитов (ресурсных ограничений).

Согласно основным положениям проектного менеджмента на временные характеристики проекта и его фаз накладываются ограничения:

$$T_0^l \leq t + T_{LF}^l \leq T_g^l \quad \text{– на время реализации фазы проекта;}$$

$$T_{LF}^l \geq \sum_{l=1}^L T_{LF}^l \quad \text{– на соотношение продолжительности проекта и продолжительности его фаз.}$$

В рамках проводимого исследования результаты выполнения фазы проекта должны влиять на состояние экологической безопасности. Для этого предлагается при оценке результатов реализации фазы сразу формировать ее в виде изменения показателей экологической безопасности.

В результате функционирования разнородных систем коммуникаций в Арктической зоне появляются нештатные ситуации, требующие оперативного вмешательства для возобновления нормального режима деятельности. Такие нештатные ситуации будем называть кризисными. Причинами появления кризисных ситуаций с точки зрения обеспечения экологической безопасности могут являться как техногенные воздействия (выход из строя оборудования, ошибки управления и т.д.), так и природные явления (высокая штормовая активность, сильное электромагнитное излучение естественных источников и др.).

В разработанной концептуальной модели обеспечения экологической безопасности арктических коммуникаций кризисные ситуации представлены следующим образом:

$$Cs = \{Cs_w\},$$

$$\text{где } Cs_w = \langle O^{CS}, S^{CS}, CCS, Re^{CS}, EC^{CS} \rangle \quad (10)$$

O^{CS} – множество объектов обеспечения ЭБ (objects of crisis situation), задействованных в кризисной ситуации;

S^{CS} – множество субъектов обеспечения ЭБ (subjects of crisis situation), задействованных в кризисной ситуации;

CCS – класс кризисной ситуации (class of crisis situation), определяющий степень потенциальной угрозы для экологической обстановки;

Re^{CS} – причины появления кризисной ситуации (reasons of crisis situation);

EC^{CS} – ожидаемые последствия кризисной ситуации (expected consequence of crisis situation). Ожидаемые последствия кризисной ситуации описываются в терминах появления различных видов загрязнения окружающей среды и уменьшения биоразнообразия природных ресурсов.

Для описания загрязнения окружающей среды используется следующая формализованная конструкция:

$$Pl = \langle Nm^{PL}, So^{PL}, Tr^{PL}, Ind^{PL} \rangle, \quad (11)$$

где

Nm^{PL} – наименование загрязнения (name of pollution);

So^{PL} – источники загрязнения (sources of pollution);

Tr^{PL} – тип загрязнения окружающей среды (type of pollution);

Ind^{PL} – множество индикаторов загрязнения (indicators of pollution).

В свою очередь множество индикаторов загрязнения Ind^{PL} включает в себя элементы Ind_j , обладающей следующей структурой:

$$Ind_j = \langle Nm^{IND}, Tr^{IND}, Val^{IND}, Con^{IND} \rangle, \quad (12)$$

где Nm^{IND} – название индикатора загрязнения (name of indicator);

Tr^{IND} – вид индикатора загрязнения (type of indicator), который может принимать два значения – качественный или количественный;

Val^{IND} – значение индикатора загрязнения (value of indicator);

Con^{IND} – ограничения на значение индикатора загрязнения (constraints of indicator), задаваемые нижней и верхней границей значения индикатора, определяющих предельно допустимый диапазон для конкретного индикатора.

В качестве отдельного компонента концептуальной модели вынесена структура информационно-аналитической среды управления экологической безопасностью арктических коммуникаций. Это сделано с целью отделения друг от друга концептов реального и виртуального мира для удобства последующего перехода к онтологическому описанию.

Структура информационно-аналитической среды (ИАС) управления экологической безопасностью арктических коммуникаций состоит из двух фрагментов: описания физического уровня организации распределенной информационной среды и описания логического уровня взаимодействия проактивных сущностей-агентов.

$$IAE = \langle PE, LE \rangle,$$

где PE – «физический» уровень организации информационно-аналитической среды (physic environment);

LE – «логический» уровень организации информационно-аналитической среды (logic environment).

Физический уровень представляет собой описание сетевой структуры размещения клиентских устройств и серверов, составляющих фундамент для функционирования информационно-аналитической среды управления экологической безопасностью арктических коммуникаций.

$$PE = \langle CI, Ser, Lk^{PE} \rangle,$$

где CI – множество клиентских устройств (clients of physic environment), предоставляющих доступ к информационно-аналитической среде;

Ser – множество серверных узлов (servers of physic environment), реализующих функции по предоставлению требуемых данных и сервисов пользователям через программные агенты;

Lk^{PE} – множество связей между клиентскими и серверными узлами «физического» уровня (links of physic environment).

Логический уровень представляет собой сетевую структуру, элементами которого являются проактивные сущности – когнитивные агенты, являющиеся виртуальными представителями конкретных субъектов обеспечения экологической безопасности в информационно-аналитической среде.

$$LE = \langle A, Lk^{LE} \rangle,$$

где A – множество агентов логического уровня (agents logic environment);

Lk^{LE} – множество связей между агентами логического уровня (links of logic environment).

Стоит отметить, что агенты интерпретируются в концептуальной модели как алгоритмические исполнители нетерминального уровня. Они являются исполнителями задач субъектов управления безопасностью в распределенной виртуальной среде. Это обеспечивает возможность моделирования деятельности каждого субъекта управления как автономной проактивной сущности с собственными интересами и целями. Агенты участвуют в решении задач управления безопасностью на всех этапах жизненного цикла угроз экологической безопасности посредством формирования коалиций между агентами, представляющих собой интерпретацию организационной структуры обеспечения безопасности в виртуальном пространстве. Такие проактивные объекты обеспечивают не только имитацию поведения субъектов управления, но и предоставляют информационные сервисы другим агентам, реализуют поиск субъектов совместной деятельности, осуществляют

информационный мониторинг и оценку показателей экологической безопасности в многомерном пространстве критериев[3].

$$A = \{A_i\},$$
$$A_i = \langle ID, R^A, Loc^A, Fn^A, Sr^A, Inf^A \rangle, \quad (13)$$

где ID – идентификатор агента на логическом уровне ИАС (identifier of agent);

R^A – роль агента (role of agent), определяет назначение агента в ИАС, может принимать значения – исполнитель, координатор, ответственное лицо, заказчик и т.д.;

Loc^A – местоположение агента в ИАС (location of agent), искусственно введенный атрибут, отвечающий за возможность однозначного определения места «физического» нахождения агента в среде. Например, при реализации ИАС на базе технологий Интернета в роли данного местоположения выступает IP и MAC адреса сетевого узла, где располагается агент.

Fn^A – набор функций (functions of agent), который может выполнять агент, зависит от текущей роли агента;

Sr^A – набор сервисов (services of agent), которые может предоставить агент другим агентам;

Inf^A – множество информационных ресурсов (information of agent), которыми владеет агент.

В рамках разработанной концептуальной модели обеспечения экологической безопасности арктических коммуникаций предлагается все отношения между концептами разбить на три общих класса в зависимости от того, между какими компонентами они определены.

$$R = \{R_{SD}, R_{SDLE}, R_{LEPE}\},$$

где

R_{SD} – обобщенный класс отношений, содержащий в себе отношения, устанавливающие взаимосвязи между концептами предметной области (relations of subject domain). Например, отношение включения субъекта управления в состав организационной структуры R_{SOS} ($R_{SOS}: S \rightarrow OS$).

R_{SDLE} – обобщенный класс отношений, содержащий в себе отношения, устанавливающие взаимосвязи между концептами предметной области и логического уровня информационно-аналитической среды (relations of subject domain and logic environment). Например, отношение представления субъекта управления в виртуальной среде R_{SA} ($R_{SA}: S \rightarrow A$).

R_{LEPE} – обобщенный класс отношений, содержащий в себе отношения, устанавливающие взаимосвязи между концептами логического и физического уровней информационно-аналитической среды (relations of logic and physic environments). Например, размещение проактивного агента в виртуальной среде R_{ACL} ($R_{ACL}: A \rightarrow CI$).

На текущем этапе реализации научно-исследовательского проекта каждый из вышеперечисленных классов отношений частично определен в концептуальной модели. Продолжается работа по формированию полностью покрывающих множеств отношений между выделенными объектами формализованного описания в зависимости от решаемых задач управления.

Для применения разработанной формализованной концептуальной модели обеспечения экологической безопасности арктических коммуникаций в

качестве структурированной базы знаний было осуществлено представление концептов и отношений модели в виде прикладной онтологии. Для построения онтологического описания использовался редактор онтологий Protégé 5.2.

На рисунке 2 представлен фрагмент онтологического описания, содержащий информацию об источниках загрязнения окружающей среды.

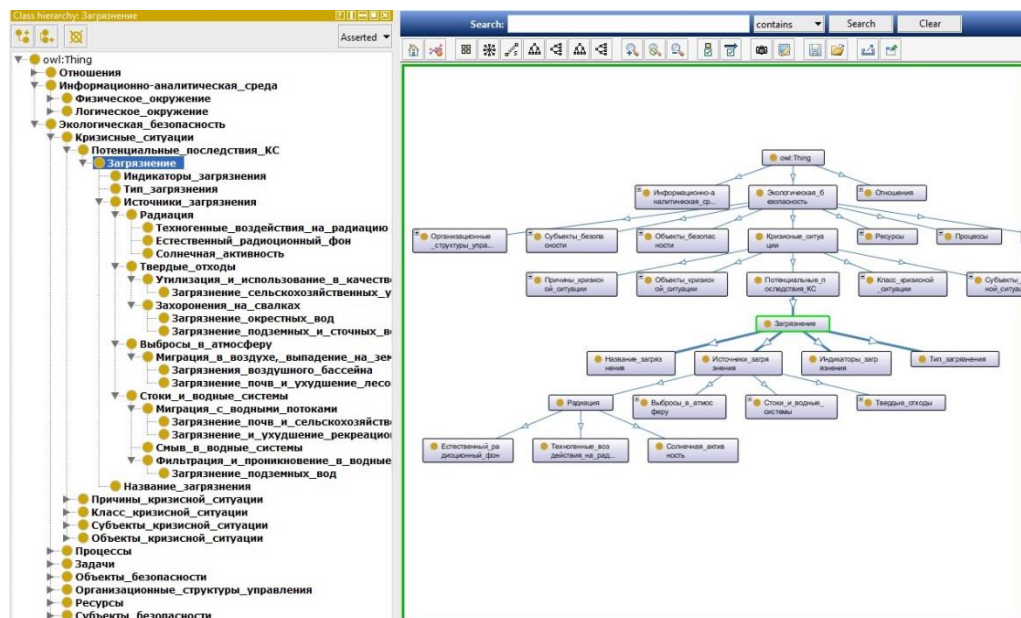


Рис. 2. Фрагмент прикладной онтологии обеспечения экологической безопасности арктических коммуникаций

В перспективе построенная прикладная онтология обеспечения экологической безопасности арктических коммуникаций будет использоваться в качестве фундамента для разработки программных средств, обеспечивающих развертывание и функционирование соответствующей информационно-аналитической среды. На данном этапе исследования прикладная онтология использовалась для автоматизированного синтеза комплекса имитационных моделей обеспечения экологической безопасности.

Заключение

В эпоху становления цифровой экономики РФ разработка методов и средств информационной поддержки управления сложными природно-промышленными и социально-экономическими системами является приоритетным направлением научно-прикладных исследований в сфере информационных технологий. К подобному типу сложных систем относятся и разнообразные компоненты арктических коммуникаций, в частности, Северный морской путь. Специфика арктических коммуникаций, проявляющаяся в погодно-климатических и технологических особенностях ведения хозяйственной деятельности, заставляет предъявлять повышенные требования к средствам и методам управления в части оперативности обработки информации и принятия

управленческих решений. Для удовлетворения таких ограничений в работе предлагается использовать принципы сетецентрического управления для организации единой виртуальной среды обеспечения экологической безопасности.

На данном этапе исследования предложено формализованное концептуальное описание обеспечения экологической безопасности арктических коммуникаций, позволяющее осуществить структуризацию знаний о предметной области с помощью выделения основных концептов и отношений между ними. Для практического использования концептуальная модель была трансформирована в прикладную OWL-онтологию, что позволит использовать ее для реализации автоматизированных процедур генерации структуры имитационных моделей и формирования хранилища данных.

Литература

1. Быстров В.В., Маслобоев А.В. Концептуальная модель жизненного цикла управления проектами в сфере обеспечения региональной безопасности // Информационные системы и технологии. – 2018. – №3(107). – С. 48-56.
2. Куватов, В.А. Потенциал Северного морского пути Арктической зоны России. Факторы и стратегия развития / В.А. Куватов, Д.В. Козьмовский, Н.В. Шаталова // Интернет-журнал «НАУКОВЕДЕНИЕ» – Выпуск 6 (25) – 2014 – С.1-16 (режим доступа: <https://naukovedenie.ru/PDF/20TVN614.pdf>)
3. Маслобоев, А.В. Исследование и разработка моделей и методов информационной поддержки управления региональной безопасностью (на примере Мурманской области) / А.В. Маслобоев / Диссертация на соискание степени доктора наук – Апатиты: КНЦ РАН – 2016, 314 с.
4. Смирнов А.С., Стариченков А.Л., Стариченкова Е.М., Малыгин И.Г. Управление безопасностью водных транспортных средств при чрезвычайных ситуациях: монография // под редакцией В.С. Артамонова. – СПб: Санкт-Петербургский университет ГПС МЧС России, 2013. –184 с.

Сведения об авторах

Быстров Виталий Викторович – к.т.н., старший научный сотрудник

e-mail: bystrov@iimm.ru

Vitaliy V. Bystrov – PhD, senior researcher

Халиуллина Дарья Николаевна – к.т.н., научный сотрудник

e-mail: khaliullina@iimm.ru

Darya N. Khaliullina – PhD, researcher

УДК 004.946, 331.5.024.52
DOI: 10.25702/KSC.2307-5252.2018.10.61-76

Е.С.Давидюк, М.Г.Шишаев, В.В.Быстров

Институт информатики и математического моделирования ФИЦ КНЦ РАН

ПРОГНОЗНОЕ МОДЕЛИРОВАНИЕ ДЛЯ МОНИТОРИНГА И УПРАВЛЕНИЯ КАДРОВЫМ ОБЕСПЕЧЕНИЕМ ПРОГРАММ РЕГИОНАЛЬНОГО РАЗВИТИЯ*

Аннотация

В статье представлен анализ проблематики прогнозирования кадрового обеспечения на региональном уровне, а также обзор существующих методов и моделей для решения этой задачи. Сформулированы качественные требования к идеальной модели (технологии) кадрового прогнозирования. Показано, что существующие модели значительно разнятся по используемым входным данным, внутренним механизмам моделирования и качеству получаемых результатов. Предложена концептуальная схема агентной технологии кадрового прогнозирования, потенциально позволяющей получить новое приближение к идеальному образу. Сформулированы ключевые задачи, решение которых необходимо для реализации технологии.

Ключевые слова:

модель, прогнозирование, кадровая потребность, региональное развитие

E.S.Davidyuk, M.G.Shishaev, V.V.Bystrov

FORECAST SIMULATION FOR MONITORING AND MANAGEMENT OF STAFFING NEED FOR REGIONAL DEVELOPMENT PROGRAMS

Abstract

The article considers an analysis of the problems of forecasting staffing at the regional level, as well as a review of existing methods and models for solving this problem. The qualitative requirements for the ideal model (technology) of personnel forecasting are formulated. It is shown that the existing models differ significantly in the input data used, the internal modeling mechanisms and the quality of the results obtained. A conceptual scheme of agent prediction technology has been proposed, potentially allowing a new approximation to the ideal image to be obtained. Formulated key tasks, the solution of which is necessary for the implementation of technology.

Keywords:

model, forecasting, staffing need, regional development

Введение

Прогнозирование кадровой потребности региональной экономики является актуальной практической задачей. В региональном управлении она возникает в различных контекстах, определяющих конечную цель прогнозирования и качественные требования к его результату. В данной работе рассматривается два из них, актуальных для региональных органов управления образованием и органов регулирования занятости населения:

* Работа выполнена при финансовой поддержке РФФИ, проект № 16-29-12878 (офи_м) «Развитие методов идентификации динамических моделей со случайными параметрами и их применение для прогнозирования миграции в Евразии».

1. определение оптимальной численности приема в региональные организации высшего и среднего профессионального образования (ВО и СПО);
2. мониторинг и управление уровнем занятости населения региона.

В первом случае конечной целью является получение структуры дополнительной кадровой потребности экономики в разрезе специальностей и направлений подготовки, а также уровней образования. Во втором случае целью является получение прогноза потребности в разрезе профессий и должностей, а также видов экономической деятельности.

Несмотря на довольно большое количество работ по данной проблематике, задача не утратила свою актуальность как с практической, так и с научной точек зрения. Кадровое обеспечение региона является системой со сложной внутренней структурой, подверженной влиянию множества разнородных внешних факторов – рыночной конъюнктуры, миграционных процессов, развитию производственных технологий и других. На основании приведенного в данной статье обзора современных отечественных подходов к прогнозному моделированию кадровой потребности можно констатировать, что по-прежнему не существует «идеальных» моделей, способных давать достаточно качественный результат при выполнении ограничений на расходуемые временные и материальные ресурсы, с учетом динамики внешних и внутренних факторов и при имеющихся параметрах доступности и качества исходных данных. С другой стороны, развитие методов системного анализа и информационных технологий открывает новые перспективы для повышения эффективности решения данной задачи. В частности, представляется перспективным применение агентного подхода к прогнозному моделированию кадровой потребности, современных методов моделирования макроэкономических и демографических процессов (в том числе миграции), методов анализа больших данных, а также информационных технологий, основанных на формализованных знаниях.

В данной работе представлен анализ проблематики прогнозирования кадрового обеспечения на региональном уровне, а также обзор существующих методов и моделей для решения этой задачи. Сформулированы качественные требования к идеальной модели (технологии) кадрового прогнозирования. Показано, что существующие модели значительно разнятся по используемым входным данным, внутренним механизмам моделирования и качеству получаемых результатов. Предложена концептуальная схема агентной технологии кадрового прогнозирования, потенциально позволяющей получить новое приближение к идеальному образу, сформулированы ключевые задачи, решение которых необходимо для ее реализации.

1. Общая характеристика проблематики прогнозирования кадрового обеспечения на региональном уровне

Первая подзадача прогнозирования кадрового обеспечения ориентирована на поддержку управления системой подготовки кадров на региональном уровне. Существующая схема управления распределением государственного задания на подготовку профессиональных кадров в образовательных организациях подразумевает вовлечение региональных органов исполнительной власти (РОИВ) либо на этапе формирования структуры госзадания (в случае с СПО), либо на

этапе его предварительного согласования (в случае с высшим образованием). С учетом сроков обучения по образовательным программам различного уровня в вузах и организациях СПО, интересующий РОИВ горизонт прогнозирования чаще всего составляет от 2 до 6 лет. В случае планирования подготовки высшего медицинского персонала или научных работников требуемый горизонт возрастает до 8 и 9-10 лет. Исходя из современной практики организации образовательного процесса, эффективного с экономической точки зрения, минимальный размер группы обучающихся по одной специальности/направлению составляет около 10 человек, что обуславливает требования к масштабу (детальности) получаемых прогнозных оценок. Задача формирования кадрового прогноза в такой постановке возникает регулярно, один раз в год.

Второй подзадачей является получение прогноза потребности в разрезе профессий и должностей, а также видов экономической деятельности. Поскольку прогноз строится для мониторинга ситуации на рынке труда с целью оперативного реагирования на ее потенциальную динамику (обеспечение занятости населения и текущих потребностей работодателей), задача прогнозирования возникает постоянно. Однако, в силу ограниченности ресурсов, РОИВ организует формирование подобного прогноза, как правило, один раз в год. В данном случае масштаб результирующих данных минимален – требуется прогноз с точностью до 1 человека. Горизонт прогнозирования, в идеале, также должен стремиться к минимуму.

Исходными данными, доступными для формирования кадрового прогноза являются: данные официальной статистики; данные опросов работодателей; экспертные оценки; прогнозы макроэкономических показателей.

Социально-экономическая ситуация, в контексте которой формируется прогноз, как правило, характеризуется такими особенностями:

- структурная динамика региональной экономики, проявляющаяся как в появлении, так и исчезновении инвестиционных проектов различного масштаба;
- зависимость региональной экономической системы от глобальной рыночной конъюнктуры и политических процессов (экономические санкции);
- наличие хозяйствующих субъектов, лишь ограниченно управляемых государственными органами;
- значительные миграционные потоки;
- глобальная и локальная динамика структуры трудовых ресурсов (появление новых и исчезновение старых профессий, изменение требований к квалификации).

Перечисленные особенности обуславливают чрезвычайно неблагоприятные для задач моделирования систем исходные условия: низкую достоверность и неполноту исходных данных; высокую структурную динамику объекта моделирования.

Проблема усугубляется, если рассматривается система масштаба макрорегиона (например, АЗРФ), при этом подобная («макрорегиональная») постановка задачи прогнозирования кадровой потребности становится все более актуальной.

Таким образом, необходима информационная технология прогнозного моделирования кадровой потребности, обладающая следующими возможностями:

1. получение прогнозов в указанных выше разрезах с заданным уровнем детализации;
2. быстрое изменение структуры модели, желательно без привлечения специалистов в сфере ИТ и моделирования;
3. учет комплекса макро- и микроэкономических факторов, влияющих на рынок труда;
4. получение адекватного прогноза в условиях низкой достоверности и неполноты исходных данных.

В следующих разделах статьи рассмотрены некоторые наиболее известные в отечественной практике методы и технологии формирования кадрового прогноза в проекции на перечисленные функциональные требования. Наименования рассмотренных подходов, методов и моделей, обозначенные в соответствующих заголовках, условны.

2. Методика прогнозирования региональной кадровой потребности РГППУ (А.Г. Мокронос)

Основные аспекты методики:

- сочетание макроэкономических расчетов потребности в кадрах в целом и по видам экономической деятельности с микроэкономическими методами анализа потребности в кадрах предприятий в разрезе уровней образования, специальностей и профессий, уровней квалификации с последующей балансировкой макроэкономических и микроэкономических расчетов;

- планирование воспроизводства рабочих мест совместно с прогнозом обеспечения экономики региона трудовыми ресурсами;

- среднесрочное и долгосрочное прогнозирование;
- многофакторная модель;
- синтез технологического и экономического подхода к прогнозированию;
- использование результатов прогноза баланса трудовых ресурсов.

Модель включает три группы факторов, влияющих на обеспечение кадрами региональной экономики:

а) факторы, формирующие потребность [8]:

- свободные вакансии;
- предполагаемое выбытие персонала по возрасту;
- создаваемые новые рабочие места;

б) факторы, уменьшающие потребность в персонале:

- ликвидация рабочих мест;
- рост производительности труда;

в) факторы, формирующие предложение на рынке труда:

- привлечение в экономику лиц, вступающих в трудоспособный возраст;
- привлечение в экономику выпускников учреждений профессионального образования;

- увеличение экономической активности трудоспособного населения в результате успешного социально-экономического развития общества;

- увеличение экономической активности населения старше трудоспособного возраста;

- привлечение в экономику дополнительного числа трудовых мигрантов.

Источниками информации являются:

- статистическая информация о численности населения, о балансе трудовых ресурсов, об отраслевой, половозрастной и образовательной структурах занятого и экономически активного населения, о занятом населении по профессионально-квалификационным группам, о выпуске учреждениями профессионального образования, о валовом региональном продукте, об обороте по видам экономической деятельности;

- информация от службы занятости населения региона о структуре незаполненных вакансий и составе зарегистрированных безработных в разрезе специальностей и профессий;

- информация о перспективном развитии экономики (исследования на основе вторичного анализа статистических данных и выявления трендовых тенденций в сфере занятости, программы и проекты социально-экономического развития региона и отраслей экономики региона, планы и программы инвестиций в реальный сектор экономики региона, планы и программы развития крупнейших предприятий региона);

- данные мониторинга заявленной организациями региона потребности в кадрах в разрезе уровней образования, специальностей, профессий и уровней квалификации;

- результаты обследования заявленного домохозяйствами спроса на услуги образовательных учреждений и отдельные образовательные программы [7];

- экспертные оценки рабочих мест на соответствие современным технологическим и социально-экономическим требованиям [7].

Основная схема моделирования кадровой потребности.

Общая потребность экономики в подготовке специалистов определяется на основе учета влияния первых двух групп факторов.

Основной составляющей региональной потребности экономики в персонале является потребность в кадрах каждой из отраслей.

Потребность отраслей экономики в кадрах складывается из потребности в персонале предприятий этих отраслей. Также учитывается движение рабочих мест в отраслевом разрезе – их ликвидация и создание. Создание и ликвидация рабочих мест в отрасли учитываются с помощью системы коэффициентов, понижающих и повышающих потребность в персонале – коэффициент увеличения объемов производства в отрасли и коэффициент технического и технологического развития отрасли.

Потребность каждого предприятия в персонале формируется исходя из имеющихся в настоящий момент вакансий, информации о выбытии персонала по возрасту, планов перспективного развития предприятия, в связи с созданием новых рабочих мест для реализации расширения производства и ликвидацией рабочих мест вследствие их полного износа или отсутствия необходимости их существования.

Заказ предприятиями специалистов и квалифицированных рабочих складывается из:

- заказа предприятиями образовательным учреждениям количества персонала с высшим уровнем профессионального образования;
- заказ предприятиями образовательным учреждениям количества персонала со средним уровнем профессионального образования;
- заказ предприятиями образовательным учреждениям количества персонала с начальным уровнем профессионального образования.

В целом модель прогноза потребностей в подготовке персонала по уровням образования, профессиям и специальностям имеет следующий вид:

$$П_{\text{Э}}^{\text{нодз}} = \sum_{g=1}^G \left[\sum_{c=1}^{C,P} \left(\sum_{i=1}^I \left(\sum_{r=1}^R Z_{ir}^{cg} \cdot k_{\text{досч}}^i \cdot k_{V_{np}}^i \cdot k_{\text{тех}}^i \right) \right) \cdot k_{\text{д/н}}^{cg} \cdot k_{\text{н/а}}^{cg} \cdot k_{\text{д/о}}^{cg} \cdot k_{\text{н/н}}^{cg} \cdot k_{\text{н/о}}^{cg} \right],$$

где $П_{\text{Э}}^{\text{нодз}}$ – количественная потребность в подготовке персонала по уровням образования, профессиям и специальностям;

G – количество уровней образования;

C, P – количество специальностей и профессий соответственно;

Z_{ir}^{cg} – заказ предприятий образовательным учреждениям на подготовку квалифицированных кадров по уровням образования;

$k_{\text{досч}}^i$ – коэффициент досчета до полного круга предприятий отрасли [8];

$k_{V_{np}}^i$ – коэффициент увеличения объемов производства в отрасли;

$k_{\text{тех}}^i$ – коэффициент технического и технологического развития отрасли;

$k_{\text{д/н}}^{cg}$ – коэффициент движения рабочей силы по профессиям, специальностям (в рамках одной отрасли);

$k_{\text{н/а}}^{cg}$ – коэффициент экономически неактивного населения (не приступающего к работе после получения образования);

$k_{\text{д/о}}^{cg}$ – коэффициент движения рабочей силы по уровням образования (специалист приступает к работе не в соответствии с полученным уровнем образования);

$k_{\text{н/н}}^{cg}$ – коэффициент привлечения специалистов из других отраслей;

$k_{\text{н/о}}^{cg}$ – коэффициент привлечения специалистов с другим уровнем образования.

Достоинства модели:

- сочетание макроэкономических расчетов с микроэкономическими методами анализа;
- планирование воспроизводства рабочих мест совместно с прогнозом обеспечения экономики региона трудовыми ресурсами;
- среднесрочное и долгосрочное прогнозирование;
- многофакторная модель;
- синтез технологического и экономического подхода к прогнозированию.

Недостатки реализации модели:

- исходная информация для прогнозирования не является оперативной и требует дополнительной обработки и расчетов;

- не указано об объеме и способе анкетирования предприятий. Современные предприятия не заинтересованы в предоставлении сведений для обеспечения конфиденциальности информации и конкурентоспособности;
- потребность в кадрах определяют сами предприятия, а методики определения в крупных, средних, малых предприятиях не согласованы;
- малые, средние и крупные предприятия имеют разные структуры занятых, что создает трудность в правильном определении коэффициента $k_{досч}^i$;
- не учитывается миграция, вахта, иностранная рабочая сила;
- модель рассчитана для прогнозирования кадровой потребности одного региона;
- система применяемых коэффициентов нуждается в постоянной корректировке в зависимости от социально-экономического развития;
- используемые в модели корректирующие коэффициенты не имеют прямых аналогов в реальном мире в виде тех или иных социально-экономических показателей или свойств реальных экономических и/или социальных объектов; это априори создает сложности для адекватного определения и оценки правильности их значений даже с привлечением экспертов.

3. Сценарный подход в прогнозировании развития рынка труда. (З. А. Васильева, Т. П. Лихачева, Н. В. Разнова, И. В. Филимоненко)

Разработана для определения кадровой потребности Красноярского края.

Основные аспекты методики:

Данная методика основана на макроэкономической модели, с учетом возможных структурных сдвигов в экономике региона – изменений темпов экономического роста [1]. Потребность в специалистах определяется валовым региональным продуктом (ВРП) и его планируемым развитием. Прогноз валового продукта на всех уровнях (федеральном, региональном, отраслевом) строится на основе использования системы показателей: инвестиции в основной капитал, численность занятых в экономике, среднедушевые денежные доходы населения региона и др. Потребность в специалистах на рынках труда городов и районов определяется сложившейся региональной структурой занятости. Потребность в специалистах в профессионально-квалификационном разрезе определяется сложившейся структурой спроса на рынке труда. Основным источником покрытия потребности в специалистах на рынке труда – ВПО, СПО, НПО.

Источниками информации являются:

- статистическая отчетность регионального комитета государственной статистики;
- данные ВПО, СПО, НПО региона о подготовке специалистов: приеме и выпуске обучающихся в разрезе специальностей;
- оценки экспертов – руководителей предприятий об оптимальной структуре занятости;
- матрицы ПКС по уровням образования, разработанные в г. Петрозаводске;
- программа социально-экономического развития региона, разработанная правительством региона;

- данные региональных отделений социальной защиты населения (СЗН) о потребности предприятий в специалистах.

Основная схема моделирования кадровой потребности.

Сбор и обработка исходной информации.

Прогнозирование среднегодовой численности занятых по сценарным вариантам («невмешательства», «экономический рост», «агрегированный») [2]. Используются макроэкономические взаимосвязи между производительностью труда и инвестиционным капиталом, при этом поправочные коэффициенты рассчитываются с помощью регрессионного анализа на основе ретроспективных данных.

Построение структуры занятости в экономике по видам экономической деятельности (ВЭД) по сценарным вариантам («невмешательства», «управления»). При этом производительность труда для всех прогнозируемых лет остается постоянной, а численность занятых по ВЭД распределяется исходя из темпов роста объемов выпуска товаров и услуг по ВЭД, взятых из программы социально-экономического развития региона.

Распределение общей численности занятых и потребности в профессиональных кадрах по ВЭД и уровням образования (УО) с учетом структуры образования региона – используются поправочные коэффициенты, определенные экспертным путем.

Расчет и распределение и потребности в профессиональных кадрах по ВЭД и УО. Распределение происходит на основе структуры занятых по ВЭД и УО, определенных на предыдущем этапе с использованием матрицы ПКС по уровням образования и поправочного коэффициента.

Достоинства модели:

- прогнозирование среднегодовой численности занятых по сценарным вариантам [6];

- используются макроэкономические взаимосвязи между производительностью труда и инвестиционным капиталом;

- построение структуры занятости в экономике по видам экономической деятельности по сценарным вариантам.

Недостатки реализации модели:

- не учтены дополнительные факторы – миграция, демография, влияние технической модернизации на качество рабочих мест;

- используются макроэкономические модели, не учитывающие структуру и развитие предприятий;

- средний и малый бизнес не принимается в расчет при построении сценарных вариантов;

- данные органов статистики и СЗН не являются оперативной и полной;

- поправочные коэффициенты определяются экспертным путем и при их неизменном значении для всего прогнозного периода не отражают происходящие перемены в структуре образования и экономики региона;

- прогнозирование экономического роста и определение сценарных вариантов базируется на сложившейся структуре региональной экономики и тенденции развития, поэтому используемая модель не учитывает и не сможет учесть существенные изменения в структуре региональной экономики, и как следствие, прогнозы кадровой потребности не будут достоверными.

4. Математическая модель прогнозирования потребностей региональных экономик в специалистах с профессиональным образованием

Данная методика была разработана в Петрозаводском государственном университете, авторы В. Н. Васильев, В. А. Гуртов и Е. А. Питухин.

Методика позволяет определить прогнозы потребности экономики в квалифицированных кадрах по уровням, направлениям и объемам подготовки и формировать на этой основе государственное задание на подготовку специалистов для системы профессионального образования Российской Федерации.

Основные аспекты методики:

Основана на системе макроэкономических математических моделей, описывающих динамику взаимосвязей экономики, рынка труда и профессионального образования.

Анализируемые параметры [3]:

- численность постоянного населения;
- численность трудоспособного населения;
- распределение численности занятого населения по отраслям экономики и социальной сферы;
- среднегодовая численность промышленно-производственного персонала;
- уровень образования среди занятых в различных отраслях экономики;
- матрицы соответствия структуры подготовки специалистов с высшим, средним и начальным профессиональным образованием по 28 группам и структуры отраслей народного хозяйства по 14 отраслям с детализацией промышленности по 15 подотраслям;
- коэффициент ротации для занятого населения.

Источниками информации являются:

- статистическая отчетность регионального комитета государственной статистики;
- данные ВПО, СПО, НПО региона о подготовке специалистов: приеме и выпуске обучающихся в разрезе специальностей;
- экспертные оценки для формирования матрицы соответствия «видов экономической деятельности – 28 укрупнённых групп специальностей» [4].

Основная схема моделирования кадровой потребности.

Моделируется численность трудовых ресурсов для каждого вида экономической деятельности:

$$L(t_{i+1}) = L(t_i) + S(t_i) - L^-(t_i),$$

где $L(t_{i+1})$ – численность трудовых ресурсов, доступная в t_{i+1} году; $L(t_i)$ – среднегодовая численность занятых в экономике в t_i году; $S(t_i)$ – дополнительное предложение на рынке труда в период t_i (складывается из выпускников системы профессионального образования, переподготовленных, переобученных безработных, приехавших трудовых мигрантов); $L^-(t_i)$ – поток выбывания трудовых ресурсов с рынка труда.

$$L^-(t_i) = L_{SC}^-(t_i) + L_M^-(t_i),$$

где $L_{SC}^-(t_i)$ – естественно-возрастное выбывание (коэффициент ротации),

$L_M^-(t_i)$ – эмигранты.

Определяется дополнительная потребность в специалистах:

$$\Delta D(t_{i+1}) = L^*(t_{i+1}) - L(t_i) + L^-(t_i) - S(t_i),$$

где $L^*(t_{i+1})$ – прогноз среднегодовой численности занятых в t_{i+1} году;

Полученный прогноз работников разделяется по уровням профессионального образования:

$$R^e(i+1) = k^e \cdot \Delta D(t_{i+1}), \quad \sum_{n=1}^e k^n = 1,$$

где $R^e(i+1)$ – прогнозные потребности в специалистах по уровням образования; e – уровень образования, k^e – доля специалистов, работающих по ВЭД, с уровнем образования e .

Расчёт прогнозной ежегодной потребности в специалистах по уровням образования и укрупнённым группам специальностей:

$$N_y^e(i+1) = \sum A_{n,y}^e \cdot R^e(i+1),$$

где $N_y^e(i+1)$ – прогноз потребности в специалистах с уровнем образования e и направлением подготовки y , $A_{n,y}^e$, – матрица профессионального квалификационного соответствия «видов экономической деятельности – 28 укрупнённых групп специальностей» [4].

Достоинства модели:

- сочетание макроэкономических моделей с микроэкономическими методами анализа;
- определение прогнозов потребности экономики в квалифицированных кадрах по уровням, направлениям и объёмам подготовки;
- учитывается взаимодействие субъектов рынка труда и сферы образования.

Недостатки реализации модели:

- прогноз среднегодовой численности занятых основан на индексе изменения ВРП и не учитывает возможные структурные сдвиги в экономике региона;
- используется постоянный коэффициент ротации, который предполагает равномерное половозрастное распределение занятого населения, что не соответствует действительности;
- коэффициенты матрицы соответствия «видов экономической деятельности – 28 укрупнённых групп специальностей» должны быть определены для каждого региона, что существенно снижает универсальность данной методики.

5. Математическая модель определения кадровой потребности реального сектора экономики республики Крым (Н.И. Клевец, Е.А.Полищук)

Разработана в Крымском федеральном университете имени В.И. Вернадского для снижения профессионально-квалификационного дисбаланса на рынке труда и покрытия кадрового дефицита с учетом перспективных направлений и потребностей для эффективного развития Республики Крым [9].

Основные аспекты методики:

В основе лежит метод математического программирования – сепарабельное квадратичное программирование. Для решения поставленной задачи использовалась система автоматизированных вычислений MathCAD.

Источниками информации являются:

- данные опроса работодателей всех форм собственности ГКУ РК «Центр занятости населения» (официальная статистика);
- данные сайтов ВПО, СПО региона о подготовке специалистов: приеме и выпуске обучающихся в разрезе специальностей;
- экспертные оценки для определения приоритетов видов экономической деятельности.

Основная схема моделирования кадровой потребности.

Выбраны 10 видов экономической деятельности, представленных в республике Крым.

Определена потребность регионального рынка труда в рабочей силе и наличии бюджетных мест в организациях ВО и СПО по данным опроса работодателей и сайтов ВПО, СПО.

Минимизируется целевая функция – взвешенная сумма квадратов отклонений принятых на бюджетное обучение абитуриентов от кадровой потребности [9]:

$$F(X) = \sum_{n=1}^{10} \left[w_n (D_n - X_n)^2 \right] \quad \text{с учетом ограничений}$$
$$\sum_n X_n \leq I_{\max}, \quad X_{\min} \leq X \leq X_{\max},$$

где X – вектор искомых оптимальных планов приема абитуриентов по направлениям подготовки; w – вектор весовых коэффициентов направлений подготовки специалистов; D – вектор спроса на специалистов по направлениям подготовки; n – номер направления подготовки специалистов; I_{\max} – суммарная численность абитуриентов по всем направлениям подготовки.

Достоинства модели:

- исходные данные не требуют предварительной подготовки и преобразования;
- модель позволяет быстро оценить примерный план приема по направлениям подготовки на текущий момент.

Недостатки реализации модели:

- исходные данные не являются строго детерминированными;
- моделирование осуществляется только на текущий момент;
- не учитываются возможные варианты развития экономики региона;
- не учитывается миграция экономически активного населения;

- весовые коэффициенты приоритетных направлений подготовки специалистов определяются экспертным путем только на текущий момент состояния экономики.

6. Моделирование кадровой потребности предприятий и отраслей экономики МГУЭСИ

Авторы – Н.Е. Егорова, д.э.н., профессор, Центральный экономико-математический институт РАН, Е.В. Ярошенко, Московский государственный университет экономики, статистики и информатики [5].

Основные аспекты методики:

В модели используется производственная функция человеческого капитала трудовых ресурсов и оптимизационная модель. Критерий оптимальности – максимальная совокупная полезность труда работников всех уровней образования. Прогноз распределения рабочей силы одной из отраслей экономики по трем уровням образования

Источниками информации являются:

- статистическая информация об объеме выпуска совокупного продукта, численности работников, физического капитала;
- затраты предприятия на содержание работников с определенным уровнем образования.

Основная схема моделирования кадровой потребности.

Накопление статистической информации за предыдущие периоды.

Оценка производственной функции по ретроспективным данным с помощью метода наименьших квадратов [5]

$$Y(t) = a_1 K(t)^{a_2} \left(\sum_{i=1}^3 L_i(t) e^{t \cdot h_i} \right)^{1-a_2},$$

где $Y(t)$ – объем выпуска совокупного продукта в году t ; $K(t)$ – объем основных производственных фондов предприятий в году t ; L_i – численность работников i -го уровня образования, участвующих в процессе производства в году t ; h_i – коэффициент человеческого капитала работника с i -м уровнем образования, характеризующий темп прироста выпуска совокупного продукта под влиянием накопления человеческого капитала работника с i -м уровнем образования (определяется с помощью МНК); a_1 – коэффициент совокупной производительности факторов производства, характеризующий уровень развития технологии (определяется с помощью МНК); a_2 – коэффициент эластичности по капиталу, позволяющий оценить вклад физического капитала в совокупный выпуск продукта (определяется с помощью МНК); $1 - a_2$ – коэффициент эластичности по труду, позволяющий оценить вклад уровня занятости в совокупный выпуск продукта; t – номер года.

Максимизация совокупной полезности труда работников всех уровней образования:

$$a_1(1-a_2) \sum_{i=1}^3 e^{t \cdot h_i} \left(\frac{K(t)}{\sum_{i=1}^3 L_i(t) e^{t \cdot h_i}} \right)^{a_2},$$

при ограничениях:

$$\sum_{i=1}^3 L_i(t) = L(t)$$

$$\sum_{i=1}^3 L_i(t) w_i(t) \leq \varphi(t) Y(t)$$

$$\sum_{i=1}^3 Q_i(t) \geq \sum_{i=1}^3 Q_i^{ucx}(t)$$

$$K(t) > 0, Y(t) > 0, L(t) > 0, L_i(t) > 0, a_1 > 0, a_2 > 0, h_i > 0,$$

где $\varphi(t)$ – удельный вес оплаты труда наемных работников в объеме совокупного выпуска продукта; $w_i(t)$ – затраты предприятия на содержание работников с i -м уровнем образования; $Q_i(t)$ – совокупная полезность труда работников всех уровней образования при оптимальном распределении работников по уровням образования.

Достоинства модели:

- прогноз численности персонала предприятий с учетом эффективности труда;
- исходные данные не требуют предварительной подготовки и преобразования.

Недостатки реализации модели:

- исследуется только одна отрасль экономики;
- наиболее точное прогнозирование только на один период;
- модель не учитывает возможные структурные сдвиги в экономике.

7. Агентная технология прогнозного моделирования кадровой потребности

Перспективным способом решения рассматриваемой задачи представляется создание технологии прогнозного моделирования с агентной архитектурой, в которой агенты представляют отдельных субъектов хозяйственной деятельности или некоторые их классы, а сцена – социально-экономическое состояние региона (рис.1).

Преимущества такого подхода заключаются в следующем: агенты имеют прообраз в реальном мире; проще получить данные для инициации их свойств, в том числе с помощью экспертных оценок (в сравнении с подходами, где компонентами модели являются полностью абстрактные сущности); удобно комбинировать макро- и микроэкономические факторы; первые определяют собственную (не зависящую от поведения агентов) динамику сцены, вторые – определяют поведение отдельных агентов и их совокупностей (коалиций); возможность масштабирования модели (от уровня отдельных экономических агентов – до отраслей или иных крупных субъектов).

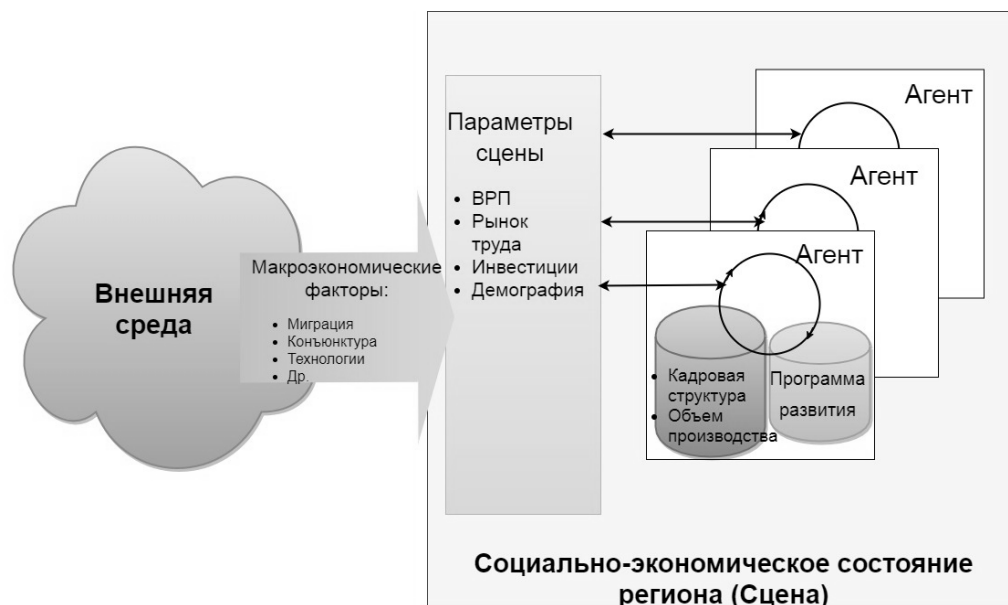


Рис.1. Концептуальная схема агентной технологии кадрового прогнозирования

Кроме того, за счет того, что агентная модель в определенной степени изоморфна реальной социально-экономической системе, упрощается формирование и анализ различных сценариев развития системы. При этом наличие развитого программного инструментария для агентного моделирования позволяет генерировать множества альтернативных траекторий системы с целью последующего экспертного анализа.

Вместе с тем, использование агентного подхода к моделированию не решает кардинально проблему неполноты и неточности исходных данных. Кроме известных способов их получения – официальной статистики, опросов и экспертных оценок, представляется перспективным использование метода аналогий: реплицирование на вновь инициируемые проекты кадровой структуры и иных параметров уже реализованных схожих проектов. Также определенные перспективы открываются и в связи с ростом объемов открытой информации в сети Интернет: ее автоматизированный анализ с применением методов data mining и искусственного интеллекта потенциально позволит извлекать неявную информацию, актуальную для построения прогнозной модели. В то же время, рост популярности массовых электронных коммуникаций создает предпосылки эффективной реализации концепции «пользователь как эксперт» для получения более достоверных экспертных оценок при формировании прогноза методом компьютерного моделирования.

Таким образом, можно обозначить следующие актуальные научно-практические задачи, подлежащие решению в контексте обозначенной проблемы:

- разработка архитектуры системы прогнозного моделирования кадровой потребности региона;
- разработка и обоснование применимости метода аналогий при инициализации модельных агентов;

- разработка методики оценки достоверности прогноза с учетом структуры модели и точности ее отдельных компонентов (степени достоверности использованных для инициализации агентов и сцены данных).

Заключение

Прогнозирование кадровой потребности региональных экономик остается актуальной задачей. Сложность и многоаспектность данной задачи обуславливает необходимость создания адекватной научной базы для ее решения в виде методов и технологий прогнозного моделирования. Обзор современных подходов к моделированию кадровой потребности демонстрирует отсутствие универсальных решений, обеспечивающих достаточную эффективность прогнозирования с учетом внешних и внутренних факторов. Наиболее существенным недостатком рассмотренных в работе моделей и подходов к прогнозированию кадровой потребности заключается в их слабом уровне изоморфности актуальной структуре экономики региона. Это принципиально снижает возможность моделировать потенциальные структурные изменения, связанные, в частности, с реализацией крупных инвестиционных проектов, а также резким (кризисным) изменением макроэкономических показателей. В большинстве рассмотренных случаев активно используется экспертный подход к идентификации параметров моделей, что также снижает возможность динамичной перестройки моделей сообразно изменениям экономической ситуации в рассматриваемых регионах.

Потенциально эффективным решением обозначенных проблем прогнозного моделирования кадровой потребности является использование агентного подхода, обеспечивающего построение динамических моделей, изоморфных изучаемой системе. Агентные модели хорошо масштабируются и позволяют учитывать при моделировании социально-экономических процессов эффекты, обусловленные деятельностью экономических агентов различного уровня – от отраслей до малых предприятий. При этом агентная архитектура обеспечивает возможность строить вложенные модели, отражающие иерархические социально-экономические структуры. Вместе с тем, для воплощения агентного подхода в контексте рассматриваемой практической проблемы необходимо решить ряд нетривиальных научно-практических задач.

Литература

1. Васильева, З. А. Проблемы моделирования кадровой потребности региональной экономики / З. А. Васильева, И. В. Филимоненко // Вестник ТГЭУ. 2012. №4 (64). – С.46-57.
2. Васильева, З.А. Разработка методики прогнозирования спроса и предложения на рынке труда и образовательных услуг экономики муниципальных образований Красноярского края / З. А. Васильева, И. В. Филимоненко, Н.В. Разнова // Спрос и предложение на рынке труда и рынке образовательных услуг в регионах России: сб. докл. Всерос. науч.-практ. интернет-конф. с междунар. участием. – Петрозаводск: Изд-во ПетрГУ, 2008. – С.54-55.
3. Гуртов, В. А. Математическая модель прогнозирования спроса и предложения на рынке труда в российских регионах / В.А. Гуртов, Е. А. Питухин //

- Обозрение прикладной и промышленной математики. Т. 11. Вып. 3. М., 2004. – С. 539.
4. Гуртов, В. А. Моделирование потребностей экономики в кадрах с профессиональным образованием / В.А. Гуртов, Е. А. Питухин, Л. М. Серова // Проблемы прогнозирования. 2007. №6. – Режим доступа: <http://cyberleninka.ru/article/n/modelirovanie-potrebnostey-ekonomiki-v-kadrah-s-professionalnym-obrazovaniem>.
 5. Егорова, Моделирование кадровой потребности предприятий и отраслей экономики /Н.Е. Егорова, Я.В. Ярошенко // Аудит и финансовый анализ. 2011. №1. – С.75-83.
 6. Козицина, А. Н. Информационная модель прогнозирования кадровой потребности региона (на примере Красноярского края) / А. Н. Козицина, И. В. Филимоненко // Вестник КрасГАУ. 2013. №7. – С.36-41.
 7. Мокроносов, А. Г. Прогнозирование обеспечения региональной экономики квалифицированными кадрами / А. Г. Мокроносов, А. А. Вершинин, Д. М. Прудников // Известия УрГЭУ. 2012. №6 (44). – Режим доступа: <http://cyberleninka.ru/article/n/prognozirovanie-obespecheniya-regionalnoy-ekonomiki-kvalifitsirovannymi-kadrami>.
 8. Мокроносов, А.Г. Прогнозирование потребности региональной экономики в подготовке кадров: монография. / А.Г. Мокроносов и др. – Екатеринбург: РГППУ, 2010. – 111с.
 9. Клевец, Н. И. Моделирование кадровых потребностей реального сектора экономики республики Крым / Н.И. Клевец, Е.А. Полищук // Теория и практика общественного развития. 2016. №9. – Режим доступа: <http://cyberleninka.ru/article/n/modelirovanie-kadrovyyh-potrebnostey-realnogo-sektora-ekonomiki-respubliki-krym>.

Сведения об авторах

Шишаев Максим Геннадьевич – д.т.н., доцент, профессор РАН, главный научный сотрудник ИИММ КНЦ РАН

e-mail: shishaev@iimm.ru

Maxim G. Shishaev – Dr.Sci. (Tech.), associate professor, professor of RAS, lead researcher

Быстров Виталий Викторович – к.т.н., старший научный сотрудник

e-mail: bystrov@iimm.ru

Vitaliy V. Bystrov – PhD (Tech. Sci.), senior researcher

Давидюк Елена Сергеевна – аспирант ИИММ КНЦ РАН,

e-mail: davelse@mail.ru

Elena S. Davidyuk – graduate student of IIMM KSC RAS

А.В. Вицентий¹, М.Г. Шишаев¹, П.А. Ломов¹, Г.Г. Гогоберидзе²

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² ФГБОУ ВО «МАГУ»

РАЗРАБОТКА ОНТОЛОГИИ ПРЕДМЕТНОЙ ОБЛАСТИ ДЛЯ ФОРМАЛЬНОГО ЛОГИЧЕСКОГО ВЫВОДА И ВИЗУАЛИЗАЦИИ В СИСТЕМАХ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ ЗАДАЧ МОРСКОГО ПРОСТРАНСТВЕННОГО ПЛАНИРОВАНИЯ*

Аннотация

Данная статья является продолжением работ коллектива по созданию подходов к решению задач разработки научно-методических основ управления морским пространственным планированием развития морехозяйственных комплексов европейской части Арктической зоны Российской Федерации на основе интеллектуальных информационных технологий (на примере морехозяйственного кластера Мурманской области). Задача, решение которой описано в этой статье, заключается в анализе и выявлении основных концептов и отношений, используемых в рамках системы пространственного планирования, и разработке их формального представления в виде онтологической модели для последующего автоматизированного выполнения аналитических операций с использованием логического вывода на примере морехозяйственного кластера Мурманской области. Предлагаемое решение основано на онтологической модели предметной области морского пространственного планирования. Модель представляет машинно-интерпретируемый набор сущностей и отношений, выявленный на примере морехозяйственного кластера Мурманской области, а также обеспечивает возможность задания набора основных концептов и отношений предметной области и осуществления формального логического вывода для решения различных задач, в том числе и задач обнаружения потенциальных конфликтов при осуществлении различных видов морской деятельности.

Ключевые слова:

онтология предметной области, морское пространственное планирование, логический вывод, управление пространственно-распределенными системами, интерфейс пользователя, системы поддержки принятия решений

A.V. Vicentiy, M.G. Shishaev, P.A. Lomov, G.G. Gogoberidze

DEVELOPMENT OF DOMAIN ONTOLOGY FOR FORMAL INFERENCE AND VISUALIZATION IN INFORMATION SUPPORT SYSTEMS FOR THE TASKS OF MARITIME SPATIAL PLANNING

Abstract

This paper is a continuation of the work to the development of scientific-methodical bases of marine spatial planning of maritime economics development in the European Arctic zone of the Russian Federation on the basis of intelligent information technologies on the example of the Murmansk region. The scientific problem, the solution of which is described in this article is to analyze and identify the main concepts and relationships used within the spatial planning system, and to develop their formal

* Исследование выполнено при частичной финансовой поддержке РФФИ и Министерства образования и науки Мурманской области в рамках научных проектов № 17-47-510298 p_a и 17-45-510097 p_a.

presentation as an ontological model for subsequent automated execution of analytical operations using logical inference on the example of the marine economic cluster of the Murmansk region. The proposed solution is based on the ontological domain model for the maritime spatial planning. The model represents a machine-interpreted set of entities and relationships, identified by the example of the marine economic cluster of the Murmansk region.

It also provides the ability to enter a set of basic concepts and relations of the subject area and implement a formal inference to solve various problems, including the problems of detecting potential conflicts in the implementation of various types of marine activities.

Keywords:

domain ontology, marine spatial planning, inference, management of spatially distributed systems, user interface, decision support systems

1. Введение

Морское пространственное планирование может определяться по-разному в зависимости от целей его использования. Одним из наиболее полных и часто используемых определений морского пространственного планирования (МПП) является определение Межправительственной океанографической комиссии ЮНЕСКО. В нем МПП определяется как общественный процесс анализа и пространственно-временного распределения деятельности человека в морских районах. В такой трактовке, морское пространственное планирование является инструментом организации рационального использования морского пространства и взаимодействия между основными заинтересованными лицами (стейкхолдерами) на основе системного и экологического подходов. Благодаря применению методов морского пространственного планирования можно оценивать кумулятивный эффект влияния различных видов морской деятельности человека на экосистемы региона и другие отрасли, а также обнаруживать и предупреждать конфликты между различными стейкхолдерами, касающиеся морехозяйственной деятельности. В этой связи, разработка научно-методических основ управления морским пространственным планированием развития морехозяйственных комплексов Российской Федерации на основе интеллектуальных информационных технологий является актуальной научной задачей.

В последние годы разработка научно-методических основ и практических инструментов в области морского пространственного планирования и комплексного управления прибрежными зонами (КУПЗ) становится особенно актуальной для российской Арктики. Так, Стратегия развития Арктической зоны Российской Федерации (АЗРФ) и обеспечения национальной безопасности на период до 2020 года, утвержденная Президентом Российской Федерации 8 февраля 2013 года [1], среди приоритетных направлений и основных мероприятий выделяет «разработку и апробацию моделей комплексного управления прибрежными зонами в арктических регионах». Планом мероприятий по реализации Стратегии развития АЗРФ и обеспечения национальной безопасности на период до 2020 года, предусматриваются разработка и апробация моделей комплексного управления прибрежными зонами в арктических регионах. Одним из главных назревающих конфликтов в АЗРФ можно назвать противоречия между освоением нефтегазовых месторождений на континентальном шельфе и промышленным рыболовством, которые ведут спор за одни и те же акватории.

Практически все намеченные недропользователями инвестиционные площадки и проекты, акваториально почти полностью совпадают с районами нерестовой миграции и нагула водных биоресурсов, особенно ценных промысловых видов гидробионтов. Усугубляется конфликт между развитием морского туристско-рекреационного бизнеса в Арктике и развитием морепромышленного производства в целом. Также, отмечается конфликтность в использовании территорий и акваторий в хозяйственных, военно-стратегических и оборонных целях [2].

Мурманская область, входящая в АЗРФ, не является исключением, поэтому вопросы морского (акваториального) пространственного планирования и эффективной организации морехозяйственной деятельности для нее также актуальны. Но, несмотря на то, что Мурманская область относится к приморским регионам, омывается Баренцевым и Белым морями, имеет относительно развитую портовую инфраструктуру, доступ к удобным международным морским путям, места базирования рыболовных, ледокольного и Северного военно-морского флотов, в регионе отсутствует четкий пространственный план развития морских акваторий и прибрежных территорий. Отсутствие системного подхода в этой области может привести к повышению межотраслевой и внутриотраслевой конфликтности и пагубно сказаться на экологии и биоразнообразии региона.

Стратегия социально-экономического развития Мурманской области до 2020 года и на период до 2025 года в качестве приоритетных направлений развития региона указывает развитие рыбопромыслового и аквакультурного кластеров, технологического кластера обеспечения шельфовой добычи в Арктике, технологического и материального обеспечения морской деятельности, туристического кластера и развитие морского круизного туризма [3]. Модель экономики Мурманской области в значительной степени ориентирована на развитие морехозяйственной деятельности и эксплуатации системообразующей роли Северного морского пути (СМП) [4]. Эти особенности обуславливают специализацию формируемых акваториальных морехозяйственных комплексов (кластеров) и значительно отличают ее от внутриконтинентальных районов Севера и циркумполярных территорий. Кроме того, переход к кластерному (узловому) развитию Мурманской области связан с очаговым освоением территории, высокой дисперсностью расселения, низкой плотностью населения и удаленностью от крупнейших промышленных центров страны. [5]

Морехозяйственный кластер Мурманской области понимается как совокупность размещенных на ограниченной территории предприятий и организаций, представляющий собой непосредственных участников кластера, и связанных с их деятельностью организаций, которая характеризуется наличием объединяющей их деятельности в одной или нескольких отраслях (видах экономической деятельности), механизма координации этой деятельности и кооперации участников кластера, синергетического эффекта, выражающегося в повышении экономической эффективности и результативности деятельности участников кластера за счет высокой степени их концентрации и кооперации. В данном контексте непосредственные участники кластера, и связанные с их деятельностью организации можно назвать совокупностью стейкхолдеров кластера. К важным характеристикам кластера относятся географическая концентрация и (или) функциональная взаимосвязанность стейкхолдеров, конкуренция и кооперация стейкхолдеров, ведение совместной деятельности,

встроенность в социально-экономическую систему. При этом, возможность использования различных видов ресурсов, сосредоточенных на одной территории потенциально ведет к появлению противоречий (конфликтов) стейкхолдеров, и оказывает негативное влияние на социально-экономическое развитие области в целом. [6]

Таким образом, в настоящее время, существует острая необходимость в разработке научно-методических основ и практических инструментов поддержки регионального управления и развития морехозяйственного кластера Мурманской области с учетом интересов и противоречий различных стейкхолдеров.

2. Постановка задачи

Большое количество субъектов экономической и иной деятельности в морских акваториях и прибрежных зонах, а также сложность затрагиваемых морехозяйственной деятельностью био-социо-экономических систем делает актуальной задачу системного представления состава и отношений разнородных объектов, так или иначе ассоциированных с рассматриваемой деятельностью на основе морского пространственного планирования и системного подхода. [7] Возникновение конфликтов, связанных с конкуренцией различных морепользователей за ресурсы и право приоритетного пространственно-временного использования территорий является естественным процессом при осуществлении морской деятельности на полиресурсных территориях, к которым относится морехозяйственный кластер Мурманской области. Одним из наиболее характерных типов конфликта в регионе является конфликт между биоресурсными и минеральноресурсными видами деятельности. Наиболее остро он проявляется на примере промышленного рыболовства и освоения шельфовых запасов нефти и газа. Почти всегда это конфликтные виды природопользования, но в зависимости от конкретных зон, возможность их совмещения колеблется от нежелательной до ограниченно допустимой. [5]

Данная статья продолжает цикл работ, направленных на создание подходов к решению задач разработки научно-методических основ управления морским пространственным планированием развития морехозяйственных комплексов европейской части Арктической зоны Российской Федерации на основе интеллектуальных информационных технологий (на примере морехозяйственного кластера Мурманской области). В рамках этого направления выделяется отдельная проблема, связанная с возможностью автоматизированного выявления потенциальных конфликтов стейкхолдеров формальными средствами для поддержки принятия решений в области МПП на примере морехозяйственного кластера Мурманской области.

Таким образом, в рамках данной работы, основная задача исследования сформулирована следующим образом: «анализ и выявление основных концептов и отношений, используемых в рамках системы пространственного планирования, и разработка их формального представления для последующего автоматизированного выполнения аналитических операций с использованием логического вывода».

3. Методы решения

3.1. Концептуальное моделирование

Решение обозначенной выше задачи возможно получить различными способами. Одним из подходов, позволяющих получить приемлемое решение подобного класса задач, является разработка концептуальной модели предметной области (КМПО) морехозяйственной деятельности. Работы, проведенные научным коллективом в этом направлении были описаны в [6]. В результате была разработана концептуальная модель предметной области, учитывающая интересы стейкхолдеров, осуществляющих морехозяйственную и иные виды деятельности в морехозяйственном кластере Мурманской области. Модель представляет собой формализованное описание качественного состава субъектов и объектов морехозяйственной деятельности региона и отношений между ними (рис. 1.). Концептуальная модель предметной области может быть использована для прогнозирования и моделирования процессов управления социально-экономического развития био- социо- экономических систем морехозяйственного кластера Мурманской области с целью синтеза приемлемых для всех заинтересованных сторон пространственных планов морепользования и обеспечения оперативной и адекватной реакции системы управления на изменения составляющих объекта управления и внешней среды.



Рисунок 1. Схема концептуальной модели предметной области

С другой стороны, КМПО необходима для обеспечения технологической основы решения задач интеллектуальной информационной поддержки по управлению морским (акваториальным) пространственным планированием развития морехозяйственных комплексов. Для этих целей, в процессе разработки модели было выделено 39 видов различных отношений, образующих иерархию и определены их свойства, потенциально значимые в контексте процедур логического вывода. Таким образом, разработанная модель представляет собой логическую систему, обеспечивающую возможность автоматизированного вывода заключений на основе имеющихся фактов, а, кроме того, может быть использована для анализа полноты и непротиворечивости коллективных знаний, представленных экспертами (рис. 2).

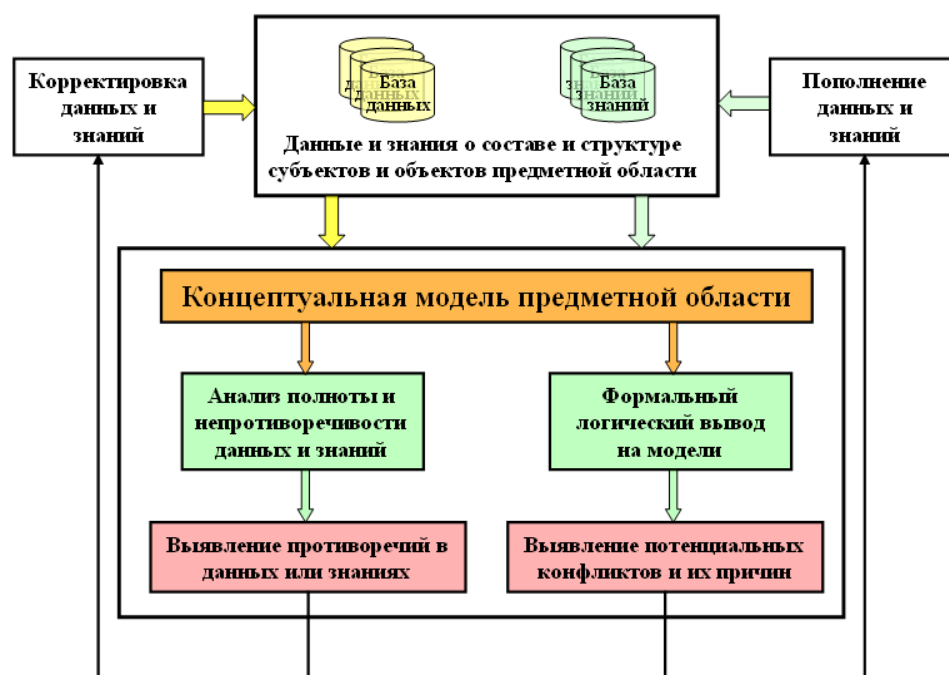


Рисунок 2. Использование концептуальной модели предметной области

3.2. Онтологическое моделирование

В работах [5, 6] была заложена основа для формализации процедур выявления конфликтов стейкхолдеров применительно к морскому пространственному планированию с помощью концептуальной модели предметной области морехозяйственной деятельности. Благодаря предложенным подходам и полученным результатам были определены состав и структура взаимодействия основных концептов. Однако, для более эффективного использования в информационных системах (ГИС, картографические сервисы, и др.) больше подходит онтология, описанная в нотации OWL (Web Ontology Language). Изначально, язык OWL разрабатывался для описания классов, присущих веб-документам и приложениям, и отношений между ними. В основе языка OWL лежит представление действительности в модели данных «объект - свойство». Современная версия OWL широко используется не только для описания веб-страниц, но и любых объектов действительности. Каждому элементу описания в этом языке (в том числе свойствам, связывающим объекты) ставится в соответствие унифицированный идентификатор ресурса (Uniform Resource Identifier – URI). Далее, в работе рассмотрено решение задачи анализа и выявления основных концептов и отношений, используемых в рамках системы пространственного планирования, и разработки их формального представления для последующего автоматизированного выполнения аналитических операций с использованием логического вывода на основе применения онтологии.

На формальном уровне онтология - это система, состоящая из набора понятий и набора утверждений об этих понятиях, на основе которых можно описывать классы, отношения, функции и индивиды. В контексте решаемой задачи под онтологией понимается система понятий предметной области, которая

представляется как набор сущностей, соединенных различными отношениями. В данной работе онтология используется для формальной спецификации понятий и отношений, которые характеризуют определенную область знаний. Важным преимуществом онтологий в качестве способа представления знаний является их формальная структура, которая упрощает их последующую компьютерную обработку. Кроме того, создание явных допущений в предметной области, лежащих в основе реализации, дает возможность легко изменить эти допущения при изменении или пополнении знаний о предметной области.

4. Результаты

Создание онтологии это итеративный процесс, требующий частого возврата к предыдущим уровням разработки и внесения изменений. В этой статье представлен текущий этап разработки, так сказать, моментальный снимок, онтологии предметной области (ОПО) морского пространственного планирования на примере Мурманской области. Для создания качественной ОПО необходимо привлечение различных экспертов предметной области. Однако, на первом этапе значительная часть информации может быть получена и из литературных источников [8-11], в том числе, с использованием методов автоматизированного анализа текстов с целью выявления их смыслов, наборов наиболее значимых концептов (классов онтологии) и связей между ними [12].

На основе подробного изучения доступных данных (отчетов по результатам научно-исследовательской деятельности, монографий, научных статей, обзоров и других открытых источников), было, в частности, установлено, что наиболее актуальны для морехозяйственного кластера Мурманской области конфликты между освоением месторождений углеводородов и добычей гидробионтов на континентальном шельфе. Причем, часто эти конфликты заданы не в явной форме, т.е. как противоречия между недропользователями и представителями промышленного рыболовства, а опосредованно, например, через претензии на одни и те же территории, которые одновременно являются как перспективными с точки зрения добычи нефти и газа, так и важными для нереста и нагула промысловых рыб.

Также в процессе анализа были выделены основные активности, характерные для морехозяйственного кластера Мурманской области. В частности, к наиболее важным из них, составляющих набор основных концептов, используемых в рамках системы пространственного планирования, можно отнести:

- добыча гидробионтов (включая ценные виды рыб) на большом удалении от берега;
- судоходство;
- рыбоводство;
- переработка гидробионтов (включая растительное сырье и ценные виды рыб) ;
- аквакультура и марикультура;
- прибрежная добыча гидробионтов (включая ценные виды рыб);
- нефтедобыча
- газодобыча,
- геологоразведка,

- транспортировка углеводородов;
- переработка углеводородов;
- ледовое судоходство,
- проводка судов;
- деятельность портов;
- поддержание инфраструктуры для судоходства;
- дноуглубительные работы;
- создание искусственных островов;
- созданию инфраструктуры;
- охрана государственных границ и сопутствующие мероприятия (учения, стрельбы, разведка, патрулирование, закрытия морских и прибрежных (акваториальных) зон, др.;
- туризм;
- охрана окружающей среды;
- прокладка коммуникаций;
- прибрежная энергетика;
- и др.

Подобным образом были выявлены и другие основные концепты и отношения, используемых в рамках системы пространственного планирования, и формально представлены в виде онтологической модели для последующего автоматизированного выполнения аналитических операций с использованием логического вывода на примере морехозяйственного кластера Мурманской области. Визуальное представление фрагмента разработанной онтологии представлено на рисунке 3.

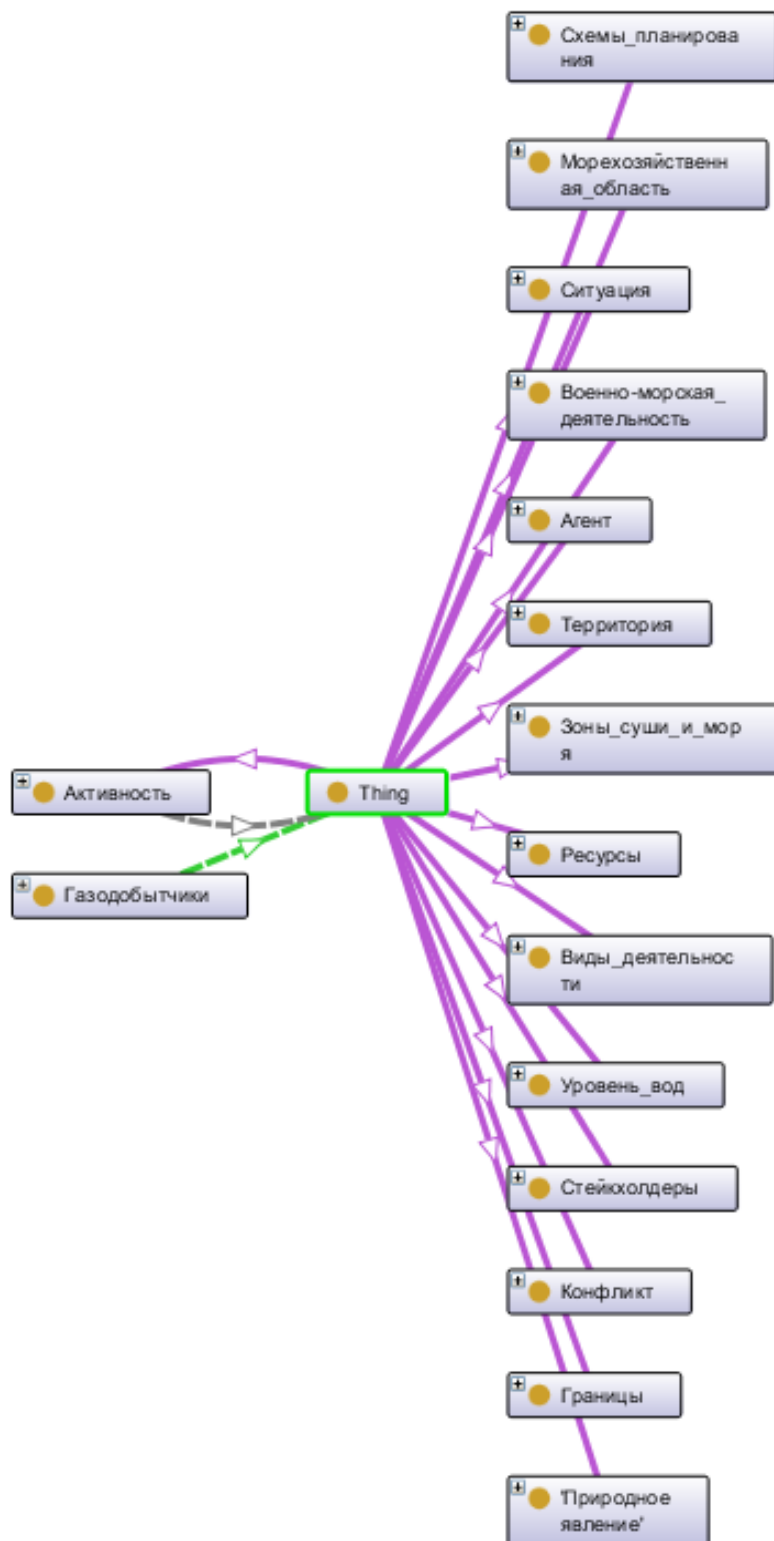


Рисунок 3. Визуальное представление фрагмента онтологии предметной области

5. Пример практического использования онтологии предметной области (use case)

Для реализации автоматизированного логического вывода в онтологии были введены специальные классы:

- Situation() – ситуация;
- ConflictSituation() – конфликтная ситуация;
- Territory() – территория;
- Agent() – агент (некоторая активная сущность);
- NaturalPhenomenon() - природное явление;
- Activity() – активность (некоторое действие).

Также были заданы следующие отношения:

- hasActivity(X,Y) – бинарное отношение «ситуация имеет активность»;
- hasTerritory(X,Y) - бинарное отношение «ситуация имеет территорию»;
- hasNegativeEffectOn(Y,Y) - бинарное отношение «имеется негативный эффект при данном сочетании активности и территории»;
- hasParticipant(X,Y) - бинарное отношение «ситуация имеет агента».

Помимо приведенных классов и отношений в онтологии были заданы дополнительные аксиомы, позволяющие применять логический вывод для решения аналитических задач. В качестве примера рассмотрим аксиому, используемую для обнаружения потенциального конфликта. Она позволяет относить к классу конфликтной ситуации некоторые экземпляры класса ситуация, имеющих некоторый набор связей (отношений) с экземплярами других классов. Данная аксиома на языке логики предикатов первого порядка будет иметь следующий вид:

Situation(X0) AND Activity(Y) AND Territory(V) AND hasTerritory(X0,V) AND hasActivity(X0,Y) AND hasNegativeEffectOn(Y,V) AND

Situation(X1) AND NaturalPhenomenon(N) AND hasActivity(X1,N) AND hasTerritory(X1,V) -> ConflictSituation(X0).

В разработанной онтологии данная аксиома представляется в виде выражения дескрипционной логики, которое в манчестерском синтаксисе языка OWL, будет иметь следующий вид:

ConflictSituation EquivalentTo

Situation and hasActivity some

(Activity and hasNegativeEffectOn some

(Territory and inverse hasTerritory some Situation))

Содержательный смысл данной аксиомы состоит в следующем: если активность в рамках ситуации негативно влияет на территорию, на которой происходит природное явление (иная ситуация), то эта ситуация является конфликтной (потенциально конфликтной) ситуацией.

Таким образом, в онтологии в виде экземпляров класса «Активность» могут быть заданы различные виды деятельности, а факты ее ведения какой-либо организацией на некоторой территории – соответствующим экземпляром ситуации, включающей экземпляр территории и агента (сущности ведущей деятельность). Далее, определяем отношения hasNegativeEffectOn между экземплярами видов активности и территориями, исходя из данных природного

мониторинга. Теперь можно также в виде ситуаций задавать природные явления, происходящие на некоторой территории.

Проведение классификации машиной вывода в этом случае позволяет автоматически выявлять конфликтные ситуации, на основе рассмотренной выше аксиомы, заданной в рамках определения класса «Конфликтная ситуация».

Визуальное представление конфликтной ситуации, полученной с помощью формального логического вывода на разработанной онтологии предметной области, представлено на рисунке 4.

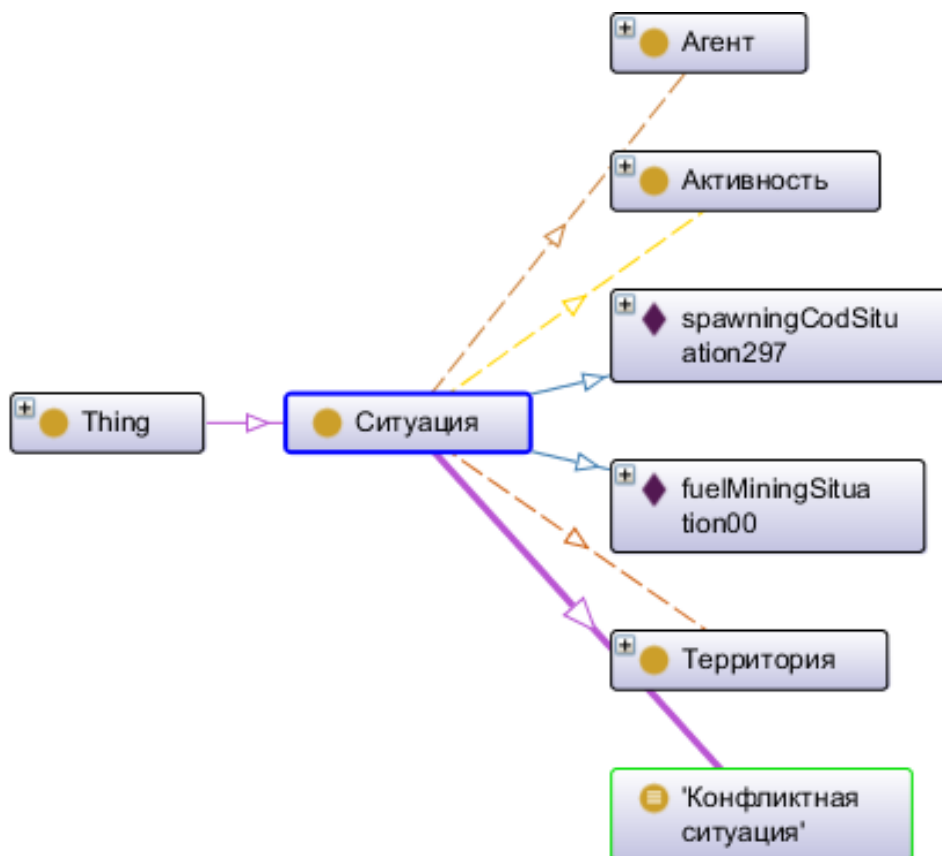


Рисунок 4. Визуальное представление конфликтной ситуации.

6. Заключение и обсуждение полученных результатов

В данной работе описано решение задачи анализа и выявления основных концептов и отношений, используемых в рамках системы пространственного планирования, и разработки их формального представления для последующего автоматизированного выполнения аналитических операций с использованием логического вывода.

В основе предложенного решения поставленной задачи лежит онтологическая модель предметной области морского пространственного планирования, представляющая машинно-интерпретируемый набор сущностей и

отношений, выявленный на примере морехозяйственного кластера Мурманской области.

Полученные результаты являются составной частью разрабатываемых научно-методических основ управления морским (акваториальным) пространственным планированием развития морехозяйственных комплексов европейской части Арктической зоны Российской Федерации (АЗРФ) на основе интеллектуальных информационных технологий.

Таким образом, можно говорить, что поставленная в работе задача была решена авторским коллективом, а результаты обоснованы и опробованы на модельном примере выявления конфликтной ситуации, связанной с совмещением деятельности разных стейкхолдеров на одной территории. При этом, нельзя с полной уверенностью утверждать, что предложенная авторами онтологическая модель предметной области находится в своем окончательном виде и ее возможности на данный момент покрывают все потенциальные потребности автоматизированного выполнения аналитических операций в рассматриваемой области. Однако, на данном этапе работ, полученный результат следует признать удовлетворительным и практически обоснованным.

Предложенный в работе подход обеспечивает возможность задания набора основных концептов и отношений предметной области морского пространственного планирования (на примере морехозяйственного кластера Мурманской области) и осуществления формального логического вывода для решения различных практических задач МПП и КУПЗ. В том числе и задач обнаружения потенциальных конфликтов при осуществлении различных видов морской деятельности различными стейкхолдерами.

Онтологическая модель предметной области морского пространственного планирования была разработана на примере Мурманской области и включает в себя классы и отношения, характерные, прежде всего, для данного региона. Однако, возможности модели могут быть существенно расширены как для использования на примере других регионов АЗРФ, так и на АЗРФ в целом, за счет пополнения знаний о предметной области и формального их описания в нотации OWL.

7. Направления дальнейшей работы

Гибкость используемого для решения поставленной проблемы подхода и предложенной онтологической модели предметной области определяют высокую вариативность дальнейших работ в этом направлении. Среди наиболее перспективных хотелось бы отметить следующие.

Дополнение и развитие предложенной онтологии за счет добавления концептов, отношений и аксиом в данной предметной области.

Разработка и интеграция модуля пополнения ОПО на основе автоматических и автоматизированных интеллектуальных методов извлечения и формализации знаний.

Расширение области применения предложенного подхода за счет пополнения знаний, заложенных в ОПО и апробация возможностей модели для решения подобных задач в других регионах арктической зоны Российской Федерации.

Особо следует отметить направление разработки интеллектуальных средств поддержки принятия решений регионального управления за счет

дополнения ОПО технологией геовизуализации, которая позволит представлять результаты анализа ситуаций лицу, принимающему решения (ЛПР) в более наглядной форме, с привязкой к конкретной территории на основе актуальных картографических данных. Некоторый задел в этом направлении уже был сделан авторами в работах [8, 10, 13 -15]. Такой интегрированный подход к управлению территориями, объединяющий возможности когнитивной динамической визуализации пространственных данных и формального представления знаний позволит существенно повысить оперативность и эффективность принятия решений в мультипредметных информационных системах поддержки регионального управления в условиях АЗРФ, а также позволит снизить когнитивную нагрузку на ЛПР в процессе работы.

Литература

1. Стратегия развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года URL: <http://docs.cntd.ru/document/499002465>
2. Коновалов А.М. Морское пространственное планирование в системе стратегического планирования развития Арктической зоны Российской Федерации. Российский Север: модернизация и развитие. 2015. С.32-38
3. Стратегия социально-экономического развития Мурманской области до 2020 года и на период до 2025 года, утвержденной постановлением Правительства Мурманской области от 25 декабря 2013 г. N 768-ПП/20
4. Карлин Л. Н., В. Н. Воробьев, В. М. Абрамов, Г. Г. Гогоберидзе Научное обеспечение стратегического планирования развития Северного морского пути как транспортного коридора с учетом изменений климата Арктики // Транспорт Российской Федерации. Журнал о науке, практике, экономике. 2014. № 1. С. 16-21
5. Ершова А.А., Вицентий А.В., Гогоберидзе Г.Г., Шишаев М.Г., Ломов П.А. Морское пространственное планирование: возможности для приморских территорий и прилегающих акваторий Мурманской области. // Национальные интересы: приоритеты и безопасность. - 2018. - Т. 14. - № 2 (359). - С. 269-287.
6. Вицентий, А.В. Шишаев, М.Г., Ершова А.А., Гогоберидзе Г.Г. Концептуальная модель морехозяйственной деятельности в регионе как основа систем информационной поддержки морского пространственного планирования //Труды Кольского научного центра РАН. – 2017. – Вып. 8. Информационные технологии. С. 77- 88.
7. Гогоберидзе Г.Г., Ершова А.А., Румянцева Е.А., Шишаев М.Г., Вицентий А.В., Ломов П.А. Возможности применения инструментария морского пространственного планирования для Мурманской области. // Материалы XXVII Международной береговой конференции. Ответственный редактор Е.А. Румянцева. 2018. С. 378-381.
8. Vicentiy A.V., Shishaev M.G., Oleynik A.G. Dynamic Cognitive Geovisualization for Information Support of Decision-Making in the Regional System of Radiological Monitoring, Control and Forecasting. // Advances in Intelligent Systems and Computing, vol. 466, pp. 483-495. Springer, Cham.
9. Vicentiy A.V., Shishaev M.G., Vicentiy I.V. The Development of Dynamic Cognitive Interfaces for Multisubject Information Systems (on the Example of

- Geosocial Service). // *Advances in Intelligent Systems and Computing*, vol. 575, pp. 449-459. Springer, Cham.
10. Вицентий А. В., Шишаев М.Г., Порядин Т.А. К вопросу о разработке когнитивных интерфейсов средств информационной поддержки управления развитием пространственно-распределенных систем // *Наука - производству. Материалы международной научно-практической конференции. Мурманский государственный технический университет. 2015. С. 109-113.*
 11. Vicentiy I.V., Eliseev S.M., Vicentiy A.V. Applying of the Classifications Trees Method in Forecasting of Risk Groups of Intolerant Behavior. // *Advances in Intelligent Systems and Computing*, vol. 662 pp. 350-359. Springer, Cham
 12. Vicentiy A.V., Dikovitsky V. V., Shishaev M.G. The Semantic Models of Arctic Zone Legal Acts Visualization for Express Content Analysis // *Advances in Intelligent Systems and Computing*, vol. 763, pp. 216-228. Springer, Cham.
 13. Вицентий А.В. Визуализация пространственных данных как подход к построению когнитивных интерфейсов мультимедийных информационных систем поддержки регионального управления // *Интернет-журнал «Науковедение»* - 2017. Т. 9. - №5 (42).
 14. Вицентий А. В. Применение дистанционного зондирования земли и космических технологий для развития арктических и субарктических территорий российской федерации // *Труды Кольского научного центра РАН. 2013. № 5 (18). С. 40-45.*
 15. Вицентий А.В. Разработка технической платформы средств динамического картографирования и визуального анализа на примере системы информационной поддержки мониторинга радиологической обстановки // *Фундаментальные проблемы системной безопасности. - 2014. С. 324-329.*

Сведения об авторах

Вицентий Александр Владимирович – к.т.н., старший научный сотрудник
e-mail: alx_2003@mail.ru

Vicentiy Alexander – PhD (Tech. Sci.), senior researcher

Шишаев Максим Геннадьевич – д.т.н, доцент, профессор РАН, главный научный сотрудник ИИММ КНЦ РАН
e-mail: shishaev@iimm.ru

Maxim G. Shishaev – Dr.Sci. (Tech.), associate professor, professor of RAS, lead researcher

Ломов Павел Андреевич – к.т.н., старший научный сотрудник, доцент
e-mail: lomov@iimm.ru

Lomov Pavel – PhD (Tech. Sci.), senior researcher, associate professor

Гогоберидзе Георгий Гививич – д.э.н., доцент
e-mail: gogoberidze.gg@gmail.com

Gogoberidze Georgii – Dr.of Sci., associate professor

А.В. Вицентий¹, М.Г. Шишаев¹, Г.Г. Гогоберидзе²

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² ФГБОУ ВО «МАГУ»

РАЗРАБОТКА ТЕХНОЛОГИИ ГЕОВИЗУАЛИЗАЦИИ ДЛЯ СИСТЕМ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ ЗАДАЧ ПРОСТРАНСТВЕННОГО ПЛАНИРОВАНИЯ С УЧЕТОМ ИНТЕРЕСОВ И ПРОТИВОРЕЧИЙ РАЗЛИЧНЫХ ГРУПП ПРИРОДОПОЛЬЗОВАТЕЛЕЙ НА ОСНОВЕ КОНЦЕПТУАЛЬНОЙ МОДЕЛИ КОНФЛИКТА*

Аннотация

В этой статье рассматривается разработка информационной технологии геовизуализации для систем поддержки принятия решений в области регионального управления на примере морехозяйственного кластера Мурманской области. Приводится постановка задачи и основные требования к разрабатываемой технологии. Кратко описывается концептуальная модель конфликтной ситуации в контексте морского пространственного планирования. Раскрывается структура, основные элементы и алгоритм работы технологии.

Ключевые слова:

геовизуализация, конфликтная ситуация, информационная система поддержки регионального управления, морское пространственное планирование, интерфейс пользователя, системы поддержки принятия решений

A.V. Vicentiy, M.G. Shishaev, G.G. Gogoberidze

DEVELOPMENT OF GEOVISUALIZATION TECHNOLOGY FOR INFORMATION SUPPORT SYSTEMS FOR SPATIAL PLANNING TASKS TAKING INTO ACCOUNT INTERESTS AND CONTRADICTIONS OF VARIOUS GROUPS OF NATURE USERS BASED ON A CONCEPTUAL MODEL OF CONFLICT

Abstract

This paper discusses the development of information geovisualization technology for decision support systems in the field of regional management on the example of the marine cluster of the Murmansk region. The statement of the problem and the basic requirements for the technology being developed are given. A conceptual model of the conflict situation in the context of marine spatial planning is briefly described. The structure, main elements and algorithm of the technology work are revealed.

Keywords:

geovisualization, conflict situation, regional government support information system, marine spatial planning, user interface, decision support systems

1. Введение

Разработка информационной технологии геовизуализации (ИТГ) для систем информационной поддержки принятия решений в области регионального управления, в том числе для решения задач морского пространственного планирования, с учетом интересов и противоречий различных групп

* Исследование выполнено при частичной финансовой поддержке РФФИ и Министерства образования и науки Мурманской области в рамках научных проектов № 17-47-510298 p_a и 17-45-510097 p_a.

природопользователей проводится в рамках работ по созданию научно-методических основ управления морским (акваториальным) пространственным планированием развития морехозяйственных комплексов европейской части Арктической зоны Российской Федерации (АЗРФ) на основе интеллектуальных информационных технологий на примере морехозяйственного кластера Мурманской области. Результаты работ в этом направлении также используются при разработке метода динамической визуализации пространственных данных для построения когнитивных интерфейсов мультипредметных информационных систем поддержки регионального управления в условиях арктической зоны Российской Федерации (АЗРФ).

Значительный научный задел в области разработки методов и средств когнитивной геовизуализации был сделан авторами в работах [1-5]. Однако информационная технология, рассматриваемая в данной статье, имеет ряд существенных отличий от полученных ранее результатов. Значительная часть этих отличий обусловлена особенностями предметной области, рассмотренными, в частности, в [6-9] и новыми подходами, используемыми для сбора и обработки данных [10, 11]. Интегрированный подход к формированию структуры и наполнению базы знаний, использованию логического вывода и визуализации результата, предлагаемый в данной работе, позволяет объединить возможности когнитивной визуализации пространственных данных и формального представления знаний для решения задач регионального управления в условиях АЗРФ и снизить когнитивную нагрузку на оператора – лицо, принимающее решения (ЛПР).

Социально-экономическое развитие АЗРФ и Мурманской области регулируется различными нормативно-правовыми актами, среди которых можно особо отметить Стратегию развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года и Стратегию социально-экономического развития Мурманской области до 2020 года и на период до 2025 года. В частности, в этих документах говорится о том, что одним из приоритетных направлений развития АЗРФ является разработка и апробация моделей комплексного управления прибрежными зонами в арктических регионах. А в качестве приоритетных направлений развития Мурманской области указаны развитие рыбопромыслового и аквакультурного кластеров, технологического кластера обеспечения шельфовой добычи в Арктике, технологического и материального обеспечения морской деятельности, туристического кластера и развитие морского круизного туризма [12,13].

При этом, наибольшее число потенциальных конфликтов связано именно с разногласиями между представителями рыбопромыслового кластера и кластера обеспечения шельфовой добычи в Арктике. Для учета интересов и противоречий всех заинтересованных лиц (стейкхолдеров), осуществляющих природопользование в регионе было предложено использовать понятие морехозяйственного кластера Мурманской области. Под морехозяйственным кластером Мурманской области понимается совокупность размещенных на ограниченной территории предприятий и организаций, представляющий собой непосредственных участников кластера, и связанных с их деятельностью организаций, которая характеризуется наличием объединяющей их деятельности в одной или нескольких отраслях (видах экономической деятельности), механизма координации этой деятельности и кооперации участников кластера,

синергетического эффекта, выражающегося в повышении экономической эффективности и результативности деятельности участников кластера за счет высокой степени их концентрации и кооперации. В данном контексте непосредственных участников кластера, и связанные с их деятельностью организации можно назвать совокупностью стейкхолдеров кластера. К важным характеристикам кластера относятся географическая концентрация и (или) функциональная взаимосвязанность стейкхолдеров, конкуренция и кооперация стейкхолдеров, ведение совместной деятельности, встроенность в социально-экономическую систему. При этом, возможность использования различных видов ресурсов, сосредоточенных на одной территории потенциально ведет к появлению противоречий (конфликтов) стейкхолдеров, и оказывает негативное влияние на социально-экономическое развитие области в целом. [10].

Регулирование деятельности всех участников и пространственное планирование морехозяйственного кластера с целью снижения конфликтности взаимодействия является нетривиальной, плохо формализуемой задачей, для решения которой необходимо разрабатывать специальные методы, технологии и программные средства. Таким образом, разработка технологии геовизуализации для систем информационной поддержки задач принятия решений в области регионального управления и морского пространственного планирования с учетом интересов и противоречий различных групп природопользователей является актуальной на сегодняшний день.

2. Постановка задачи

Выявление, описание, предупреждение и разрешение различных конфликтов, возникающих между стейкхолдерами, является первостепенной задачей повышения эффективности регионального управления и обеспечения устойчивого социально-экономического развития территорий. В этой связи, для повышения эффективности пространственного планирования морехозяйственного кластера Мурманской области, была поставлена задача разработки технологии геовизуализации для систем информационной поддержки задач морского пространственного планирования с учетом интересов и противоречий различных групп природопользователей.

3. Описание полученных результатов

3.1. Формальное описание конфликтной ситуации

Необходимо отметить, что различия в интересах стейкхолдеров и возникающие в связи с этим противоречия, часто приводят к конфликтам, описать которые без формального аппарата довольно сложно. В связи с этим, важным этапом при работе с конфликтными ситуациями, необходимым для их всестороннего понимания, является формальное описание конфликта. А большое количество субъектов социальной, экономической, политической и иной деятельности в регионе управления, и сложность затрагиваемых био-социально-экономических систем делает актуальной задачу системного представления состава и конфликтного взаимодействия разнородных субъектов, так или иначе ассоциированных с рассматриваемой деятельностью.

В контексте данной работы определим конфликт как взаимодействие двух (или более) стейкхолдеров, имеющих различные (несовместимые) цели и (или)

способы их достижения. Таких стейкхолдеров можно называть субъектами конфликта или игроками или участниками конфликта. А сам конфликт можно называть также конфликтной ситуацией.

Тогда, концептуальная модель конфликта (КМК), являющаяся ключевым элементом системы поддержки принятия решений (СППР) для регионального управления и описывающая конфликтные ситуации, возникающие между стейкхолдерами в процессе осуществления их деятельности, может быть представлена в терминах теоретико-множественных отношений следующим образом: $CS = \langle SC, OC, GSO, R \rangle$, где:

CS – концептуальная модель конфликтной ситуации, описывающая все множество конфликтных ситуаций, задаваемое кортежем $\langle SC, OC, GSO, R \rangle$;

SC – множество субъектов конфликта (стейкхолдеров), осуществляющих деятельность на территории регионального управления. На самом деле, элементами множества SC могут быть не только непосредственные субъекты, осуществляющие деятельность в регионе, но и прочие субъекты био-социо-экономической системы, которые могут иметь заинтересованность в различных исходах конфликта или на которых может существенно повлиять тот или иной исход конфликта. Такие субъекты называются косвенными участниками конфликта (например, муниципальные, региональные и федеральные органы власти, экологические организации, местное население и др.);

OC – множество объектов конфликта. В это множество входят те сущности, которые непосредственно являются объектом (предметом) конфликта. Для решения практических задач в области поддержки регионального управления необходимо специфицировать основные характеристики объекта конфликта. Таким образом, множество объектов конфликта может быть представлено в виде:

$OC = \langle S, KA, Rs, RI \rangle$, где:

S – множество пространственных областей (в географическом смысле) или территорий (акваторий), связанных с объектом конфликта;

KA – множество видов деятельности осуществляемых (или планируемых к осуществлению) стейкхолдерами из множества SC , которые связаны с объектом конфликта;

Rs – множество ресурсов био-социо-экономической системы, используемых или необходимых для осуществления некоторой деятельности стейкхолдерами из множества SC , которые связаны с объектом конфликта;

RI – множество связей на множествах $\langle S, KA, Rs \rangle$.

GSO – множество целей субъектов конфликта (стейкхолдеров), которые они пытаются реализовать в рамках осуществляемой деятельности.

R – множество отношений, заданных на множествах $\langle SC, OC, GSO \rangle$.

Кроме того, для целей поддержки регионального управления, на основе анализа и обобщения информации по основным видам деятельности наиболее важных стейкхолдеров региона (на примере Мурманской области) с использованием КМК были выявлены потенциальные конфликты и определены некоторые наиболее значимые их виды [12, 13]. Все возможные конфликты в регионе управления были условно разделены на три вида:

1. конфликты, связанные с осуществлением различных (исключающих друг друга или несовместимых полностью или частично) видов деятельности стейкхолдеров. Например, «охрана природных территорий – добыча

полезных ископаемых», «охрана морских государственных границ – водный туризм», «питьевой водозабор – сброс (утечка) сточных (технических) вод» и др.

2. конфликты, связанные с использованием определенных ресурсов территории. При этом, некоторая территория также может рассматриваться как особый вид ресурса. Здесь возможны как «внутрисекторальные», так и «межсекторальные» конфликты. Например, конкуренция различных добывающих компаний за право освоения одного месторождения, конкуренция рыболовецких компаний за квоту или территории вылова, и т.д.
3. конфликты, связанные с использованием или осуществлением некоторой деятельности на определенной территории. В таких конфликтах объектом является некоторое «ограниченное пространство» или его часть. Другими словами, можно сказать, что в этом виде конфликтов объектом является «право пользования» некоторой территорией. Примерами таких конфликтов могут быть закрытие территории для военных целей или для обеспечения безопасности, организация особо охраняемых природных территорий и т.п.

Можно отметить, что все типы конфликтов имеют территориальную и/или временную привязку. Поэтому, для информационной поддержки решения задач пространственного планирования с учетом интересов и противоречий различных групп природопользователей на основе концептуальной модели конфликта эффективным инструментом будут являться технологии геовизуализации.

3.2. Описание технологии геовизуализации

Далее приводится краткое описание состава и структуры разработанной технологии геовизуализации для систем информационной поддержки задач морского пространственного планирования с учетом интересов и противоречий различных групп природопользователей (рис. 1).



Рисунок 5. Технология геовизуализации для систем информационной поддержки задач морского пространственного планирования с учетом интересов и противоречий различных групп природопользователей

Обобщенный алгоритм работы технологии выглядит следующим образом:

1. Пользователь подает на вход описание ситуации;
2. Осуществляется формализация описания ситуации;

3. Осуществляется обработка ситуации (в соответствии с имеющейся концептуальной моделью предметной области);
4. Осуществляется логический вывод (наличие/отсутствие конфликта);
5. Осуществляется визуализация ситуации, результатов логического вывода, и рекомендаций (при наличии) по разрешению конфликта;
6. Пользователь проводит анализ полученных данных;
7. Пользователь уточняет описание ситуации либо заканчивает работу с ней.

Дадим некоторые пояснения по работе пользователя с технологией. На первом шаге пользователь должен предоставить описание анализируемой ситуации. Чем точнее и подробнее будет это описание, тем эффективнее сможет работать с ним система. На втором шаге, проводится формализация описанной ситуации в терминах предметной области. Для этих целей используется концептуальная модель предметной области. На третьем и четвертом шагах осуществляется обработка (анализ) ситуации на наличие или отсутствие потенциального конфликта. Для этого, кроме модели предметной области, используется также онтология и набор аксиом, позволяющие обнаруживать «неявные» конфликты. На пятом шаге, с привлечением внешних источников данных и средств визуализации (географические информационные системы, картографические сервисы, базы космических снимков, карты и атласы, средства визуализации онтологий и др.) осуществляется визуализация ситуации, результатов логического вывода, и рекомендаций (при наличии) по разрешению конфликта. Так как все ситуации так или иначе привязаны к территории морехозяйственного кластера, можно осуществить их визуализацию на какой-либо картографической подложке, то есть, осуществить геовизуализацию анализируемой пользователем ситуации. На шестом и седьмом шагах пользователь может на основе визуального анализа ситуации либо уточнить ее описание, либо закончить анализ ситуации.

4. Заключение и будущие работы

Таким образом в процессе работы была предложена информационная технология геовизуализации для систем информационной поддержки задач морского пространственного планирования с учетом интересов и противоречий различных групп природопользователей. Определены структура и основные составляющие технологии, а также входные и выходные данные. Результатом работы технологии геовизуализации является геоизображение, отображающее основные концепты конфликтной ситуации с привязкой к геооснове (картографическая подложка). Кроме того, при наличии дополнительной атрибутивной информации в базах данных, возможен ее вывод при анализе полученного геоизображения конечным пользователем. Такой подход к визуализации данных в системах информационной поддержки задач морского пространственного планирования позволяет снизить когнитивную нагрузку на ЛПП, так как информация по анализируемой ситуации представляется конечному пользователю в комплексном, интегрированном виде (формальное описание (активированные концепты и связи предметной области), логический вывод и его

объяснение (последовательность активированных аксиом), визуализация формального представления ситуации (визуальное представление фрагмента онтологии, описывающего анализируемую ситуацию), представление конфликтной ситуации на карте (в виде геоизображения и атрибутивных данных), вывод приемлемых решений и/или рекомендаций (при наличии такой возможности).

В будущих работах планируется развитие технологии геовизуализации пространственных данных для построения на их основе когнитивных интерфейсов мультипредметных информационных систем поддержки регионального управления в условиях арктической зоны Российской Федерации. В частности, планируются работы в направлении повышения когнитивности получаемого геоизображения, разработки и внедрение методов и инструментальных средств интерактивного визуального анализа геоданных, повышение эффективности восприятия информации за счет адаптации получаемых визуальных образов к когнитивным особенностям ЛПР и характеристикам устройств отображения.

Литература

1. Вицентий А.В. Разработка технической платформы средств динамического картографирования и визуального анализа на примере системы информационной поддержки мониторинга радиологической обстановки // *Фундаментальные проблемы системной безопасности*. - 2014. С. 324-329.
2. Вицентий А. В., Шишаев М.Г., Порядин Т.А. К вопросу о разработке когнитивных интерфейсов средств информационной поддержки управления развитием пространственно-распределенных систем // *Наука - производству*. Материалы международной научно-практической конференции. Мурманский государственный технический университет. 2015. С. 109-113.
3. Vicentiy A.V., Shishaev M.G., Oleynik A.G. Dynamic Cognitive Geovisualization for Information Support of Decision-Making in the Regional System of Radiological Monitoring, Control and Forecasting. // *Advances in Intelligent Systems and Computing*, vol. 466, pp. 483-495. Springer, Cham.
4. Vicentiy A.V., Shishaev M.G., Vicentiy I.V. The Development of Dynamic Cognitive Interfaces for Multisubject Information Systems (on the Example of Geosocial Service). // *Advances in Intelligent Systems and Computing*, vol. 575, pp. 449-459. Springer, Cham.
5. Вицентий А.В. Визуализация пространственных данных как подход к построению когнитивных интерфейсов мультипредметных информационных систем поддержки регионального управления // *Интернет-журнал «Науковедение»* - 2017. Т. 9. - №5 (42).
6. Ершова А.А., Вицентий А.В., Гогоберидзе Г.Г., Шишаев М.Г., Ломов П.А. Морское пространственное планирование: возможности для приморских территорий и прилегающих акваторий Мурманской области. // *Национальные интересы: приоритеты и безопасность*. - 2018. - Т. 14. - № 2 (359). - С. 269-287.
7. Гогоберидзе Г.Г., Ершова А.А., Румянцева Е.А., Шишаев М.Г., Вицентий А.В., Ломов П.А. Возможности применения инструментария морского пространственного планирования для Мурманской области. // *Материалы*

- XXVII Международной береговой конференции. Ответственный редактор Е.А. Румянцева. 2018. С. 378-381.
8. Коновалов А.М. Морское пространственное планирование в системе стратегического планирования развития Арктической зоны Российской Федерации. Российский Север: модернизация и развитие. 2015. С.32-38
 9. Вицентий А. В. Применение дистанционного зондирования земли и космических технологий для развития арктических и субарктических территорий российской федерации // Труды Кольского научного центра РАН. 2013. № 5 (18). С. 40-45.
 10. Вицентий, А.В. Шишаев, М.Г., Ершова А.А., Гогоберидзе Г.Г. Концептуальная модель морехозяйственной деятельности в регионе как основа систем информационной поддержки морского пространственного планирования //Труды Кольского научного центра РАН. – 2017. – Вып. 8. Информационные технологии. С. 77- 88.
 11. Vicentiy A.V., Dikovitsky V. V., Shishaev M.G. The Semantic Models of Arctic Zone Legal Acts Visualization for Express Content Analysis // Advances in Intelligent Systems and Computing, vol. 763, pp. 216-228. Springer, Cham.
 12. Стратегия развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года URL: <http://docs.cntd.ru/document/499002465>
 13. Стратегия социально-экономического развития Мурманской области до 2020 года и на период до 2025 года, утвержденной постановлением Правительства Мурманской области от 25 декабря 2013 г. N 768-ПП/20

Сведения об авторах

Вицентий Александр Владимирович – к.т.н., старший научный сотрудник

e-mail: alx_2003@mail.ru

Vicentiy Alexander – PhD (Tech. Sci.), senior researcher

Шишаев Максим Геннадьевич – д.т.н, доцент, профессор РАН, главный научный сотрудник ИИММ КНЦ РАН

e-mail: shishaev@iimm.ru

Maxim G. Shishaev – Dr.Sci. (Tech.), associate professor, professor of RAS, lead researcher

Гогоберидзе Георгий Гививич – д.э.н., доцент

e-mail: gogoberidze.gg@gmail.com

Gogoberidze Georgii – Dr.of Sci., associate professor

УДК 004.5, 004.9

DOI: 10.25702/KSC.2307-5252.2018.10.99-111

А.В. Вицентий, М.Г. Шишаев, В.В. Диковицкий

Институт информатики и математического моделирования ФИЦ КНЦ РАН

ОПЫТ ИСПОЛЬЗОВАНИЯ ВИЗУАЛИЗАЦИИ СЕМАНТИЧЕСКИХ МОДЕЛЕЙ ДОКУМЕНТОВ ДЛЯ ЭКСПРЕСС КОНТЕНТ-АНАЛИЗА (НА ПРИМЕРЕ НОРМАТИВНО-ПРАВОВЫХ АКТОВ, РЕГЛАМЕНТИРУЮЩИХ РАЗВИТИЕ ПРОСТРАНСТВЕННО-РАСПРЕДЕЛЕННЫХ СИСТЕМ)*

Аннотация

В настоящее время большие объемы данных доступны в текстовой форме. Однако из-за характерных особенностей текста на естественных языках, разработка полностью автоматических методов анализа семантики текстов является сложной задачей. В настоящей статье описывается состав, структура и некоторые области применения разработанных технологий семантического анализа и визуализации семантических моделей текстовых документов. Также описываются методы визуального экспресс контент-анализа документов. Эти методы являются частью технологии визуализации семантических моделей текстовых документов и реализованы в виде независимых программных инструментов. Для демонстрации основных возможностей технологии, подробно описывается опыт использования визуализации семантических моделей документов для визуального экспресс контент-анализа нормативно-правовых актов, регламентирующих развитие пространственно-распределенных систем различного уровня и анализ полученных результатов. В заключении делается вывод о перспективных областях использования разработанных технологий, а также определяются основные направления дальнейшей работы и возможности по расширению функциональности методов визуального экспресс контент-анализа текстовых документов.

Ключевые слова:

визуальный анализ документов, контент-анализ, человек-компьютерный интерфейс, управление пространственно-распределенными системами, tensorflow, TF-IDF

A.V. Vicentiy, M.G. Shishaev, V.V. Dikovitsky

EXPERIENCE IN USING THE VISUALIZATION OF DOCUMENT SEMANTIC MODELS FOR EXPRESS CONTENT ANALYSIS (ON THE EXAMPLE OF LEGAL ACTS REGULATING THE DEVELOPMENT OF SPATIALLY DISTRIBUTED SYSTEMS)

Abstract

Currently, large amounts of data are available in text form. However, due to the characteristic features of the text in natural languages, the development of fully automatic methods for analyzing the semantics of texts is a difficult task. This paper describes the composition, structure and some areas of application of the developed technologies of semantic analysis and visualization of semantic models of text documents. Also, methods for visual express content analysis of documents are described. These methods are part of the technology for visualizing semantic models of text documents and implemented as independent software tools. To demonstrate the main features of the technology, the experience of using the visualization of semantic document models for visual express content analysis of legal acts regulating

* Исследование выполнено при финансовой поддержке РФФИ и Министерства образования и науки Мурманской области в рамках научных проектов № 17-47-510298 p_a и 17-45-510097 p_a.

the development of spatially-distributed systems of various levels and analysis of the results is described in detail. The final part of the paper identifies some promising areas of application of the developed technologies, as well as determines the main directions for further work and the possibilities to expand the functionality of the methods of visual express content analysis of text documents.

Keywords:

documents visual analysis, content analysis, human-computer interface, management of spatially-distributed systems, tensorflow, TF-IDF

1. Введение

С увеличением количества и объема документов, используемых для решения различных прикладных задач, становится все труднее выделить достаточное время для их изучения. Это приводит к распространению ситуаций, когда приходится отказаться от подробного изучения документа, ограничившись кратким обзором, чтобы понять основной смысл. Таким образом, мы жертвуем точностью понимания смысла документа в пользу скорости его интерпретации. Что касается текстовых документов, этот процесс часто называют «диагональным чтением». С диагональным чтением мы получаем некоторый набор основных мыслей, обобщающих смысл документа. Получив общее представление о значении документа, мы интегрируем его как неотъемлемый объект в нашу ментальную систему понятий и, при необходимости, можем «извлечь» его для более детального изучения. В этой работе мы делаем следующий шаг в этом направлении: получение наиболее общего представления о значении документа, которое может быть выражено в одном изображении, интерпретированном одним взглядом. Мы будем называть этот процесс «быстрым пониманием». Благодаря такому подходу, мы можем применять подход визуального анализа для анализа семантики текстовых документов.

Работа с электронными документами воплощает задачу быстрого распознавания смысла в разных ситуациях. В этой статье рассматриваются три случая:

1. Представление результатов поиска. Даже самая сложная поисковая система гарантирует только релевантность результата, то есть его соответствие сформулированному запросу, но не реальным ожиданиям пользователя. Удовлетворение реальных ожиданий пользователя должно являться главным свойством результата поиска. Это свойство называется пертинентностью. Для его априорной гарантии можно использовать разные подходы к организации поисковых систем, в том числе предложенные авторами [1]. Однако, для этого необходимо иметь некоторую модель психических стереотипов каждого пользователя, которую чрезвычайно сложно обеспечить в информационных системах массового использования. Поэтому потенциально эффективным подходом может быть дополнение документа в списке результатов поиска изображением, характеризующим его обобщенное значение. Это обеспечит апостериорную быструю оценку соответствия результатов поиска ожиданиям пользователей. В этом случае главная задача визуализации - убедиться, что пользователь в сжатом виде понимает смысл документа.

2. Идентификация противоречивых документов. Этот случай основан на предположении, что документы, которые идентичны в терминах набора понятий, но имеют разную семантическую структуру, вероятно, будут противоречить друг

другу. Методика быстрого распознавания смыслов поможет выявить такие потенциальные конфликты в случае большого количества или большого объема исходных документов. Согласно первоначальному предположению, главная задача визуализации в этом случае - отразить сходство терминов, используемых в рассматриваемых документах, а также силу отношений между терминами.

3. Разделение набора документов на группы по смыслу. Хотя обычной практикой является априорное распределение похожих документов в разные папки, во многих случаях папка может содержать огромное количество документов, что заставляет пользователя кластеризовать набор документов. Этот случай похож на предыдущий, за исключением того, что нет необходимости понимать значение как таковое для разделения большого количества документов на аналогичные кластеры - достаточно идентифицировать только сходство значения двух или более документов. Потенциально это делает другие методы визуализации обобщенного смысла документа более эффективными.

В этой работе мы рассмотрим двухэтапный способ получения общего семантического представления документа. На первом этапе в автоматическом режиме выполняется семантический анализ документа. Это делается путем применения методов анализа контента и лексикографического анализа к текстовым документам. Первое реализовано путем подсчета TF и других мер, а второе – с помощью нейронной сети, предназначенной для семантического разрешения текстовых предложений. Результатом первого этапа является взвешенная семантическая сеть, которая характеризует частоту использования понятий в документе, а также наличие и силу семантических связей между понятиями. Этот семантический образ документа визуализируется как единое целое на втором этапе. Затем этот визуальный образ представляется пользователю для быстрого понимания смысла документа. Таким образом, предлагаемый подход представляет собой комбинацию семантического и визуального анализа.

Необходимость быстро обрабатывать постоянно увеличивающиеся объемы информации в текстовых документах существует в самых разных сферах деятельности. В этой статье мы рассматриваем проблему с точки зрения регионального административного управления. В этой области существует много объектов управления с довольно сложными отношениями, вызванными административной иерархией, территориальными особенностями и другими факторами. Это создает благодатную почву для появления большого количества документов с перекрывающимися, а также противоречивыми смыслами. В качестве примера для исследования мы взяли два документа, отражающих стратегические приоритеты России в Арктическом регионе:

1. Стратегия социально-экономического развития Мурманской области до 2020 года и на период до 2025 года (далее – «Стратегия МО») [2];
2. Стратегия развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года (далее – «Стратегия АЗРФ») [3].

Выпущенные на разных административных уровнях и, очевидно, имеющие пересекающиеся субъекты рассмотрения (Мурманская область является частью арктической зоны Российской Федерации), эти документы являются хорошим примером для рассмотрения.

Остальная часть статьи разделена на два основных раздела. В первом разделе рассматриваются подходы, методы и программные средства для

формирования семантической модели документа. Во втором разделе представлены результаты формирования семантических изображений вышеуказанных документов, а также предварительные выводы о перспективах и проблемах использования различных методов визуализации для быстрого распознавания смысла в контексте трех перечисленных выше проблем.

2. Технология формирования семантической модели документа

Для анализа текста используются различные методы извлечения формализованной семантики, основанные на аппарате лингвистики, статистических методах, математической логике, кластерном анализе, методах искусственного интеллекта и других методах и технологиях. Подходы к обработке и анализу текста можно разделить на статистические и лингвистические. Статистический подход основан на предположении, что содержание и, частично, семантика текста отражается наиболее часто встречающимися словами. Суть статистического анализа - учет появления слов в документе или предложении. Общим подходом является присвоение каждому термину t в документе неотрицательного веса.

Вес термина можно рассчитать по-разному, простейшим является «вес», равный числу вхождений термина t в документе d , обозначаемому tf (term frequency) [4]. Этот метод взвешивания не учитывает дискриминационную силу этого термина. Поэтому, когда доступны статистические данные об использовании терминов в наборе документов, выполняется схема вычисления веса $tf-idf$, определяемая следующим образом:

$$tf - idf_{i,d} = tf_{i,d} \times idf_i, idf_i = \log \frac{N}{df_i} \quad (1)$$

df - частота документа, определяемая как количество документов в коллекции, содержащих t , idf - обратная частота документа, N - общее количество документов в коллекции. Основным недостатком статистических методов текстового анализа является невозможность учета семантики. Использование лингвистических методов текстового анализа наряду со статистическими позволяет преодолеть этот недостаток.

Предлагаемая технология анализа текста включает несколько этапов: графематический, морфологический, синтаксический и семантический анализ. Результаты работы каждого уровня используются следующим уровнем анализа в качестве входных данных. Графематический анализ выделяет элементы структуры текста: параграфов, абзацев, предложений, отдельных слов и т. д. Целью морфологического анализа является определение морфологических характеристик слова и его основной словоформы. Целью синтаксического анализа является определение синтаксической зависимости слов в предложении. В связи с присутствием в русском языке большого количества синтаксически омонимичных конструкций, наличием тесной связи между семантикой и синтаксисом, процедура автоматизированного синтаксического анализа текста является недостаточной для определения зависимостей между понятиями предложения. Сложность увеличивается экспоненциально при увеличении количества слов в предложении и числа используемых правил. Семантический этап определяет формальное представление смысла составляющих входной текст

слов и конструкций. Для учета и хранения контекста понятий предметной области, учета различных форм передачи синтаксиса, а также на решение проблемы равнозначности слов направлено формирование на основе коллекции документов семантической модели предметной области (СМПО) в виде структуры взвешенных семантических отношений. СМПО позволяет реализовать процедуры извлечения и хранения множественного контекста употребленных в документах понятий, частично решая проблему совместимости новой информации с уже накопленными знаниями, а также выявить противоречия в семантических образах документов, в случае, если новая информация противоречит накопленной. Ниже приведена процедура формирования СМПО.

На начальном этапе текст делится на предложения и подвергается грамматическому анализу и лемматизации. Для синтаксического анализа и определения морфологических характеристик слов используется грамматический словарь русского языка [5], тезаурус WordNet [6], а также библиотека определения синтаксических связей SyntaxNet, основанная на аппарате искусственных нейронных сетей Tensor-Flow [7]. Особенностью такого подхода является возможность производить анализ морфологии и синтаксиса для слов, отсутствующих в тезаурусе. TensorFlow представляет собой библиотеку для машинного обучения и глубокого исследования нейронных сетей в рамках научно-исследовательской организации Machine Intelligence. Система масштабируема и может быть использована на множестве устройств. Основу библиотеки составляют графы потоков данных, библиотека функционирует на уровне задания архитектуры нейронной сети и ее параметров. Данные в TensorFlow представлены в виде многомерных массивов данных с переменным размером – тензоров. Вычисления представляются в виде направленного графа, пути, по которым эти данные перемещаются - это ребра графа. Тензоры переходят от узла к узлу по ребрам графа [7]. Множество морфологических признаков, определяемых SyntaxNet (374), грамматических категорий (49), и типов зависимостей (37) заданы в нотации Universal Dependency [8]. Для обработки предложений библиотекой SyntaxNet слова предложений преобразуются в вектор библиотекой Word2Vec [9]. Векторные представления слов позволяют рассчитать смысловую близость между словами. Поскольку алгоритмы Word2Vec основаны на обучении нейронной сети, для достижения эффективной работы необходимо использовать большой текстовый корпус для обучения. Для обучения Word2Vec использовались подготовленные векторы, полученные в наборе данных Google News [11]. Модель содержит векторы для 3 миллионов слов и фраз. Фразы были получены с использованием подхода skip-gram [10]. Далее предложения в векторной форме подаются на входной слой нейронной сети, реализованной на основе TensorFlow и обученной на корпусе Universal Dependences. Русскоязычная часть текстового корпуса в Universal Dependences представлена корпусам SinTagRus [12] и Google Russian Treebank [13]. Результатом работы SyntaxNet является дерево зависимостей предложения и морфологические характеристики слов. Результатом анализа документа является взвешенный семантический образ документа. Семантический образ документа представляет собой семантическую сеть, набор понятий и набор ребер - отношений над понятиями. СМПО формируется в результате интеграции семантических образов документов. Структура сервиса семантического анализа представлена на рисунке.

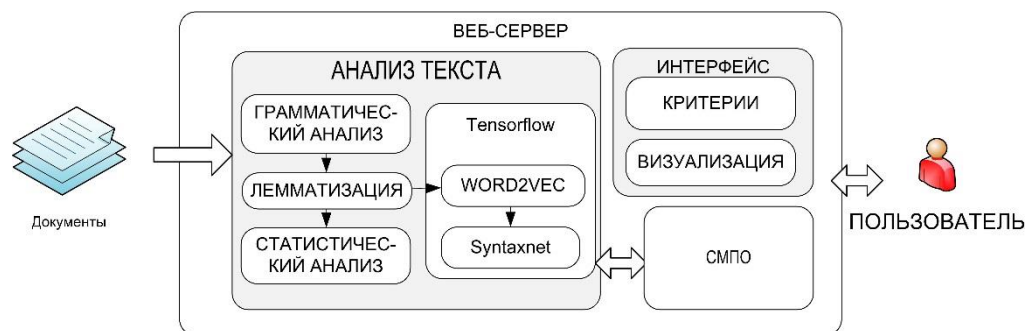


Рисунок 1. Веб-сервис семантического анализа.

С помощью данного сервиса проанализированы два документа - «Стратегия АЗРФ» и «Стратегия МО», получены две взвешенные семантические сети, содержащие 1993 слова и 2957 отношений для первого документа и 7847 слов и 18044 отношения для второго. Следующий шаг состоял в визуализации документов с целью обеспечения анализа контента.

3. Опыт визуализации семантических моделей документов в контексте различных задач визуального экспресс контент-анализа.

Для демонстрации основных возможностей технологии визуализации семантических моделей текстовых документов ниже приводятся примеры использования методов визуального экспресс контент-анализа документов. Эти методы являются частью технологии визуализации семантических моделей текстовых документов. Эти методы могут использоваться как в комплексе, так и по отдельности. Способы использования методов и интерпретации полученных результатов зависят от поставленной пользователем задачи.

Для проведения процедуры визуального контент анализа упомянутых выше «Стратегии социально-экономического развития Мурманской области до 2020 года и на период до 2025 года» и «Стратегии развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года» мы использовали три метода, которые были реализованы нами как независимые программные инструменты:

1. «Top 100 Words» - этот метод визуализирует сто самых важных понятий, которые характеризуют документ. Визуализация осуществляется в виде списка понятий документа, упорядоченного по весу понятий;
2. «Semantic Network» - этот метод визуализирует семантическую модель документа в виде семантической сети. Семантическая сеть документа состоит из понятий документа и различных типов отношений между ними;
3. «Semantic Networks Imposition» - этот метод визуализирует семантические модели документов в виде круга. Понятия документа помещаются на окружности, а отношения между ними отображаются внутри круга.

Ниже описан опыт практического применения этих методов для визуального экспресс контент-анализа нормативно-правовых актов, регламентирующих развитие пространственно-распределенных систем, и краткая интерпретация полученных результатов.

3.1 Пример использования метода «Top 100 Words»

Первым методом, для которого мы опишем пример использования для визуального экспресс-анализа содержимого документов, является метод «Top 100 Words». Этот метод, с одной стороны, обеспечивает быстрое создание визуального образа документа даже для относительно больших документов а, с другой стороны, хорошо подходит для оценки тематики анализируемого документа в целом. Фактически, этот метод визуализирует наиболее важные слова документа, которые можно считать ключевыми для описания его контента.

Для вычисления значения важности или веса слова используется статистическая мера TF-IDF. В текущей реализации метода используется абсолютное значение важности слов. То есть при вычислении веса слова, объем документа не учитывается.

Мы применили метод «Top 100 Words» к «Стратегии МО» и «Стратегии АЗРФ» независимо друг от друга. Фрагмент результатов работы метода представлен на рисунке 2. На рисунке показано только одиннадцать концептов (слов) с максимальным весом из каждого документа.

«Стратегия МО»			«Стратегия АЗРФ»		
	WORD	TF-IDF		WORD	TF-IDF
1	развитие	161	1	Арктической	190
2	области	141	2	Российской	173
3	региона	126	3	Федерации	161
4	Мурманской	120	4	развитие	90
5	государственной	93	5	обеспечение	79
6	населения	79	6	зоны	68
7	повышение	65	7	государственной	61
8	обеспечение	63	8	системы	52
9	Российской	62	9	безопасности	33
10	создание	48	10	деятельности	30
11	системы	42	11	населения	26

Рисунок 2. Результаты анализа документов с помощью метода «Top 100 Words» (фрагмент)

В зависимости от задачи анализа, пользователь может по-разному интерпретировать полученные результаты. Но, даже на первый взгляд, видно, что из множества слов с максимальным весом, более 50% слов совпадают в обоих документах. Это, в частности, может свидетельствовать о том, что эти документы имеют общую тематику и схожую терминологию. На основании этого вывода, можно говорить о том, что при проведении разбиения коллекции документов на несколько кластеров, эти документы с высокой вероятностью могут попасть в один и тот же кластер, так как обладают некоторой общностью.

Результаты анализа документов с использованием метода «Top 100 Words» могут быть использованы для повышения эффективности задач поиска документов на этапе анализа результатов поиска. Чтобы повысить эффективность анализа результатов поиска, каждый документ должен быть дополнен визуальным образом документа. В этом случае визуальный образ документа представляет собой список основных концептов (понятий) документа. Список понятий упорядочивается по значению весов понятий.

Однако для решения задач выявления потенциальных противоречий в документах метод «Top 100 Words» не подходит, поскольку он не предоставляет никакой дополнительной информации о документе.

Данный метод имеет и некоторые недостатки. На самом деле, метод «Top 100 Words» позволяет настраивать количество концептов, которые выводятся на экран. При этом возникает вопрос о том, какое количество концептов является оптимальным. С одной стороны, количество концептов должно быть достаточным для решения задачи пользователя, но с другой стороны, количество концептов не должно быть слишком большим. Если выводить на экран слишком много концептов, то сложность задачи визуального анализа этих концептов может приближаться к сложности задачи анализа исходного документа. В этом случае, применение метода «Top 100 Words» для визуального экспресс контент-анализа документа теряет смысл.

Если руководствоваться особенностями восприятия визуальной информации человеком, в частности «магическим числом Миллера» [14], и выводить на экран от 5 до 9 концептов с максимальным весом, то в случае больших документов существует высокая вероятность того, что результаты визуального анализа малого количества концептов будут бесполезны для пользователя.

На данный момент у нас нет однозначного ответа на вопрос о том, какое количество концептов является оптимальным для визуального экспресс контент-анализа документа. Решение данного вопроса запланировано нами в будущей работе.

3.2. Пример использования метода «Semantic Network»

Вторым методом, для которого мы опишем пример использования для визуального экспресс-анализа содержимого документов, является метод «Semantic Network». Результатом визуализации в этом случае является семантическая сеть документа.

Эта семантическая сеть не является статическим визуальным образом документа. Пользователь может взаимодействовать с ним в интерактивном режиме: разворачивать сетевые узлы, исследовать связи между основными концептами документа, визуализировать семантическую сеть документа, начиная с определенного пользователем концепта, ограничивать количество отображаемых концептов, изменять параметры отображения семантической сети и т.д.

Для демонстрации возможностей метода «Semantic Network», мы визуализировали семантические модели документов «Стратегия МО» и «Стратегия АЗРФ». Точкой входа для визуализации семантических сетей и для первого и для второго документа был выбран концепт «развитие». Мы выбрали этот концепт в качестве точки входа визуализации по двум причинам. Во-первых,

по результатам использования метода «Top 100 Words» этот концепт попал в список наиболее важных слов как в «Стратегии МО» (первое место списка), так и в «Стратегии АЗРФ» (четвертое место списка). Во-вторых, сумма весов концепта «развитие» для этих двух документов, является максимальной, в сравнении с суммами весов других пар совпадающих концептов.

Для сравнения, фрагменты визуальных образов семантических сетей «Стратегии МО» и «Стратегии АЗРФ», полученные с помощью метода «Semantic Network» приведены на рисунке 3.

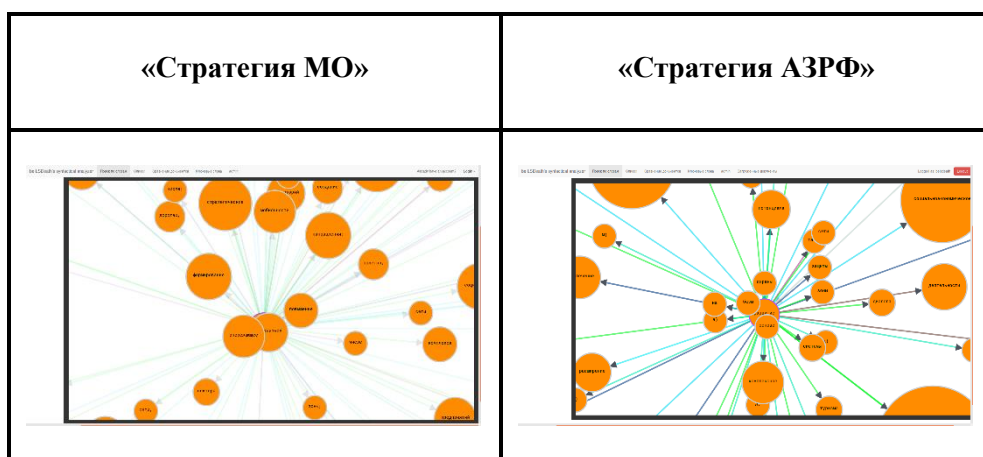


Рисунок 3. Визуальное представление семантических сетей анализируемых документов

Визуальный анализ и работа с семантическими сетями документов осуществляется конечным пользователем в соответствии с заранее определенными задачами анализа. Для проведения более подробного анализа, пользователю предоставляются разнообразные возможности для интерактивного взаимодействия с семантическими сетями (масштабирование, разворачивание узлов, отображение типов связей, перестроение семантической сети и др.). Даже без расчетов, полагаясь только на визуальный анализ семантических сетей документов, можно сделать несколько выводов. В частности, можно обнаружить, что количество концептов, связанных с понятием «развитие» в «Стратегии МО», значительно превышает количество концептов, связанных с тем же понятием в «Стратегии АЗРФ». Это может говорить о том, что развитию различных видов деятельности в «Стратегии МО» уделяется больше внимания, чем в «Стратегии АЗРФ». Масштабируя визуальное отображение семантической сети «Стратегии МО», можно более подробно изучить связи концепта «развитие» с другими концептами и сделать вывод о том, развитию каких именно видов деятельности уделяется больше всего внимания в этом документе.

Приведенный пример визуального экспресс контент-анализа документов не является единственным вариантом использования метода «Semantic Network». В частности, этот метод может быть использован для решения задач повышения pertinентности поиска по корпусу документов и других задач, для которых важны вес и связи концептов в анализируемом документе. Что касается задачи

выявления потенциальных противоречий в документах, метод «Semantic Network» может дать лишь предварительный ответ о потенциальной возможности наличия противоречий в документах. Для получения более точного ответа, требуются дополнительные усилия со стороны пользователя, связанные с более глубоким анализом смысла концептов и типов связей между ними.

Метод «Semantic Network» можно также использовать и для визуальной кластеризации документов. Для эффективного решения этой задачи важно обеспечить одинаковую ориентацию для одинаковых концептов.

В свою очередь, задача визуализации концептов в пространстве относительно друг друга является нетривиальной задачей. В частности, библиотека Word2Vec, которая используется в технологии семантического анализа документов, использует 300-мерные вектора для описания слов и фраз. Отображение слов в трехсотмерном пространстве не имеет смысла для визуального экспресс контент-анализа документа. Поэтому, необходимо проводить редуцирование пространства визуализации семантической модели документа в зависимости от решаемой пользователем задачи. Разработку метода редуцирования пространства визуализации семантической модели документа в зависимости от решаемой пользователем задачи планируется реализовать в будущих работах.

Как и для метода «Top 100 Words», для метода «Semantic Network» также остается актуальным вопрос об оптимальном количестве концептов и связей, которые следует выводить пользователю для визуального экспресс контент-анализа документа. Стоит отметить, что за счет возможностей интерактивного взаимодействия пользователя с визуальным образом семантической сети документа при использовании метода «Semantic Network» этот вопрос не является чрезвычайно важным. Прежде всего, это связано с тем, что пользователь может самостоятельно управлять настройками визуализации, такими как количество выводимых концептов, минимальный вес выводимого концепта, сворачивание и разворачивание узлов сети и т.д. Таким образом, пользователь может самостоятельно настроить оптимальное для него отображение семантической сети документа с учетом решаемой задачи и индивидуальных особенностей восприятия визуальной информации.

Решение вопроса об оптимальном способе визуализации семантической сети для конечного пользователя требует проведения дополнительных исследований и будет рассмотрено нами в будущей работе.

3.3 Пример использования метода «Semantic Networks Imposition»

Для визуализации различий в структуре семантических отношений между понятиями двух документов мы использовали JavaScript библиотеку визуализации D3.js. Понятия, присутствующие в обоих сравниваемых документах, расположены одинаково на обеих диаграммах. Линии представляют семантические отношения между основными понятиями в семантическом образе каждого документа. Визуализация семантических образов документов «Стратегия МО» и «Стратегия АЗРФ» представлена на рисунке 4.

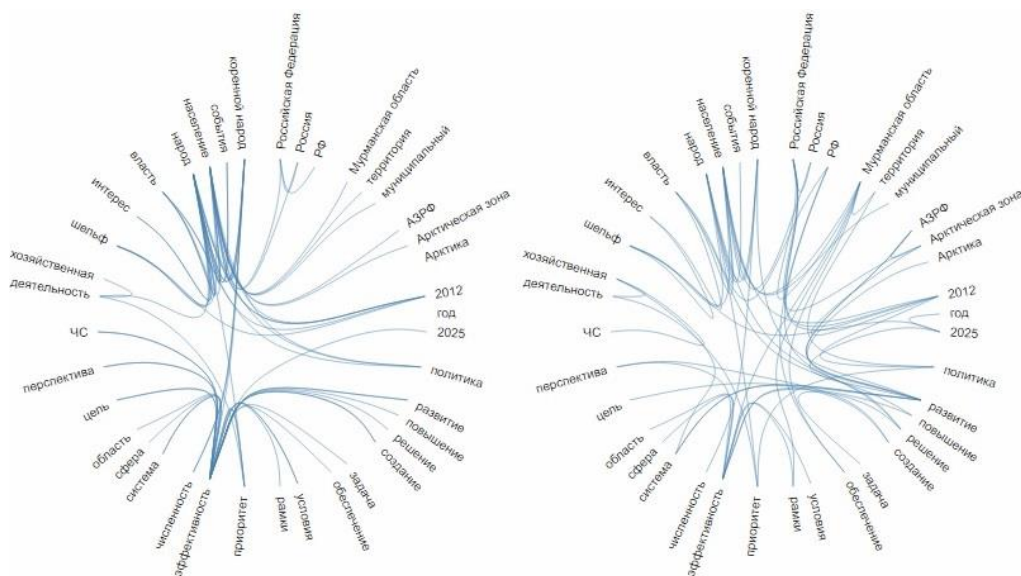


Рисунок 4. Визуальное представление семантических сетей анализируемых документов

Анализ визуализированных семантических образов позволяет идентифицировать наиболее часто встречающиеся отношения между понятиями. В то же время можно быстро определить разницу между структурами отношений в разных документах. На основе этого анализа можно сформировать гипотезы о важности отношений между основными понятиями анализируемых документов. Например, очевидный вывод состоит в том, что более подробно излагается второй документ, в то время как авторы первого документа оперируют более общими понятиями. Рассмотрим структуру отношений понятия «развитие» при визуализации семантических образов обоих документов. Понятие «развитие» рассматривается в первом документе в контексте показателя эффективности, в то время как во втором «развитие» рассматривается в контексте конкретных мест и времени. Это может быть косвенным признаком наличия несоответствий и противоречий в этих документах. Однако, чтобы подтвердить это предположение, необходимо провести более глубокий анализ и сравнение документов.

4. Заключение

В этой статье мы описали примеры применения визуализации семантических моделей документа для визуального экспресс-анализа содержимого. В качестве методов визуального экспресс-анализа содержимого документов были рассмотрены методы «Top 100 Words», «Semantic Network» и «Semantic Networks Imposition». Используя методы визуализации, описанные в статье, были получены следующие выводы об анализируемых документах:

1. В анализируемых документах список наиболее значимых концептов аналогичен. Пример применения метода «Top 100 Words», описанного выше, показал, что совпадение понятий составляет более 50%.

2. Количество связей между концептами в рассмотренных документах значительно различается. Визуализация семантической модели документов, выполненная с использованием метода «Semantic Network», позволяет визуально оценить разницу в количестве связей, не выполняя вычислений.
3. В документе «Стратегия МО» связи между концептами описаны более подробно, чем в «Стратегии АЗРФ». Такой вывод можно сделать на основе анализа визуализации, выполненного методом «Semantic Networks Imposition».

Также мы можем сделать некоторые выводы о методах:

1. Основываясь на нашем опыте анализа документов, можно сделать вывод, что методы визуального экспресс-анализа содержимого, рассмотренные в этой статье, не являются универсальными.
2. Эффективность конкретного метода зависит от типа решаемой задачи (кластеризация, повышение пертинентности, выявление противоречий).
3. Примеры визуального анализа документов, описанные в статье, позволяют нам говорить о том, что все рассмотренные методы визуального экспресс-анализа контента могут быть успешно применены при решении задач, для которых достаточно понять общий смысл документа на основе визуального отображение его наиболее значимых семантических аспектов.
4. Основываясь на полученном опыте, мы можем сделать предварительный вывод о том, что метод «Top 100 Words» лучше использовать в задачах кластеризации документов, метод «Semantic Network» - для быстрого понимания сути документа и повышения пертинентности результатов поиска, а метод «Semantic Networks Imposition» - для сравнения и идентификации потенциальных противоречий в структуре документа.

Но окончательные выводы можно сделать только после проверки этих методов на большом наборе документов.

5. Обсуждение результатов и будущая работа

В целом, наш опыт использования методов визуализации для экспресс-анализа содержания правовых актов, регулирующих развитие пространственно распределенных систем, может быть оценен как положительный. Визуализации, созданные этими методами, позволяют получить новые знания об анализируемых документах без детального изучения контента документов.

В своих будущих работах мы планируем провести ряд более подробных экспериментов для больших коллекций документов, чтобы найти ответы на следующие вопросы:

1. Для каких типов задач визуализация семантических моделей документа для экспресс-анализа контента может быть наиболее эффективной?
2. Как эффективность визуального анализа зависит от характеристик анализируемого документа?
3. Как наилучшим образом представить результаты визуализации пользователю?
4. Как можно редуцировать визуальный образ с учетом характеристик задачи анализа?

Литература

1. Shishaev, M.G. Architecture and Technologies of Knowledge-Based Multi-Domain Information Systems for Industrial Purposes/ V.V. Dikovitsky M. G. Shishaev , N. V. Nikulina //Automation Control Theory Perspectives in Intelligent Systems. Proceedings of the 5th Computer Science On-line Conference 2016 (CSOC2016), Vol 3. pp. 359 - 369
2. Стратегия социально-экономического развития Мурманской области до 2020 года и на период до 2025 года, утвержденной постановлением Правительства Мурманской области. – Режим доступа: <http://docs.cntd.ru/document/465602093>
3. Стратегия развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года. – Режим доступа: <http://docs.cntd.ru/document/499002465>
4. Salton, G. and McGill, M. J. 1983 Introduction to modern information retrieval. McGraw-Hill, ISBN 0-07-054484-0
5. Зализняк, А.А. Грамматический словарь русского языка. <http://odict.ru>
6. Тезаурус WordNet. <http://wordnet.ru>
7. TensorFlow. <https://www.tensorflow.org>
8. Universal Dependencies. <http://universaldependencies.org>
9. Word2Vec. <https://code.google.com/archive/p/word2vec/>
10. Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed Representations of Words and Phrases and their Compositionality. In Proceedings of NIPS, 2013.
11. Google News. Режим доступа – <https://drive.google.com/file/d/0B7XkCwpI5KDYNINUTTISS21pQmM/edit?usp=sharing>
12. SinTagRus. Режим доступа – <http://www.ruscorpora.ru/search-syntax.html>
13. Google Russian Treebank. Режим доступа – <https://old.datahub.io/dataset/universal-dependencies-treebank-russian>
14. Miller, G.: The Magical Number Seven, Plus or Minus Two. In: The Psychological Re-view, vol. 63, 1956 pp. 81 – 97

Сведения об авторах

Шишаев Максим Геннадьевич – д.т.н, доцент, профессор РАН, главный научный сотрудник ИИММ КНЦ РАН
e-mail: shishaev@iimm.ru

Maxim G. Shishaev – Dr.Sci. (Tech.), associate professor, professor of RAS, lead researcher

Диковицкий Владимир Витальевич – к.т.н, старший научный сотрудник
e-mail: dikovitsky@gmail.com

Vladimir V. Dikovitsky – PhD (Tech. Sci.), senior researcher

Вицентий Александр Владимирович – к.т.н., старший научный сотрудник, доцент
e-mail: alx_2003@mail.ru

Vicentiy Alexander – PhD (Tech. Sci.), senior researcher, associate professor

В.В. Диковицкий^{1,2}, А.А. Менькова^{1,2} Л.М. Шишаева³

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² Филиал ФГБОУ ВО «МАГУ» в г. Апатиты

³ Санкт-Петербургский государственный университет

МОДАЛЬНАЯ ЛОГИКА КАК СРЕДСТВО ФОРМАЛИЗАЦИИ НОРМАТИВНЫХ ДОКУМЕНТОВ

Аннотация

В статье рассмотрены предпосылки появления противоречий в нормативных документах, аргументы в пользу необходимости автоматизации поиска противоречий с помощью модальной логики. Описаны основные составляющие модальной логики, сделан упор на деонтические модальности. Рассмотрены основные обозначения, аксиомы и правила вывода деонтической логики. Проанализирован существующий опыт создания логических систем с применением аппарата модальной логики.

Ключевые слова:

Противоречия в нормативных документах, модальная логика, модальность, деонтическая логика

V.V. Dikovitskiy, A.A. Menkova, L.M. Shishaeva

MODAL LOGIC AS A MEANS OF FORMALIZATION OF NORMATIVE DOCUMENTS

Abstract

The article describes the prerequisites for the emergence of contradictions in regulatory documents, the arguments in favor of the need to automate the search for contradictions using modal logic. The main components of modal logic are described, emphasis is placed on deontic modalities. The main notation, axioms and rules for the derivation of deontic logic are considered. The existing experience of creating logical systems using the modal logic apparatus is analyzed.

Keywords:

Contradictions in regulatory documents, modal logic, modality, deontic logic

Введение

Нормативные документы – особый вид документов. [1] С одной стороны, законы, распоряжения, нормативные акты, указы, постановления и т.п. являются текстами на естественном языке (русском, английском и т.д.), с другой стороны, назначение этих текстов – применение к конкретной ситуации в предметной области и вывод о том, соответствует ли данная ситуация Закону. [2]

Обширность базы нормативно-правовой документации, ее постоянное увеличение и модернизация приводят к появлению противоречий. Многие аспекты противоречий не поддаются строгой формализации что делает затруднительным их автоматическое выявление с применением вычислительной техники [3].

Тем не менее формализация закона (или любого нормативного акта) необходима для того, чтобы средствами математической логики выявлять в нем

возможные противоречия, неполноту, противоречия с другими законами [4]. Это способствовало бы улучшению, как законодательной работы, так и правоприменительной деятельности.

Важнейшим условием обеспечения целостности правового поля [5] является выявление противоречий, которые нередко обнаруживаются при сопоставлении действующих или формировании новых нормативно-правовых актов, касающихся различных аспектов одного и того же объекта права. Для решения данных проблем разрабатываются технологии автоматизированного контроля целостности правового поля при разработке и использовании нормативно-правовых документов [6].

Практическая ценность подобных разработок очевидна. Автоматизированная обработка нормативных актов позволит выявить противоречия в законодательстве. Она может быть использована на предприятиях и в учреждениях при подготовке локальных актов для предупреждения противоречий с законом и с ранее принятыми локальными актами, что позволит избежать массы проблем при работе с нормативными документами [2]. Исходя из всего вышеизложенного, автоматизированный поиск противоречивых высказываний в нормативных документах – актуальная задача на сегодняшний день.

Нормативные понятия всегда были в центре внимания философов и логиков. Так, особое внимание к возможности применения модальных логик в сфере нормативных систем уделял аргентинский логик и философ права – Карлос Эдуардо Альчуррон.

Деонтическая логика (подвид модальной логики, см. ниже), которая имеет дело с нормативным использованием языка и с такими нормативными понятиями, как обязательство, запрет и дозволение, позволила пролить свет на различные аспекты феномена нормативности; а разработка немалого количества систем деонтической логики демонстрирует интерес к нормативному дискурсу, и не малый. Несмотря на то, что такая наука, как юриспруденция, изначально занимается описанием и систематизацией правовых норм, формальному исследованию нормативных понятий должного внимания юристами не уделялось [7].

Невозможность строго классифицировать науку о праве как эмпирическую, так и точную; но необходимость в создании четкой системы норм с определенными стандартами рациональности и целостной непротиворечивости правовых норм, представляет модальную логику наиболее удачным инструментом для автоматизации поиска противоречий в нормативных документах. Но для начала необходимо выяснить, что такое модальная логика, и из каких частей она состоит.

О модальной логике

Модальная логика – раздел логики, в котором исследуются логические связи модальных высказываний [8], то есть, помимо стандартных логических связок, переменных или предикатов (утверждения, высказанные о субъекте) имеются модальности. Стоит отметить, что модальная логика слагается из ряда разделов, или направлений, каждое из которых занимается модальными высказываниями определенного типа.

Модальность – это оценка высказывания, данная с той или иной точки зрения. Она выражается с помощью понятий: «необходимо», «возможно»,

«доказуемо», «опровержимо», «обязательно», «разрешено» и т.п. Модальными высказываниями являются высказывания, содержащие хотя бы одно из таких понятий. Модальные высказывания делятся на типы (логический, физический, эпистемический, оценочный, нормативный) в соответствии с точкой зрения, на основе которой формулируются выражаемые ими характеристики [8].

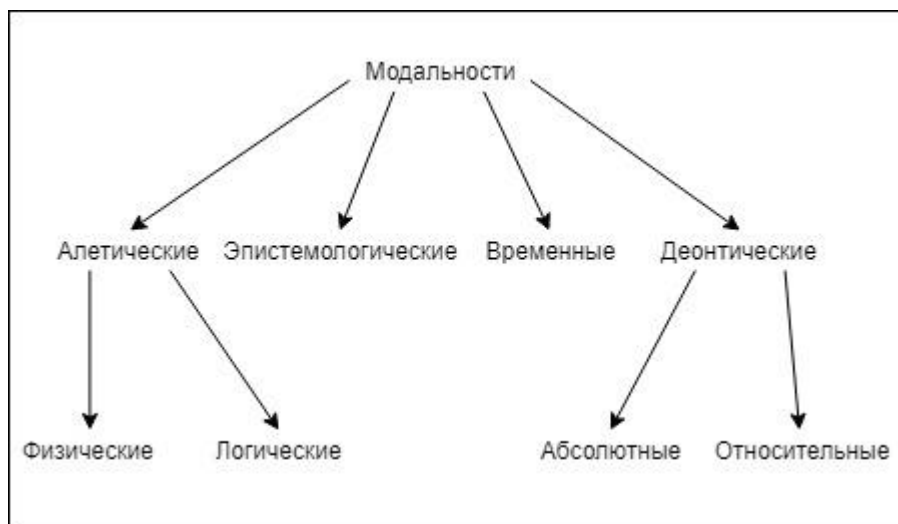


Рис.1. Разновидности модальностей.

Существует несколько разновидностей модальностей (рис. 1):

- Алетические модальности выражают информацию о логической либо фактической обоснованности суждения с помощью операторов «необходимо», «случайно», «возможно», «невозможно». В свою очередь алетические модальности могут быть физическими и логическими.

Физические модальные высказывания формируются с помощью физических модальных понятий (физически необходимо, физически возможно и т.п.), называемых также онтологическими или каузальными.

Логические модальные понятия связаны с «механикой» человеческого мышления и используются для характеристики существенных ее моментов. Физические модальные понятия касаются устройства самого реального мира.

Нечто необходимо, если оно не может быть иным, чем оно есть. Именно в зависимости от того, на какое основание опирается утверждение о необходимости, выделяются логическая и физическая необходимость. Логическая необходимость связана с логическим законом: логически необходимы законы логики и все, что вытекает из них. Физическая необходимость связана с законами природы: физически необходимо то, отрицание чего нарушает законы природы. Физически необходимы, например, высказывания: «Все планеты вращаются вокруг своей оси» и «Электрон, движущийся по стационарной орбите, не излучает энергию». Отрицания этих высказываний противоречили бы законам физики: отрицание первого высказывания несовместимо с законами небесной механики, отрицание второго – с законами квантовой механики [9].

- Эпистемологические модальности степенью достоверности и обоснованностью информационного суждения.

Эпистемическая логика – неклассическая логика, один из видов, модальной логики, а именно – логика модальностей знания и мнения (эпистемических модальностей) некоторого идеализированного агента. Объектами изучения для нее выступают пропозиции вида «х знает, что Р» и «х считает, что Р» [10].

- Временные модальности использует такие операторы как «всегда», «никогда», «одновременно», «раньше», «позже».

Темпоральная логика – подраздел модальной логики, в высказываниях которой учитывается временной аспект. Используется для описания последовательностей явлений и их взаимосвязи по временной шкале.

Есть два подхода темпоральной логики, основанные на принципах здравого смысла и диалектики: «после этого» означает «по причине этого», либо «после этого» означает «позже» в хронологическом смысле. В темпоральных логиках бывает два вида операторов: логические и модальные. Модальные операторы, используются в логике линейного времени и логике деревьев вычислений.

Временные логики применяются для описания и исследования причинно-следственных зависимостей, развивающихся во времени. Семантика темпоральных логик существенно зависит от той математической модели, которая используется для описания феномена времени. В самом общем случае в качестве модели времени можно взять любое частично упорядоченное множество. Элементы этого множества соответствуют различным моментам времени. В качестве темпоральных моделей могут выступать любые модели Крипке, построенные на основе частично упорядоченных шкал.

- Деонтические модальности – это выраженная в суждении информация, побуждающая людей к определенным поступкам.

Деонтическая логика – раздел модальной логики, в котором исследуются возможности применения средств символической логики к анализу морального, правового, политического, экономического и подобных им, нормативных, типов суждений. Деонтическую логику также называют логикой норм или нормативной логикой.

Деонтические логики подразделяются на абсолютные (монадические) и относительные (диадические) логики. В абсолютной логике перед высказыванием p может быть поставлен любой из префиксов «обязательно», «запрещено», «разрешено». В относительной деонтической логике к этой конструкции добавляется еще и условие, например, «обязательно p при условии S », где p и S – высказывания [11]. При этом ни один из видов деонтической логики не учитывает субъекта, который обязан или может выполнить определенное действие. Из-за этого в диадической деонтической логике возникают парадоксы.

В монадической схеме деонтической логики (в дальнейшем именуется просто «деонтическая логика») имеется пять возможных префиксных дополнений к высказыванию p :

- обязательно p (O p);
- разрешено p (PE p);

- запрещено p (IM p);
- необязательно p (GR p);
- опционально p ; (OP p).

Все утверждения деонтической логики могут быть выражены через дополнение O следующим образом:

$$PEp = \neg O\neg p \quad (1)$$

$$IMp = O\neg p \quad (2)$$

$$GRp = \neg Op \quad (3)$$

$$OPp = (\neg Op) \wedge (\neg O\neg p) \quad (4)$$

Таким образом, утверждается, что разрешено p тогда и только тогда, когда отрицание p не обязательно; запрещено, когда отрицание обязательно; опционально, когда ни p , ни отрицание p не являются обязательными (т.е. необязательно и разрешено).

Также следует понимать, что высказывание p – это высказывание о действии, возможном или невозможном, а выражение Op уже не имеет такого характера, поэтому повторное применение оператора O , т.е. запись OOp , не имеет никакого смысла, хотя синтаксически возможна.

Система аксиом деонтической логики, включающая аксиомы и правила вывода [13], строится на основе оператора O :

$$DA1: Op \rightarrow \neg O\neg p \quad (5)$$

$$DA2: O(p \wedge q) \rightarrow Op \wedge Oq \quad (6)$$

$$DA3: O(p \vee \neg p) \quad (7)$$

$$DR: p \leftrightarrow q \text{ выводимо} \rightarrow Op \leftrightarrow Oq \text{ выводимо.} \quad (8)$$

Здесь формула 8 задает аксиому DA1: «необязательно p » или «отказ от требования, права выполнить действие p », формула 9 – аксиому DA2: обязательны одновременно тогда и только тогда, когда обязательны поодиночке, 10 – аксиому DA3: обязательно выполнить или не выполнить действие, 11 – правила вывода DR: Если $p \leftrightarrow q$ выводимо, то и $Op \leftrightarrow Oq$ выводимо. Иными словами, если p и q логически эквивалентны, то Op и Oq логически эквивалентны.

Семантика деонтической логики строится на понятии возможных миров (или идеальных миров). Предполагается, что существует некоторое подмножество множества возможных миров, такое, что для некоторого предложения p предложение Op (означающее, что p обязательно) имеет место (выполнимо) тогда и только тогда, когда p выполнимо во всех этих мирах [14].

Модальная логика в задачах анализа утверждений

Анализ информации и моделирование рассуждений относятся к числу интенсивно развивающихся в настоящее время областей знания. Результаты этих исследований применимы во всех областях деятельности, связанных с формированием, изменением и обоснованием утверждений. Моделирование рассуждений традиционно основывается на логике, однако в последние десятилетия в данном направлении исследований применяются и методы других наук – лингвистики, статистики, психологии, методов искусственного интеллекта. Моделированию рассуждений и формализации высказываний посвящены работы [15,16] В работе [17] предложен механизм для задания иерархии и дополнений к деонтической логике с целью сохранения согласованности в присутствии конфликтов. Формальное представление и анализ

нормативной информации представлено в работах [18-21]. Последние десятилетия получило развитие направление автоматизации рассуждений. В работе [22] представлен логический вывод на основе деонтической логики. В рамках проектов ИПИ РАН разработан язык расширенных семантических сетей и продукционный язык ДЕКЛ для их обработки. [23] Они образуют законченный технологический комплекс, ориентированный на сложные задачи, связанные с логическим выводом, преобразованием представлений, экспертными решениями. В работе [24] предложена автоматизация проверки правильности спецификаций протоколов инициирования сеансов в виде мультиагентной системы, представляемой процессными моделями и описанием требований их правильности на языке временной модальной логики. Автоматизацию проверки правильности спецификаций предложено осуществлять логическими программами, получаемыми с помощью предлагаемой методики перехода от процессной модели описания спецификации и требований правильности на языке модальной логики к логической программе проверки правильности на языке логического программирования ПРОЛОГ. В работе [25] предложена система понятий, отношений и аксиом для нормативных документов, позволяющая строить математические модели для множеств взаимосвязанных документов и проводить их логический анализ.

В работе [26] предложено использовать эвристики наряду с логическим выводом. Наличие эвристик является неотъемлемой частью реально работающей ЭС, поэтому попытка «зажать» механизм вывода только в жесткий математический аппарат не принесет практического успеха. Следует применить более гибкий подход с использованием аппарата семантических и нейронных сетей, т.к. для использования только автоэпистемической логики нужны неограниченные временные и компьютерные ресурсы (в силу идеализации рассуждений) и требуется эвристика для ограничения этих рассуждений [26].

Заключение

Формализация нормативных актов необходима для того, чтобы средствами математической логики выявлять в нем возможные противоречия, неполноту, противоречия с другими законами. Это является важнейшим условием обеспечения целостности правового поля. Для решения данных проблем разрабатываются технологии автоматизированного анализа нормативных документов и контроля целостности правового. Автоматизированная обработка нормативных документов и их логический анализ могут быть использованы при подготовке локальных актов для предупреждения противоречий с законом и с ранее принятыми локальными актами.

Литература

1. Миков А.И. Представление онтологий нормативных документов с использованием прикладных логик // Известия ЮФУ. Технические науки, 2014 (в печати)
2. Аксиоматизация нормативных отношений на основе деонтической логики: методы и алгоритмы // Пекшева Марина Владимировна
3. <http://www.inteltec.ru/publish/articles/textan/RCDL2004.shtml#1>
4. Миков А.И. Нормативные мультиагентные системы // Известия ЮФУ. Технические науки, 2014 (в печати)

5. Капустина М.А. О понятии «правовое поле» / М.А. Капустина // Правоведение. 2006. № 6. С. 220-222
6. Разработка технологии проверки и согласования нормативно-правовой базы на основе онтологий // П.А. Ломов, А.Г. Олейник
7. Альчуррон К.Э. Отделение и отменяемость в деонтической логике / пер. с англ. Е.Н. Лисанюк, А.М. Павловой // «Нормативные системы» и другие работы по философии права и логике норм / К.Э. Альчуррон, Е.В. Булыгин, П. Герденфорс, Д. Макинсон; под ред. Е.Н. Лисанюк. СПб., 2013. 380 с
8. А. Ивин. Логика. Режим доступа: <https://www.e-reading.club/book.php?book=24044>
9. А. Ивин, Логика, учебник для гуманитарных факультетов. Психологическая библиотека Киевского Фонда содействия развитию психической культуры. Режим доступа: <http://psylib.org.ua/books/ivina01/txt08.htm>
10. Эпистемическая логика в проблематике искусственного интеллекта. Магистерская диссертация студента 2 курса Кляшторного А.С
11. Миков А.И. Формализация целей и ограничений ИТС средствами деонтической логики // АНО «Институт компьютеринга», г. Пермь, с.47
12. Горский Д.П., Ивин А.А., Никифоров А.Л. Краткий словарь по логике // Просвещение, Москва, 1991, с.208
13. Клини С. Математическая логика // Издательство Мир, Москва, 1973, с.480, Колмогоров А.Н., Драгалин А.Г. Введение в математическую логику // Издательство Московского университета, Москва, 1982, с.240]
14. Сидоренко Е.А. Логика. Парадоксы. Возможные миры. Размышления о мышлении в девяти очерках. // Едиториал УРСС, Москва, 2002, с.312
15. Пospelов Д.А. МОДЕЛИРОВАНИЕ РАССУЖДЕНИЙ. Опыт анализа мыслительных актов. Москва «Радио и связь» 1989 г.
16. Alchourrón Carlos E. And Bulygin Eugenio. Normative Systems. Library of Exact Philosophy, No. 5. Springer-Verlag, New York and Vienna 1971, XVIII + 208 pp.
17. Alchourrón, Carlos E. (1993). Philosophical Foundations of Deontic Logic and the Logic of Defeasible Conditionals. In Meyer & Wieringa 1993, 43–84.
18. Allen L. E., Saxon C. S. Analysis of the logical structure of legal rules by a modernized and formalized version of Hohfeld's fundamental legal conceptions. – 1985.
19. Mann B. H. The Formalization of Informal Law: Arbitration Before the American Revolution // NYUL Rev. – 1984. – Т. 59. – С. 443.
20. Francesconi E. et al. (ed.). Semantic processing of legal texts: Where the language of law meets the law of language. – Springer, 2010. – Т. 6036.
21. Brighi R., Palmirani M. Legal text analysis of the modification provisions: a pattern oriented approach // Proceedings of the 12th International Conference on Artificial Intelligence and Law. – ACM, 2009. – С. 238-239.
22. Von Wright G. H. Explanation and understanding. – Cornell University Press, 2004
23. Кузнецов И.П. Семантико-ориентированные системы на основе баз знаний / И.П.Кузнецов, А.Г. Мацкевич // МТУСИ, 2007 ISBN 5948740277. – Москва, 2007, 173с.
24. Мьё Тхет Наунг, Девятков Владимир Валентинович Формальная проверка правильности сетевого взаимодействия агентов, специфицируемых на

языке процессных выражений // Интернет-журнал Науковедение. 2016. №5 (36).

25. Миков Александр Иванович Представление онтологий нормативных документов с использованием прикладных логик // Известия ЮФУ. Технические науки. 2014. №6 (155). URL: <https://cyberleninka.ru/article/n/predstavlenie-ontologiy-normativnyh-dokumentov-s-ispolzovaniem-prikladnyh-logik> (дата обращения: 29.10.2018)
26. Ларионов, Д. С. (2005). Использование модальной логики для проектирования оболочек экспертных систем. Известия Томского политехнического университета. Инжиниринг георесурсов, 308 (4), 173-177.

Сведения об авторах

Диковицкий Владимир Витальевич – к.т.н, научный сотрудник
e-mail: dikovitsky@gmail.com

Vladimir V. Dikovitskiy – Ph.D. (Tech. Sci.), researcher

Менькова Анастасия Александровна – младший научный сотрудник,
e-mail: amenkova@iimm.ru

Anastasia A. Menkova – junior researcher

Шишаева Любовь Максимовна – студент СПбГУ,
e-mail: liubashishaeva@gmail.com

Liubov M. Shishaeva – student of SPbSU

П.А. Ломов

Институт информатики и математического моделирования ФИЦ КНЦ РАН

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ТЕХНОЛОГИИ ГЕНЕРАЦИИ ЛЕКСИКО-СИНТАКСИЧЕСКИХ ПАТТЕРНОВ ДЛЯ ПОДДЕРЖКИ РЕШЕНИЯ ЗАДАЧ ОБУЧЕНИЯ ОНТОЛОГИЙ*

Аннотация

Работа посвящена применению онтологических паттернов проектирования в рамках решения задач генерации, расширения и наполнения онтологий на основе анализа тематических естественно-языковых текстов. Рассмотрена технология автоматизированной генерации и применения лексико-синтаксических шаблонов для наполнения онтологии, а также ее программная реализация.

Ключевые слова:

онтология, лексико-синтаксические паттерны, обучение онтологий, онтологические паттерны содержания

P. A. Lomov

SOFTWARE IMPLEMENTATION OF THE TECHNOLOGY FOR GENERATION OF LEXICAL-SYNTACTIC PATTERNS FOR SUPPORT OF ONTOLOGY LEARNING

Abstract

The work is devoted to the application of ontological design patterns for solving problems of generation, expansion and population of ontologies based on the analysis of natural language texts. The technology of automated generation and use of lexico-syntactic patterns for ontology population, as well as its software implementation, are considered.

Keywords:

ontology, lexical-syntactic patterns, ontology learning, ontological content patterns

Введение

На сегодняшний день онтологические модели (онтологии) стали распространенным средством представления машиночитаемых знаний. Однако их разработка и сопровождение подразумевают решения целого ряда разноплановых задач, что требует существенных затрат труда и времени квалифицированных специалистов. Одной из таких задач является наполнения онтологии (ontology population) [1]. Она рассматривается в литературе как одна из подзадач обучения онтологий (ontology learning) [2] и заключается в добавлении в нее экземпляров к заданным в ней классам и определения отношений между этими экземплярами. При этом каких-либо структурных изменений иерархии классов, а также набора отношений не производится. Для автоматизации решения данной задачи могут использоваться лексико-синтаксические паттерны (шаблоны), которые обычно применяются в компьютерной лингвистике для решения задач автоматической обработки текстов. В

* Работа выполнена при финансовой поддержке РФФИ (проект № 16-07-00562) и поддержке РФФИ и Министерства образования и науки Мурманской области (проект № 17-47-510298 p_a).

рассматриваемом же случае они выступают абстрактными описаниями фрагментов текста, содержащих лексемы, которые могут рассматриваться как экземпляры классов онтологии и отношения между ними.

Такое применение лексико-синтаксических паттернов выступило предпосылкой появления еще одного вида онтологических паттернов проектирования (Ontology Design Patterns, ODP)[3] - лексико-синтаксических онтологические паттернов (Lexico-Syntactic ODPs, LSOP). Они представляют собой своего рода надстройку над лексико-синтаксическими паттернами. Отличие состоит в том, что LSOP позволяют не только распознавать языковые конструкции, но определяют соответствие между их составляющими и онтологическими элементами (классами, экземплярами, отношениями). В ряде случаев для отдельного LSOP даже может быть задан определенный онтологический паттерн содержания (Content ontology Design Patterns, CDP), который представляет фрагмент онтологии, описывающий типовую ситуацию (например, участие объектов в событии, выполнение роли объектом, наличие частей у объекта и др.) в предметной области. Наличие такого соответствия между синтаксическими конструкциями и структурными элементами онтологий позволяет применять LSOP при решении задач их автоматического построения и наполнения на основе текстовых корпусов.

В предыдущей работе [4] был рассмотрен вопрос автоматизированного синтеза разработчиком составных CDP, которые определяют более сложный онтологический фрагмент, позволяющий обеспечить представление знаний о некотором объекте предметной области с необходимой степенью детализации. В рамках данной работы предлагается технология генерации лексико-синтаксических паттернов на основе CDP для последующего получения текстовых фрагментов, которые могут содержать факты, удовлетворяющие ему. Генерация именно LSOP на данном этапе исследования не рассматривается. Комбинирование возможности автоматизированного синтеза CDP и последующей генерацией на его основе лексико-синтаксических паттернов позволит разработчику не только получить структуру представления понятий предметной области в онтологии, но и обеспечит его поддержку в решении задачи наполнения онтологии.

Использование лексико-синтаксических паттернов в задачах обучения и наполнения онтологии

Наиболее ранней зарубежной работой, в которой рассматривается применение лексико-синтаксическим паттернов для решения задач генерации и наполнения онтологий считается работа Марти Херста [5]. В ней анализировалась встречаемость в специализированном англоязычном текстовом корпусе фрагментов, удовлетворяющих паттернам, которые указывают на наличие родовых отношений между объектами. Например, из выражение «...works by authors such as Herrick, Goldsmith, and Shakespeare.», на основе паттерна «NP1 such as NP» (NP – именная группа) определяется отношение гипонимии между объектами, заданными лексемами «Herrick», «Goldsmith», «Shakespeare» и объектом, заданным лексемой «authors». Спецификой, отмечаемой автором, данных паттернов является их высокая точность и как следствие небольшое количество распознаваемых гипонимов/гиперонимов.

В последующих работах другие авторы часто используют паттерны Херста и делают попытки расширить их с учетом специфических языковых конструкций, характерных для предметной области [6,7], других видов отношений [8, 9], а также

дополнительных правил и эвристик [10]. Так, например, в работе [10] помимо лексико-синтаксическим паттернов Херста применялись некоторые LSOP [11]. Наряду с этим, были заданы контекстные паттерны (context patterns), которые позволяют при анализе текста учитывать классы, уже существующие в наполняемой онтологии. Дополнительно были определены смысловые ограничения, заданные на основе семантических групп англоязычного тезауруса WordNet.

Необходимо также отметить ряд успешных российских проектов, ориентированных на извлечение информации из русскоязычных текстов с применением лексико-синтаксическим паттернов. Данную задачу можно рассматривать как составляющую задачи наполнения онтологии, так как в результате ее решения из текстовых источников извлекаются структурированные описания фактов, которые потенциально могут быть использованы для формирования содержания онтологий. К таковым проектам относятся LSPL [12], RCO Pattern Extractor [13], Alex[14], Tomita-парсер [15]. Каждая из данных систем обладает своим языком записи паттернов, обладающих примерно одинаковой выразительностью. Среди них свободно доступными для использования в исследовательских целях являются системы LSPL и Tomita-парсер.

Особенности генерации лексико-синтаксических паттернов для наполнения онтологии

Как правило решение задачи наполнения онтологии предполагает наличие целевой онтологии, мультимедийного корпуса (обычно текстового) и средства извлечения (extraction engine) [16-18]. Средство извлечения используется для распознавания в корпусе наименований сущностей с целью последующего их представления в виде экземпляров классов, а также экземпляров отношений, заданных между классами. Оно также может дополнительно включать некоторые абстрактные описания структуры текстовых фрагментов, а также различные правила и эвристики, отражающих специфику предметной области.

В рассматриваемом нами случае основной задачей является анализ CDP и последующее построение на основе полученных результатов соответствующего лексико-синтаксического паттерна, который мог бы быть использован в рамках средства извлечения. При этом генерация такого паттерна должна производиться без учета какой-либо предметной специфики текстов, так как заранее неизвестно онтологию какой предметной области или задачи будет наполнять пользователь.

Заметим также, что представленные в каталоге [11] LSOP ориентированы на обработку отдельных предложений для извлечения из них фактов наличия единичного отношения между понятиями. Например, участие объекта в событии или выполнение некоторой роли объектом. Таким LSOP соответствуют простые CDP, включающие не более двух понятий, связанных отношением, например, паттерн «Участие в событии» (CDP Participation) или «Роль объекта» (CDP ObjectRole). В случае же синтезированного пользователем составного CDP, он может включать в виде частей несколько других паттернов. Тогда в нем будет представлено несколько понятий, связанных разными отношениями. Это в свою очередь будет приводить к необходимости поиска связанных по смыслу текстовых фрагментов, описывающих эти понятия, при этом каждый из них может включать несколько предложений.

Данная проблема установления кореферентных связей [19] представляет высокую сложность, так как требует выявления смысловых отношений между лексемами, которые далеко не всегда явно представлены в синтаксической

структуре содержащих их фрагментов текста. Поэтому решение данной проблемы за счет описания таких связей с помощью только лексико-синтаксических паттернов затруднительно и специально в рамках данного этапа исследование не рассматривается. В рамках ее решения в дальнейшем для улучшения результатов применения сформированных ЛСП планируется осуществлять дополнительный анализ текста помимо применения анализатора с набором лексико-синтаксических паттернов и/или производить дополнительную обработку результатов работы последнего с целью определения связанных по смыслу текстовых фрагментов среди найденных.

Еще одной особенностью генерации лексико-синтаксических паттернов является необходимость использования языка их описания, который принят в используемом анализаторе. Он как правило позволяет задавать иерархию правил распознавания лексических конструкций, указывать условия извлечения из них отдельных лексем, а также определять дополнительные ограничения, такие как морфологические признаки лексем и требования их согласования. В данной работе для выполнения анализа текста будет использоваться LSPL-анализатор [20] и соответственно его язык формального описания языковых конструкций для распознавания и извлечения цепочек лексем из текста. Такой выбор обусловлен наличием доступной программной реализации и поддержки со стороны действующего коллектива разработчиков. По сравнению с tomita-парсером, обладающим сходными возможностями, LSPL-анализатор отличает отсутствие фиксированного размера фрагмента текста, сопоставляемого с паттерном на каждом шаге анализа. В tomita-парсере такой размер равен одному предложению, поэтому поиск цепочки лексем, удовлетворяющей некоторому лексико-синтаксическому паттерну, будет производиться только в его рамках. Это не позволяет применять tomita-парсер для выявления фрагментов текста, между лексемами которых могут присутствовать кореферентные связи.

Генерация лексико-синтаксических паттернов для основе заданного CDP

Решение задачи наполнения онтологии заключается в добавлении в нее экземпляров классов и задании отношений между найденными экземплярами. В этом случае некоторый лексико-синтаксический паттерн должен позволять найти в тексте фрагменты, которые можно было бы проинтерпретировать в виде дуг: «экземпляр – отношение –экземпляр».

В рамках различных CDP, такие дуги обычно представляются логическими выражениями (OWL аксиомами) следующего вида:

ClassA SubClassOf relationX some ClassB (1)

ClassA EquivalentTo relationX some ClassB (2)

где ClassA, ClassB – классы онтологии (Ontology classes), relationX – отношение между экземплярами класса ClassA и экземплярами класса ClassB.

Смысл первой аксиомы состоит в том, что если какой-либо экземпляр причисляется к классу ClassA, тогда он должен быть связан хотя бы с одним экземпляром класса ClassB. Вторая аксиома определяет более сильное ограничение, которое состоит в том, что экземпляр причисляется к классу ClassA, тогда и только тогда, когда он связан хотя бы с одним экземпляром класса ClassB.

Данные аксиомы часто применяются для описания структуры CDP, поэтому их можно использовать для формирования основы соответствующего лексико-

синтаксического паттерна. В общем случае основным индикатором фрагмента текста, соответствующего аксиомам рассмотренного вида, является глагол, согласованный по числу и роду с существительным или местоимением, которое (отдельно или в составе именной группы) предшествует ему. Наряду с этим следует учитывать сходную ситуацию с глаголом в страдательном залоге, а также возможность употребления в тексте синонимов лексем. Таким образом, во фрагментах такого вида глагол будет представлять отношение, а существительные или местоимения – экземпляры классов.

Рассмотрим представление лексико-синтаксического паттерна, соответствующей рассмотренным аксиомам (1) и (2) на языке LSPL. Подробное описание синтаксиса языка приведено на сайте разработчиков [20]. Лексико-синтаксический паттерн на языке LSPL записывается в виде расширенных формул Бэкуса-Наура (правил грамматики) и задает лексический состав и поверхностно-синтаксические свойства искомой языковой конструкции для ее распознавания. Дополнительно для каждого лексико-синтаксического паттерна указывается паттерн извлечения, который определяет какие компоненты найденной синтаксической конструкции должны быть извлечены. Таким образом общий вид паттерна на языке LSPL будет следующим:

имя_паттерна =

паттерн_расознавания[=*text*> *паттерны_извлечения_текста*]

Лексико-синтаксические паттерны, соответствующие OWL-аксиоме вида (1) или (2) на языке LSPL будут выглядеть следующим образом:

- BranchAB = ClassA Space Relation Space ClassB =*text*> ClassA Relation ClassB
- BranchBA = ClassB Space RevRelation Space ClassA =*text*>
ClassB RevRelation ClassA
- Space = {W}<0, N> =*text*> Space
- ClassA = {N1 | Pn1 | N1 <synonymA1> | ... | N1 <synonymAn> }<1> =*text*>
N1 Pn1
- ClassB = {N2 | Pn2 | N2 <synonymB1> | ... | N2 <synonymBn> }<1> =*text*>
N2 Pn2
- Relation = {V<synonymRel1> <N1.n=V.n, N1.g=V.g> | ... | V<synonymReln>
<N1.n=V.n, N1.g=V.g> | V<synonymRel1> <Pn1.n=V.n, Pn1.g=V.g> | ... |
V<synonymReln> <Pn1.n=V.n, Pn1.g=V.g>}<1> =*text*> V
- RevRelation = {V<synonymRel1> <N1.n=V.n, N1.g=V.g, V.r=yes> | ... |
V<synonymReln> <N1.n=V.n, N1.g=V.g, V.r=yes> | V<synonymRel1> <Pn1.n=V.n,
Pn1.g=V.g, V.r=yes> | ... | V<synonymReln> <Pn1.n=V.n, Pn1.g=V.g, V.r=yes>}<1>
=*text*> V

Среди приведенных паттернов основными являются паттерны «BranchAB» и «BranchBA». Они задают общую структуру распознаваемых цепочек, состоящих из лексем, которые представляют экземпляры классов и отношение между ними. При этом паттерн «BranchAB» определяет цепочку с глаголом, в активном залоге, а «BranchBA» – в пассивном. Остальные паттерны описывают их отдельные элементы.

Паттерн «Space» соответствует цепочке, состоящей из N-го числа любых лексем. Использование данного шаблона в основных позволяет распознавать те цепочки, в которых между лексемами, представляющими экземпляры классов, и лексемами, представляющими отношения, могут быть предлоги, прилагательные, наречия. Например, это позволяет распознать цепочку «Емкость является нежелательной величиной», где между глаголом «является» и существительным

«величиной» присутствует прилагательное «нежелательной», которое соответствует паттерну «Space». Число N устанавливается равным среднему числу лексем в предложениях текста, так как предполагается, что искомая цепочка будет в пределах одного предложения.

Лексико-синтаксические паттерны «ClassA» и «ClassB» ориентированы на обнаружение лексем, соответствующих вероятным экземплярам классов. Это могут быть существительные (N1, N2), местоимения (Pn1, Pn2) или конкретные синонимы наименований классов (N1 <synonymAn>, N2 <synonymBn>). Набор синонимов при генерации лексико-синтаксического паттерна берется из внешнего словаря. На данный момент используется онлайн словарь от компании «Яндекс» Яндекс.Словарь, доступ к которому осуществляется через его программный интерфейс (API).

Паттерны «Relation» и «RevRelation» позволяют обнаружить в тексте глагол, соответствующий отношению между классами. Он должен быть согласован по роду и числу (условия $N1.n=V.n$ и $N1.g=V.g$) с лексемой (существительным или местоимением), обнаруженной с помощью паттерна «ClassA». Паттерн «RevRelation» обнаруживает цепочки с глаголом в страдательном залоге (условие $V.r=yes$). Паттерны извлечения позволяют получить из распознанных цепочек отдельные лексемы, соответствующие вероятным экземплярам классов и отношениям. Полученный набор лексико-синтаксических паттернов, сохраняется в онтологии того CDP, по которому производилась его генерация.

Программная реализация технологии

Представленная технология была реализована на языке Java в виде настольного приложения – генератора лексико-синтаксических паттернов LSPatGen. Для разработки графического интерфейса использовалась платформа JavaFX. Приложение состоит из 2 основных модулей: модуля генерации паттернов и модуля текстового анализа.

Модуль генерации паттернов включает java-классы, реализующие функции анализа онтологии и выявления подходящих для генерации OWL-аксиом, настройки параметров генерации (корректировка лексем, полученных из онтологии, их лемматизация и формирование списка их синонимов) и формирования лексических паттернов. Лемматизация производится с путем вызова программы *mystem* – лемматизатора от компании Yandex. Список синонимов исходных лексем для включения в формируемый лексико-синтаксический паттерн формируется путем обращения к веб-сервису «Яндекс.Словарь» – <https://dictionary.yandex.net>.

Модуль текстового анализа выполняет функции вызова внешнего лексико-синтаксического анализатора LSPL для анализа указанного пользователем текста с применением сформированного паттерна, обработку полученных результатов и представление найденных текстовых фрагментов.

Графический интерфейс включает 3 оконные формы, соответствующие этапам формирования и использования лексико-синтаксического паттерна. Первая форма отображается при загрузке онтологии и позволяет выбрать аксиомы, которые необходимо использовать при генерации (рис. 1).

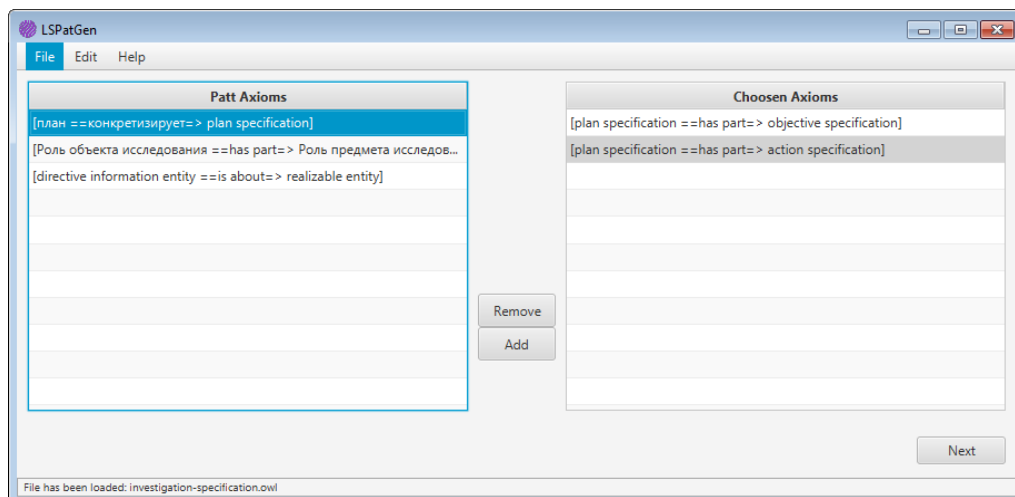


Рис. 1. Форма выбора аксиом для генерации

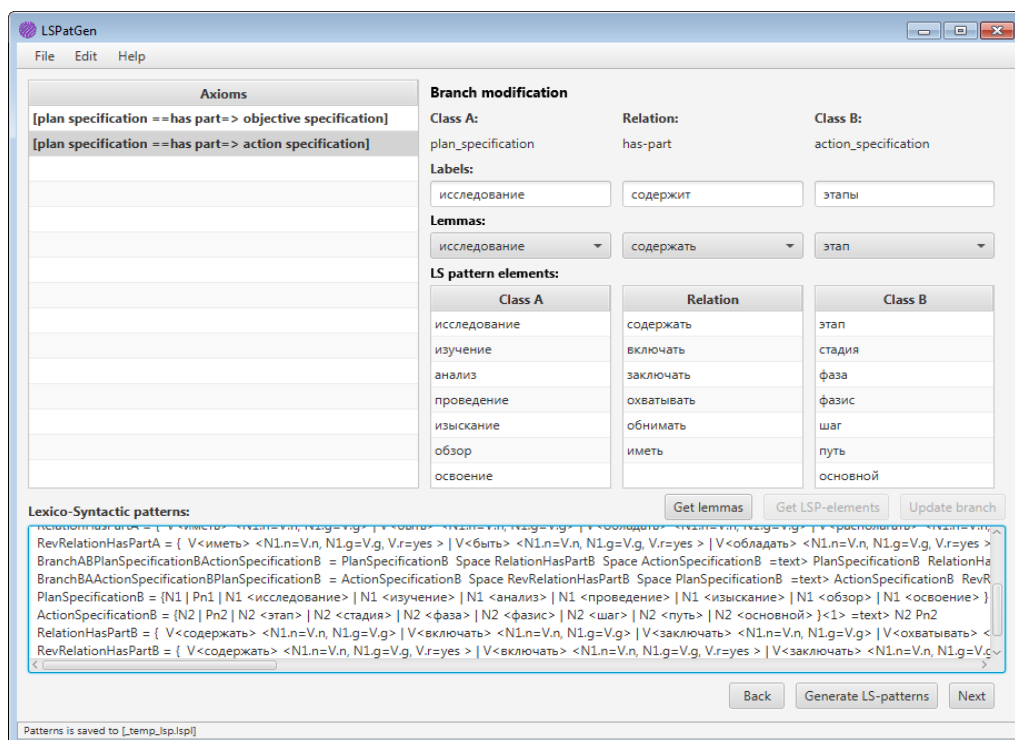


Рис. 2. Настройка параметров генерации паттернов

Далее производится преобразование выбранных аксиом в дуги и определение параметров генерации паттернов для каждой дуги. Это предполагает корректировку лексемы – наименования названия класса (Label), ее лемматизацию, получение и редактирование списка ее синонимов. После чего

производится генерация лексико-синтаксических паттернов и их применение для анализа выбранного пользователем текста (рис. 2, рис. 3)

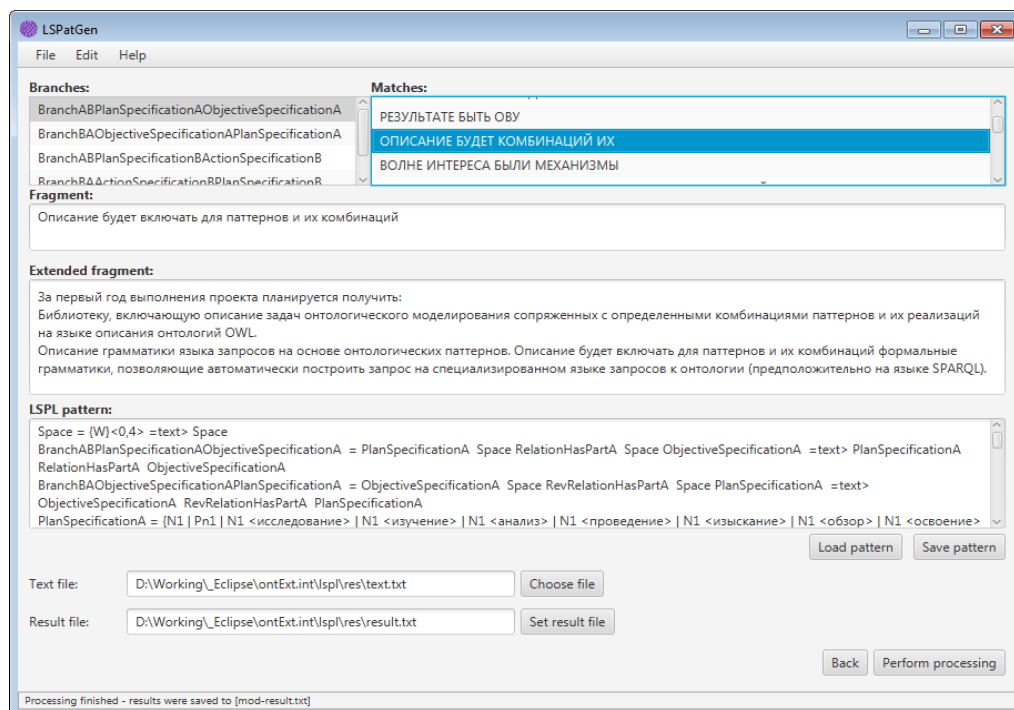


Рис. 3. Анализ текста с помощью полученных паттернов

В результате пользователю для каждого правила (список - Branches) сгенерированного паттерна представляется список найденных цепочек лексем (Matches), а также фрагменты текста, содержащие выбранную цепочку.

Заключение

В данной статье была представлена технология генерации лексико-синтаксических паттернов на основе заданного CDP, а также ее программная реализация в виде настольного приложения. Были рассмотрены основные проблемы формирования лексико-синтаксических паттернов, возникающие в общем случае при последующем использовании полученных паттернов для поиска фрагментов текста, которые могут содержать элементы онтологии. К каковым были отнесены: отсутствие возможности опоры на предметную специфику текста, необходимость обнаружения кореферентных связей между лексемами, различие языков описания паттернов в зависимости от применяемого анализатора.

Наряду с этим был определен общий вид OWL-аксиом, которые могут использоваться для формирования лексико-синтаксических паттернов, было рассмотрено их обобщенное представление в виде паттерна на языке LSPL, а также принципы формирования и использования сгенерированных паттернов с применением разработанного приложения LSPatGen.

Литература

1. Petasis G., Karkaletsis V., et al. Ontology Population and Enrichment: State of the Art. In Knowledge-Driven Multimedia Information Extraction and Ontology Evolution - Bridging the Semantic Gap, Paliouras G., Spyropoulos C.D., Tsatsaronis G., Lecture Notes in Artificial Intelligence, n. 6050, pp. 1-17, Springer-Verlag, 2011.
2. Lehmann J., Voelker J. Introduction to Ontology Learning Perspectives on Ontology Learning, AKA / IOS Press, (2014).
3. Blomqvist E., Hitzler P., Janowicz K., Krisnadhi A., Narock T., Solanki M. (2015). Considerations regarding Ontology Design Patterns. *Semantic Web*, 7 (1), pp. 1-7.
4. Ломов, П.А., Автоматизация синтеза составных онтологических паттернов содержания / П.А. Ломов, // Научный журнал «Онтология проектирования» - Том 6, №2(20)/2016. – Самара: Новая техника, 2016, сс. 162-172 ISSN 2223-9537.
5. Hearst M., Automatic acquisition of hyponyms from large text corpora. In Proceedings of the Fourteenth International Conference on Computational Linguistics, pp. 539--545, Nantes, France, July 1992.
6. Liu K., Chapman W., Savova G., Chute C., Sioutos, N., Crowley R. (2011). Effectiveness of Lexico-Syntactic Pattern Matching for Ontology Enrichment with Clinical Documents. *Methods of Information in Medicine*, 50(5), pp. 397–407.
7. Mukherjea S., Sahay S. Discovering biomedical relations utilizing the World-Wide Web. Proceedings of Pacific Symposium on Biocomputing; Maui, HI. 2006. p. 164-175.
8. Berland M., Charniak, E. Finding parts in very large corpora. Proceedings of the 37th Conference on Computational Linguistics; College Park, MD. 1999. p. 57-64.
9. Sundblad H. Automatic acquisition of hyponyms and meronyms from question corpora. Proceedings of the 15th European Conference on Artificial Intelligence; Lyon, France. 2002.
10. Maynard D., Funk A., Peters W. (2009). Using Lexico-Syntactic Ontology Design Patterns for ontology creation and population. Proceedings of WOP2009, Nov.: CEUR-WS.org.
11. Web portal dedicated to ontology design patterns (ODPs) [Электронный ресурс]. – URL: <http://ontologydesignpatterns.org/>
12. Ефремова Н.Э., Большакова Е.И., Носков А.А., Антонов В.Ю. Терминологический анализ текста на основе лексико-синтаксических шаблонов // Компьютерная лингвистика и интеллектуальные технологии: Труды Международной конференции Диалог'2010. – М.: Изд-во РГГУ, 2010. – С. 124-129.
13. Ермаков А.Е., Плешко В.В., Митюнин В.А. RCO Pattern Extractor: компонент выделения особых объектов в тексте // Информатизация и информационная безопасность правоохранительных органов: XI Международная научная конференция. Сборник трудов. М.: 2003.
14. Жигалов В. А., Жигалов Д.В., Жуков А. А., Кононенко И.С., Соколова Е.Г., Толдова С.Ю. Система Alex как средство для многоцелевой автоматизированной обработки текстов // Труды международного семинара Диалог 2002 «Компьютерная лингвистика и интеллектуальные технологии». М.: Наука, 2002. Т.2. С. 192-208.
15. Томита-парсер – Режим доступа: <http://api.yandex.ru/tomita/>

16. Buitelaar P., Cimiano P., Racioppa S., Siegel M.: Ontology-based Information Extraction with SOBA. In: Proceedings of the International Conference on Language Resources and Evaluation, pp. 2321–2324. ELRA (May 2006)
17. Kim S., Son W., Park B., Park Y., Lee C., Wang H., Jang G., Park G.: OPTIMA: An Ontology Population System. In: 3rd Workshop on Ontology Learning and Population (July 2008)
18. Weber N., Buitelaar P. Web-based Ontology Learning with ISOLDE. In: Proceedings of the Workshop on Web Content Mining with Human Language at the International Semantic Web Conference, USA (2006)
19. Мальковский М. Г., Старостин А. С., Шилов И. А. Метод разрешения местоименной анафоры в процессе синтаксического анализа // Сборник научных трудов SWorld по материалам международной научно-практической конференции. 2013. Т. 11, № 4. С. 41—49.
20. Lexico-Syntactic Pattern Language – Режим доступа: <https://github.com/cmc-msu-ai/lspl>

Сведения об авторах

Ломов Павел Андреевич – к.т.н, старший научный сотрудник
e-mail: lomov@iimm.ru

Pavel A. Lomov – Ph.D., senior researcher

А.Г. Олейник

Институт информатики и математического моделирования ФИЦ КНЦ РАН

**ПРЕДВАРИТЕЛЬНЫЕ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ ПО ПРОЕКТУ
«РАЗРАБОТКА МОДЕЛИ, МЕТОДОВ ИНТЕГРАЦИИ И СИСТЕМНОГО
ИСПОЛЬЗОВАНИЯ МЕЖДИСЦИПЛИНАРНЫХ ЗНАНИЙ И ДАННЫХ ДЛЯ
ПОДДЕРЖКИ УПРАВЛЕНИЯ КОМПЛЕКСНЫМ РАЗВИТИЕМ АЗРФ»***

Аннотация

Представлен обзор результатов первых двух лет работ по проекту. Рассмотрено решение задачи определения области компетенции и структуры онтологии знаний об АЗРФ. Отличительной особенностью разрабатываемых решений является представление онтологии в виде специализированных матрицеподобных структур алгебры кортежей. Это дает возможность рассматривать задачи вывода на онтологии как задачи удовлетворения ограничений. Применение паттернов онтологического проектирования и технологий программирования в ограничениях позволило разработать методы оптимизации выполнения запросов доступа к данным на основе онтологий.

Ключевые слова:

междисциплинарные знания, интеграция знаний, онтологическое моделирование, вывод над онтологией, алгебра кортежей, удовлетворение ограничений

A.G. Oleynik

**PRELIMINARY RESULTS OF RESEARCH ON THE PROJECT «DEVELOPMENT
OF THE MODEL, METHODS OF INTEGRATION AND USE OF
INTERDISCIPLINARY KNOWLEDGE AND DATA FOR MANAGEMENT SUPPORT
OF THE RUSSIAN ARCTIC COMPLEX DEVELOPMENT»**

Abstract

An overview of the results obtained in the first two years of work on the project is presented. The solution of the problem of determining the competence area and the structure of ontology of knowledge about the Russian Arctic is considered. The ontology representation in the form of specialized matrix-like structures of the algebra of tuples is distinctive feature of the solutions being developed is. The use of Ontology Content Design Patterns and constraint programming technologies made it possible to develop methods for optimizing the execution queries in the frame of ontology-based data access.

Keywords:

interdisciplinary knowledge, knowledge integration, ontology modeling, inference on ontology, tuple algebra, constraint satisfaction

Введение

Эффективное обеспечение национальных интересов России в Арктике требует разработки и реализации комплексных, системных решений в области развития данных территорий, учитывающих их специфические особенности. Для выработки таких решений существует задел в виде накопленных знаний и

* Работа выполнена при финансовой поддержке РФФИ (проект №16-07-00562-а)

фактографических данных о различных аспектах существования, развития и взаимодействия локализованных в арктической зоне объектов и систем различной природы. Однако разнородность и разобоченность имеющихся данных и знаний существенно осложняет их совместное использование при выработке системных решений. Это обуславливает актуальность разработки структуры, методов и средств формирования и использования формальной модели, обеспечивающей целостное представление междисциплинарных знаний и данных об Арктической зоне РФ (АЗРФ) – объектах, локализованных в АЗРФ, и процессах, характерных как для отдельных крупных объектов, так и для образуемых этими объектами систем. Реализация формальной модели и средств работы с ней должны обеспечить возможность комплексного использования накопленных данных и междисциплинарных знаний для выработки системных решений в области управления территориями АЗ РФ, анализа и обоснования планов развития этих территорий с учетом влияния разнородных факторов.

Устранить, или, по крайней мере, существенно снизить негативное влияние исходной неоднозначности накопленной информации, а также обеспечить возможность реализации компьютерных методов ее обработки позволяет применение онтологий. На сегодняшний день существует множество различных инструментальных средств и технологий, позволяющих разрабатывать информационные системы и программные продукты, основанные на формальных онтологиях, а также работать с онтологиями непосредственно. Наиболее популярным языком описания веб-онтологий можно назвать язык *Ontology Web Language (OWL)* [1]. В стэндфордском университете создана система *Ontolingua* [2], обеспечивающая интероперабельность создаваемых онтологий. Распространенным инструментом редактирования онтологий является бесплатное программное средство с открытым исходным кодом *Protege*. В качестве альтернативы хранения онтологий в файлах применяют программные средства – *RDF-репозитории* [3] или *RDF-хранилища*.

При решении практических задач возникают ситуации, когда базовых механизмов вывода, предоставляемых существующими средами разработки онтологий, оказывается недостаточно и требуется совместно в единой математической системе использовать различные виды рассуждений. Ситуация осложняется, если параметры модели не полностью определены. В настоящем проекте в качестве единой методологической базы для организации рассуждений на онтологиях используется и развивается математический аппарат алгебры кортежей [4]. Разрабатываемые методы реализуются в рамках парадигмы программирования в ограничениях. При этом существующие в теории удовлетворения ограничений стратегии ускорения поиска дополняются оригинальными методами, основанными на свойствах АК-объектов.

В ходе реализации проекта, наряду с распространенными методами и технологиями, используются обладающих новизной методы и информационные технологии, разработанные и развиваемые с участием членов коллектива исполнителей проекта. К ним, в частности, относятся: модифицированная методика формирования и анализа формализованных концептуальных описаний структуры и задач различных областей деятельности, играющих существенное значение для АЗРФ; технология онтологических паттернов; методы удовлетворения ограничений для реализации процедур вывода на онтологии.

Проект "Разработка модели, методов интеграции и системного использования междисциплинарных знаний и данных для поддержки управления комплексным развитием Арктической зоны Российской Федерации" поддержан грантом РФФИ № 16-07-00562-а и выполняется коллективом в составе: д.т.н. Олейник А. Г. (руководитель), к.т.н. Зуенко А.А., к.т.н. Ломов П. А., Олейник Ю.А., к.т.н. Фридман О.В.

В настоящей работе представлен обзор результатов, полученных в ходе двухлетних работ по решению заявленных в проекте задач. Более детально результаты изложены в публикациях, указанных в списке использованных литературных источников настоящей работы.

Определение области компетенции и разработка структуры онтологии знаний об АЗРФ

С целью определения области компетенции разрабатываемой онтологии проведен анализ основных видов деятельности в АЗРФ. При этом использовались основополагающие документы, характеризующие текущее социально-экономическое состояние территорий АЗРФ и определяющие стратегические планы развития этих территорий. В первую очередь это «Основы государственной политики Российской Федерации в Арктике на период до 2020 г. и дальнейшую перспективу» и «Стратегия развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года». Кроме этого, были проанализированы стратегии развития субъектов и территорий, входящих в АЗРФ согласно Указу Президента РФ № 296 от 2.05.2014г. и ряд отчетов о реализации этих стратегий.

Анализ показал, что перспективы социально-экономического развития территорий АЗРФ традиционно связываются с добычей и переработкой природных ресурсов и, в первую очередь, углеводородного сырья. Важное место в стратегических планах занимает развитие Северного морского пути. Архангельская область выделяется весомым вкладом судостроения в региональную экономику. Определенные надежды в европейской части АЗРФ возлагаются на развитие туризма. Отдельно отмечается необходимость поддержки и развития традиционной деятельности коренных народов Севера. С перспективами реализации «ресурсных» проектов тесно связаны виды деятельности, которые можно условно отнести к инфраструктурным. Это развитие транспортно-логистической сети, финансовой системы, систем кадрового обеспечения, здравоохранения, жилищно-коммунального хозяйства. Специфические условия территорий АЗРФ требуют обязательного включения в рассмотрение деятельности, связанной с охраной окружающей среды и снижением отрицательных экологических последствий всех остальных видов деятельности.

Важным аспектом развития АЗРФ является деятельность, связанная с укреплением обороноспособности. Но в рамках настоящего проекта она не рассматривалась.

Проект ориентирован на разработку средств поддержки управления функционированием и развитием различных подсистем и объектов, локализованных на территории АЗРФ. В качестве объектов управления при решении различных задач может выступать как конкретный физический объект (предприятие, муниципальное образование и т.п.) или процесс (строительство

предприятия, формирование экономического кластера и т.п.), так и сложная пространственно-распределенная подсистема (региональный топливно-энергетический комплекс, северный морской путь и др.). Управление осуществляется на различных уровнях различными органами – от федерального до «объектового». Соответственно инструменты поддержки, создаваемые на основе результатов проекта, должны обеспечивать целостное представление и целевую обработку мультidisциплинарной информации (данных и знаний) в интересах решения задач управления «арктическими объектами». Именно общность объектов исследований и управления определена исполнителями проекта в качестве «связующего звена» интегрируемых в рамках целостной информационной системы гетерогенных знаний и данных. В работах [5, 6] исполнителями проекта представлен подход к формированию структуры мультidisциплинарной информационной системы и некоторые результаты его реализации. При построении системы в проекте используются и развиваются механизмы обеспечения доступа к базам данных на основе онтологии (*Ontology-based data access, OBDA*) [7], что дает возможность воспользоваться преимуществами как отработанных механизмов работы с базами данных, так и методов работы с онтологиями [8].

На основе результатов анализа определена базовая система классов и отношений онтологии междисциплинарных знаний и данных об АЗРФ. Описание онтологической модели приведено в работе [5], где для формируемой структуры используется название «Интегрированное пространство знаний» (ИПЗ). Разрабатываемая комплексная онтология формируется из модулей, определяемых в соответствии с уровнем абстракции их понятий и функционального назначения их содержимого. Каждый модуль представляет собой файл, содержащий отдельную онтологию, описанную с помощью языка веб-онтологий OWL. При построении онтологии использованы - базовая формальная онтология (Basic Formal Ontology, BFO) [9] и ее расширение - онтология информационных артефактов (Information Artifact Ontology, IAO) [10]. При формировании исследовательской версии онтологии знаний об АЗРФ использовалась как информация об исследованиях, проводимых разнопрофильными научными институтами и подразделениями Кольского научного центра РАН, так и представленная в зарубежных информационных ресурсах с открытым доступом. В частности, использован каталог полярных данных (PolarData Catalogue: <https://www.polardata.ca>), данные Arctic Data Ecosystem Map (<http://arcticdc.org/products/data-ecosystem-map>) и данные, представленные на web-сайте геологической службы Финляндии (Fennoscandian Ore Deposit interactive map and Database (FODD): <http://en.gtk.fi/information-services/databases/fodd/>). Сформированная на текущий момент версия онтологии доступна в GIT-репозитории по адресу: <https://github.com/palandlom/ontology-of-integrated-knowledge-space>. Онтология содержит 56 классов, 30 объектных свойств и тезаурус со 116 экземплярами терминов, относящихся к «арктической» деятельности. При формировании тезауруса использовался проект ГОСТ Р «Проведение исследований в полярных регионах. Термины и определения», разработанный АНО НИЦ «Полярная инициатива». Указанный стандарт в 2018 году был рекомендован к утверждению после рассмотрения Техническим комитетом по стандартизации и прохождению экспертизы [11].

Для упрощения разработки и оперирования прикладными онтологиями, ориентированными на использование в рамках создаваемой в проекте единой информационной среды, было принято решение формировать их как совокупность реализаций онтологических паттернов содержания (Ontology Content Design Patterns, CDP) [12]. Это позволит экспертам предметных областей работать не со специфическими языковыми конструкциями (аксиомами OWL, триплетами RDF), а их комбинациями, имеющими описание своего назначения и способов применения. В проекте используются онтологические паттерны содержания, которые описывают возможные варианты представления знаний предметной области в виде фрагментов онтологии, представляющих собой связанные отношения наборы классов онтологии. Для каждого паттерна задается обобщенное определение ситуации (General Use Case, GUC), в которой необходимо его применять, а также набор квалификационных вопросов на естественном языке, определяющие область компетенции паттерна. Каждый из таких вопросов может быть преобразован в запрос к онтологии, ответ на который гарантированно будет получен в случае реализации соответствующего паттерна. При формировании разрабатываемой онтологии были выделены целостные фрагменты понятийной системы ИАО, представляющие концепты. Выделенные фрагменты расширены с учетом организации научной деятельности в РФ и задач применения разрабатываемой онтологии и локализованы в виде отдельных онтологических паттернов содержания. Каждый такой паттерн представляется в виде мини-онтологии в отдельном файле и решает одну задачу онтологического моделирования [5].

При решении практических задач разработчик прикладной онтологии может не обнаружить в системе готового паттерна, полностью отвечающего требованиям конкретной задачи. В рамках проекта разработана технология автоматизированного создания новых паттернов на основе уже имеющихся в онтологии. Создание нового «комбинированного» паттерна не всегда может быть реализовано простым объединением существующих фрагментов онтологий, а требует модификации как их структуры, так и отношений между ними. Технология автоматизированного синтеза составных онтологических паттернов содержания для их последующего применения подробно представлена в работе [13]. В указанной работе описана общая процедура синтеза составного паттерна содержания, рассмотрены способы подбора паттернов – участников синтеза, а также специфические случаи комбинирования распространенных онтологических паттернов содержания с использованием, в частности, процедур «поглощение» и «включение». В качестве примера рассмотрена генерация составного паттерна для представления действия, состав участников которого определяет его вероятную продолжительность. Разработанная технология реализована в виде JAVA-библиотеки для применения в составе программного инструментального средства разработки онтологий на основе онтологических паттернов содержания.

Представление онтологии матрицеподобными структурами алгебры кортежей

Управление большим RDF-хранилищем онтологии является ресурсоемкой задачей с точки зрения физической организации хранилища, проведения индексирования триплетов, а также выполнения запросов. Производительность информационной системы падает с возрастанием объема данных и сложности

запросов к онтологии, включающих конъюнкцию нескольких шаблонов, так как для ответа на запрос перебирается большое число связей триплетов. Использование более сложных способов группировки триплетов в таблицы (cluster-property table) [14] далеко не всегда обеспечивает устранение проблемы снижения производительности при обработке запросов. В качестве варианта решения указанной проблемы в проекте предложено структурировать представление RDF-хранилища в виде специализированных матрицеподобных структур. Такое представление RDF-хранилища позволит сократить расход памяти компьютера для хранения RDF-триплетов, а также повысить эффективность обработки SPARQL-запросов. Для реализации предлагаемого представления используется формальный аппарат описания задач удовлетворения ограничений с помощью матрицеподобных структур. Применение данного аппарата позволяет более эффективно, по сравнению с табличным представлением, описывать, хранить и обрабатывать нечисловые ограничения предметной области.

Нечисловые или качественные ограничения (правила, логические формулы, многоместные отношения) играют важную роль при моделировании многих предметных областей. Для представления и обработки таких ограничений предложено использовать два типа матрицеподобных структур: *C*-системы и *D*-системы. Эти структуры предлагается рассматривать как ограничения над конечными доменами, а рассуждения на данных структурах реализовывать в форме процедур решения задачи удовлетворения ограничений (Constraint satisfaction problems, CSP) [15]. С помощью *C*-систем моделируются дизъюнктивные нормальные формы (ДНФ) конечных предикатов, а с помощью *D*-систем – конъюнктивные нормальные формы (КНФ).

Формирование матрицеподобных структур, соответствующих фрагментам (паттернам) онтологии, осуществляется на основе системы правил отображения. Каждому классу фрагмента онтологии сопоставляется определенная переменная (атрибут), а каждому экземпляру класса определенное значение данной переменной (атрибута). Каждому типу связи онтологии сопоставим бинарное ограничение (матрицу ограничений). Каждая строка матрицы ограничений описывает некоторое множество дуг между экземплярами соответствующих классов онтологии.

Представление онтологии *C*- и *D*-системами позволяет решать задачи ускорения логического вывода и обеспечения семантической корректности онтологии в рамках парадигмы программирования в ограничениях, используя накопленный в этой области опыт решения задач комбинаторного поиска. Учет дополнительных (нелогических) ограничений предлагается выполнять в виде, так называемых, глобальных ограничений, для которых в теории удовлетворения ограничений имеются специализированные эффективные алгоритмы обработки [15].

Хранение исходных данных и результатов вычислений над матрицеподобными структурами осуществляется в формате XML (Extensible Markup Language). Для корректного представления матрицеподобных структур разработана специальная структура XML-файлов.

Корневой элемент XML-документа – *ntDoc* – содержит элементы для хранения доменов, атрибутов, схем отношений и систем.

Элемент для хранения доменов *ntDomain* содержит два атрибута. Атрибут *name* служит для обозначения имени домена, а атрибут *type* указывает тип данных домена и может принимать одно из следующих значений: “*string*” для строкового типа данных, “*int*” – для целочисленного, “*float*” – для числа с плавающей запятой. Внутри *ntDomain* размещаются элементы *element*, которые содержат значения, принадлежащие домену.

Элемент *ntAttribute* содержит имя атрибута *name* и имя соответствующего ему домена *domain*.

Элемент *ntScheme* имеет атрибут *name* для обозначения имени схемы и содержит элементы *element*, в каждом из которых указывается атрибут, входящий в схему.

Элемент *ntSystem* имеет более сложную структуру. Атрибут *name* указывает имя системы. *Scheme* содержит имя схемы для описываемой системы. *Type* может принимать значение “*c*” для *C*-систем или “*d*” для *D*-систем.

Каждая система состоит из элементов *nt*, служащих для представления кортежей. Каждый из элементов *nt* содержит элементы *component*, обозначающие компоненты. Каждая компонента содержит элементы *element*, указывающие, какие элементы домена входят в данную компоненту.

Для различных целей XML документы, содержащие описания матрицеподобных структур, могут преобразовываться в другой формат. Например, в HTML-документ с использованием стилей CSS (Cascading Style Sheets) для отображения.

В разрабатываемой системе предусматриваются следующие возможности: сохранение документа в долговременной памяти в виде дерева XML с использованием XMLDoc; загрузка документа из XML-файла; визуализация документа в HTML-странице с использованием стилей для отображения в браузере.

RDF-формат представления онтологии является подмножеством XML. Поэтому сохранение онтологии в этом формате (файле с расширением «*.owl»), позволяет оперировать с представленными данными и знаниями как с XML-документом. В проекте для преобразования онтологии, созданной в редакторе Protégé и сохраненной в RDF-формате, в XML-файл, содержащий описание онтологии уже в виде совокупности матрицеподобных структур, были сформулированы специализированные правила на языке XSLT (eXtensible Style Language for Transformations).

Основными тегами RDF-формата, используемыми для представления онтологии, являются: тег <owl:Class> - служит для описания некоторого класса онтологии; тег <owl:Restriction> - служит для описания связей между классами, является вложенным тегом в тег <owl:Class>, который выступает в качестве источника связи; <owl:onClass> - является вложенным тегом тега <owl:Restriction> и описывает класс-приемник для данной связи, тег <owl:NamedIndividual> - позволяет задать конкретные экземпляры классов, и кортежи отношений. В основе предложенного алгоритма преобразования исходного XML-файла онтологии в требуемый XML-файл описания соответствующих онтологии матрицеподобных структур лежит набор правил трансформации, устанавливающих соответствия между тегами исходного и результирующего файлов.

Использование контекстно-ориентированного метода управлению ограничениями

Для повышения релевантности извлекаемой информации и скорости обработки запросов конечного пользователя при решении задачи интеграции разнородных знаний и данных об Арктической зоне Российской Федерации исполнителями проекта развит контекстно-ориентированный метод к управлению ограничениями. Данный метод представляет собой модификацию ранее разработанных методов, ориентированных на повышение эффективности обработки информации.

В работе с создаваемой информационной системой участвуют следующие объекты: пользователь, прикладная программа (приложение), запрос приложения и задача. Каждый из перечисленных типов объектов характеризуется собственным контекстом. Контекст взаимодействия складывается из контекстов взаимодействующих объектов, а каждый из них, в свою очередь, подразделяется на абстрактный и прикладной контексты. Абстрактный контекст объекта соответствует интенционалу этого объекта. Прикладной контекст объекта позволяет описать некоторое множество его экземпляров. Контексты объектов взаимодействия формализуются с использованием технологии шаблонов онтологического проектирования (ODP), а также авторских разработок в области представления и обработки нечисловых ограничений предметной области в виде специализированных матрицеподобных структур.

Абстрактный контекст задачи декларативно описывается в виде отношений онтологии, формализующей задачу исследования. Прикладной контекст исследования задается посредством приписывания классам онтологии конкретных экземпляров. Для облегчения генерации абстрактного контекста задачи (того или иного исследования) в проекте разработаны такие шаблоны онтологического проектирования, как: «Информационная сущность», «Планируемый процесс», «Исследование», «Предмет исследования», «План исследования», «Цели и задачи исследования», «Метод исследования», «Публикация по исследованию».

При работе с онтологией для каждого квалификационного вопроса пользователя формируется соответствующий запрос на специализированном языке SPARQL. Для выполнения процедур анализа SPARQL-запрос должен быть разложен в набор конъюнктивных запросов. Абстрактный контекст конъюнктивного запроса описывается тройками-триплетами, которые соответствуют связям в графе онтологии. Прикладной контекст запроса получается из абстрактного после приписывания классам конкретных экземпляров, значений (где это необходимо). При представлении контекстов SPARQL-запроса с помощью предложенных матрицеподобных структур каждому триплету вида «субъект X – свойство – объект Y », где в качестве субъекта и объекта используются имена, сопоставляется нечисловая матрица $R[XY]$, а каждому триплету, содержащему в позициях объекта и/или субъекта конкретные значения, сопоставляется определенный кортеж матрицы $R[XY]$.

Контекст приложения описывает возможности взаимодействующей с онтологией прикладной программы. Абстрактный контекст приложения – это список типов выполняемых приложением запросов (шаблонов запросов). Прикладной контекст приложения характеризует конкретный экземпляр программного модуля и ссылается на прикладные контексты генерируемых запросов.

Контекст пользователя также подразделяется на абстрактный и прикладной. Абстрактный контекст позволяет описать в виде специализированной онтологии некий тип пользователя, а прикладной контекст сопоставляется с конкретным человеком, работающим в рассматриваемой информационной системе, и его полномочиями (реализуемыми функциями) в системе [6].

С точки зрения решения задач проекта контекстно-ориентированное управление ограничениями по мнению авторов обладает следующими достоинствами:

- Возможность ставить и решать задачи интеллектуальной обработки данных и знаний в виде задач удовлетворения ограничений с применением современных библиотек и сред программирования в ограничениях.
- Возможность активировать только те контекстные ограничения, которые актуальны для интересующего исследования (задачи). Это позволяет гибко перенастраивать и оперативно анализировать ограничения, специфичные для конкретной предметной области, способствуя уменьшению трудоемкости решаемых задач.
- На основе анализа контекстов на этапе построения онтологии обеспечивается более детальный контроль корректности ее структуры. В результате удается автоматизировать проверки с точки зрения широкого набора ограничений.
- При совместной обработке пользовательских SPARQL-запросов и контекстных ограничений отслеживаются некорректные обращения к онтологии, что позволяет выявлять некорректности в структуре запроса еще до исполнения запроса и тем самым повысить быстродействие всей разрабатываемой информационной системы.

Применение паттернов онтологического проектирования для оптимизации выполнения запросов доступа к данным

Подход к построению систем, основанных на знаниях, подразумевающий комбинирование онтологии, представляющей понятийную систему предметной области, и реляционной базы данных, используемой для хранения данных, соответствующих заданным в онтологии понятиям получил название *Ontology-based data access (OBDA)* [7,8]. В рамках указанного подхода онтология используется в качестве основы для формирования пользовательских запросов, а база данных для получения результатов их выполнения. Это позволяет обеспечить скорость выполнения онтологического запроса сходную со скоростью выполнения запроса в рамках традиционного подхода с применением только базы данных. Повысить производительность базы данных можно за счет предварительного построения индексов для ее таблиц. Однако формирование эффективных индексов требует априорного знания перечня запросов, которые будут выполняться над базой данных.

Как было указано выше в данной статье, разрабатываемая онтология формируется с использованием онтологических паттернов содержания (CDP). С каждым паттерном связан набор квалификационных вопросов, которые формулируются на естественном языке. Предполагается, что именно ответы на данные вопросы будут преимущественно интересовать пользователя при работе с информационной системой. На основе этого предположения в проекте

разработана технология проведения априорной индексации базы данных, что обеспечивает ускорение отклика системы на запросы пользователя. При реализации информационной системы на основе OBDA подхода для каждого из включенных в онтологию паттернов определяется фрагмент схемы базы данных. Принцип соотнесения элементов онтологии и схемы данных соответствует рекомендациям консорциума W3C [16]. В качестве программной среды для реализации подхода OBDA был использован Ontop [17]. Данное программное средство реализует трансляцию онтологического запроса (SPARQL-запроса) в SQL-запрос над базой данных, а также представление результатов выполнения SQL-запроса в виде набора связанных экземпляров онтологии пользователю [18].

Разработанная технология применения паттернов онтологического проектирования для оптимизации запросов доступа к базе данных более подробно представлена исполнителями проекта в работе [19]. Для каждого паттерна формируемой онтологии разрабатывается набор правил мэппинга в формате Ontop для переписывания исходного запроса SPARQL-запроса в SQL-запрос. Каждое правило соответствует одной таблице реляционной базы данных и включает компоненты: source - SQL-запрос и target - множество шаблонов для генерации выражений АBox, представляющие собой конкретные экземпляры классов онтологии, из результатов выполнения SQL-запроса.

В качестве примера в работе [19] рассмотрен паттерн «Спецификация исследования» (*Investigation specification*), который позволяет определить такие «части» спецификации исследования, как: объект; предмет; цель; действия/методы. Приведен набор квалификационных вопросов данного паттерна, соответствующий паттерну фрагмент схемы реляционной базы данных, включающий 9 таблиц, и пример правила мэппинга.

Далее в [19] рассмотрено формирование индексов базы данных на основе квалификационного вопроса: «Какой метод планировалось применять для исследования данного объекта?» для паттерна «Спецификация исследования». Экспериментальная оценка применения технологии априорной индексации базы данных на основе квалификационных вопросов онтологических паттернов содержания проводилась на тестовом наборе данных (около 10 000 000 связанных экземпляров онтологии, представленных в 9 таблицах реляционной базы данных по 1000 000 строк в каждой). Эксперименты, заключающиеся в выполнении серий SPARQL-запросов с помощью системы Ontop, показали, что среднее время выполнения запроса после индексации уменьшается в 7-8 раз.

Обработка SPARQL-запросов к онтологии с использованием программирования в ограничениях

Представление RDF-триплетов онтологии в виде специализированных матрицеподобных структур (С-систем) дает возможность рассматривать задачи вывода на онтологии как задачи удовлетворения нечисловых ограничений (CSP). Среди разработанных алгоритмов удовлетворения нечисловых ограничений особый интерес представляют алгоритмы распространения ограничений, которые позволяют преобразовать за полиномиальное время описание задачи CSP к более простому виду с сохранением всех решений, то есть сократить пространство поиска. Для случая, когда задача CSP выражена в виде совокупности С-систем,

сформулированы правила редукции, осуществляющие упомянутые преобразования [20].

Основная идея разработанного метода обработки SPARQL-запросов к онтологии состоит в том, чтобы анализу подвергается не вся исходная CSP целиком, а выбираются отдельные матрицы, которые соответствуют унарным и бинарным ограничениям (ограничениям на одну или две переменные). Анализ таких матриц позволяет обрабатывать меньшее число строк и/или столбцов, чем в исходной задаче. При этом выявление «лишних» значений в доменах атрибутов какой-либо матрицы влечет удаление этих же значений из доменов атрибутов всей задачи CSP, то есть влияет на все ограничения. В результате могут появляться новые унарные и бинарные ограничения, к которым вновь применяются алгоритмы распространения ограничений.

При распространении ограничений сначала обрабатываются унарные ограничения, поскольку с их помощью можно непосредственно «сузить» домены переменных. Затем активируются бинарные ограничения, поскольку их обработка более трудоемка. Активация ограничений осуществляется по событиям. К таким событиям, в частности, относится «усечение» домена.

Типовой SPARQL-запрос на извлечение данных, как правило, представляет собой длинную связку (join) шаблонов триплетов, включающих различные и труднопредсказуемые комбинации предикатов, что негативно влияет на скорость его выполнения. При больших объемах RDF-хранилищ применение разработанных методов удовлетворения ограничений позволяет ускорить обработку SPARQL-запросов по сравнению с традиционным подходом, базирующимся на динамическом программировании. Это обусловлено тем, что динамическое программирование предполагает реализацию стратегии поиска в «ширину», а методы удовлетворения ограничений используют поиск в «глубину».

Применение методов распространения ограничений, разработанных ранее одним из исполнителей проекта, для ускорения обработки SPARQL-запросов к онтологиям рассматривается в работе [21]. В качестве примера представлено формирование ответа на один из квалификационных вопросов онтологического паттерна содержания для понятия «Процесс исследования» (Investigation). Описаны ограничения, сопоставляемые каждой связи онтологии, которая фигурирует в рассматриваемом SPARQL-запросе, и их представление в виде C- систем с учетом введенных конкретных экземпляров классов. Описанная далее в работе [21] пошаговая схема процедуры распространения ограничений наглядно иллюстрирует механизм редукции пространства поиска в процессе вывода решения CSP.

Количественная оценка эффективности разработанных процедур распространения ограничений

Как показали исследования авторов [22], обработка качественных зависимостей, представленных в виде логических выражений и правил, в современных средах программирования в ограничениях недостаточно эффективна, и не может быть выполнена за приемлемое время даже при сравнительно небольшой размерности задачи. Отличие качественных (нечисловых) ограничений от числовых состоит в том, что переменные даже в рамках одного ограничения могут иметь различные конечные области определения, а сами ограничения, как правило, не являются бинарными.

Для реализации авторских алгоритмов вывода на ограничениях, представленных в виде специализированных матрицеподобных структур (С-систем), была выбрана библиотека Choco, которая является бесплатным программным обеспечением с открытым кодом, созданным для описания и решения задач удовлетворения ограничений [23]. На основе библиотеки Choco, был разработан набор классов, расширяющий функционал основной библиотеки для представления и решения задач CSP в виде набора С-систем. В отличие от типовых ограничений библиотеки Choco и их распространителей, работающих с присваиваниями переменных, распространитель на С-системах работает с компонентами системы, которые могут содержать несколько значений. В процессе распространения он упрощает саму С-систему, с каждой итерацией уменьшая объем обрабатываемой информации и увеличивая, тем самым, скорость нахождения решения.

Проведенные исследования, представленные в работе [22], продемонстрировали, что применение предложенного подхода к обработке качественных ограничений, частным случаем которых являются связи онтологии, дает существенный выигрыш времени по сравнению со встроенными в библиотеку Choco алгоритмами распространения качественных ограничений. В качестве тестовой была выбрана задача N-Queens, ввиду удобства ее масштабирования. Кроме того, данная задача также моделируется в виде совокупности бинарных С-систем, как и исследуемые онтологии. При решении задачи N-Queens, когда требовалось найти все ее решения с лимитом по времени в две минуты, реализуемый в рамках проекта подход позволил обработать пространство поиска размерности 12 в 12 степени. При этом, использование стандартных средств Choco не дало возможности исследовать пространство поиска более 9 в 9 степени. В случае, когда ставилась задача, найти хотя бы одно решение за фиксированное время (две минуты), стандартные средства Choco переставали укладываться в отведенное время уже при размерности пространства поиска, оцениваемого числом 19 в 19 степени. Предлагаемые же авторами методы справились с пространством поиска размерности 76 в 76 степени.

Заключение

Созданная и развиваемая исследовательская версия междисциплинарной онтологии Арктической зоны РФ обеспечивает возможности отработки методов, технологий и средств манипулирования формализованными междисциплинарными знаниями. При разработке онтологии использована технология онтологических паттернов. В развитии данной технологии исполнителями разработана технология автоматизированной генерации составного паттерна по требованиям пользователя. Использование наборов определяемых для паттернов квалификационных вопросов положено в основу разработанных механизмов оптимизации выполнения запросов доступа к данным в OBDA – системах.

Отличительной особенностью разрабатываемых в проекте решений является представление концептов и связей онтологии в виде матрицеподобных структур алгебры кортежей. Это дает возможность рассматривать задачи вывода на онтологии как задачи удовлетворения ограничений над конечными доменами. В настоящее время существует большое количество инструментальных сред программирования в ограничениях, однако, стандартов для подобного рода программных продуктов пока не существует, что свидетельствует об

актуальности и новизне исследований в рамках данного подхода, а также об актуальности поиска новых методов и способов программного представления ограничений. В проекте особое внимание уделяется работе с качественными ограничениями. Именно такой тип ограничений свойственен сложным системам, к которым, в частности, относятся и разнородные системы, локализованные в АЗРФ. Отличие качественных ограничений от числовых состоит в том, что переменные даже в рамках одного ограничения могут иметь различные области определения. Существующие на данный момент в теории удовлетворения ограничений методы, основанные на табличном представлении нечисловых ограничений, являются недостаточно эффективными для работы с ограничениями такого типа. Разрабатываемые методы и алгоритмы позволяют расширить область применения технологии программирования в ограничениях для организации процедур вывода на онтологиях. В частности, реализующий методологию программирования в ограничениях метод обработки SPARQL-запросов к онтологии междисциплинарных знаний обеспечивает поэтапное снижение размерности пространства поиска решения. При больших объемах RDF-хранилищ компактное представление онтологии в виде C-систем в сочетании с реализацией разработанных алгоритмов распространения ограничений позволяет значительно ускорить обработку запроса по сравнению с «традиционной» реализацией обработки на основе динамического программирования.

Литература

1. Web Ontology Language (OWL). – Режим доступа: <https://www.w3.org/OWL>
2. Ontolingua. – Режим доступа: <http://www-ksl.stanford.edu/software/ontolingua>
3. Resource Description Framework (RDF). – Режим доступа: <https://www.w3.org/RDF>
4. Кулик, Б. А. Алгебраический подход к интеллектуальной обработке данных и знаний / Б. А. Кулик, А. А. Зуенко, А. Я. Фридман - СПб.: Изд-во Политехнического ун-та, 2010. – 235 с.
5. Олейник, А.Г. Разработка онтологии интегрированного пространства знаний / А.Г. Олейник, П.А. Ломов // Онтология проектирования. - 2016. - Т. 6, №4(22). - С. 465-474.
6. Oleynik Andrey. Solutions for System Analysis and Information Support of the Various Activities in the Arctic / Andrey Oleynik, Pavel Lomov, Alexey Shemyakin, Alexey Avdeev // Czech Polar Reports, 2017. V. 7, № 2. - P. 271-279.
7. Ontology-Based Data Access: A Data Management Paradigm – Режим доступа: <http://www.obdasystems.com/obda>.
8. Когаловский, М.Р. Системы доступа к данным, основанные на онтологиях / М.Р. Когаловский // Программирование. 2012, №4. - С. 55-77.
9. Arp, R. Building Ontologies with Basic Formal Ontology/ Arp, R., Smith, B. Spear, A // The Mit Press, 2015 — Режим доступа: <https://mitpress.mit.edu/books/building-ontologies-basic-formal-ontology>
10. Smith, B. Aboutness: Towards Foundations for the Information Artifact Ontology / C. Werner, B. Smith, C. Werner // In Proceedings of the Sixth International Conference on Biomedical Ontology (ICBO). CEUR Vol. 1515 1-5. – Режим доступа: <http://ceur-ws.org/Vol-1515/regular10.pdf>

- 11.ГОСТ Р Проведение исследований в полярных регионах. Термины определения - Режим доступа: <http://tk187.ru/gostr1>
12. Gangemi, A. *Ontology Design Patterns.*/ Aldo Gangemi, Valentina Presutti // In: *Handbook on Ontologies. International Handbooks on Information Systems.* Staab S., Studer R. (eds). Springer, Berlin, Heidelberg, 2009. - Режим доступа: https://doi.org/10.1007/978-3-540-92673-3_10.
13. Ломов, П.А. Автоматизация синтеза составных онтологических паттернов содержания / П.А. Ломов // *Онтология проектирования.* 2016. Т. 6, №2(20). - С. 162-172.
14. Yuksel, A.S. *An Analysis of RDF Storage Models and Query Optimization Techniques/* Asim Sinan Yuksel, Ibrahim Arda Cankaya, Mehmet Erkan Yuksel// *International Journal of Information Engineering and Electronic Business(IJIEEB).* Vol. 7, No. 2, Mar. 2015. - P. 20-26.
15. Щербина О.А. Удовлетворение ограничений и программирование в ограничениях / О.А. Щербина // *Интеллектуальные системы.* 2011. Т.15, N 1-4. - С. 53-170.
16. *A Direct Mapping of Relational Data to RDF* - Режим доступа: <https://www.w3.org/TR/2012/REC-rdb-direct-mapping-20120927>
17. *Ontop*- Режим доступа: <https://ontop.inf.unibz.it>
18. Calvanese, D. *Ontop: Answering SPARQL queries over relational databases /* D. Calvanese, B. Cogrel, S. Komla-Ebri, R. Kontchakov, D. Lanti, M. Rezk, M. Rodriguez-Muro, G. Xiao // *Semantic Web,* 2017, Vol. 8, No. 3. - P. 471-487.
19. Ломов, Л.А. Технология применения паттернов онтологического проектирования для оптимизации выполнения запросов в системах обеспечения доступа к данным на основе онтологии / П.А. Ломов, А.Г. Олейник // *Онтология проектирования.* - 2017. - Т. 7, №4(26). - С. 443-452.
20. Зуенко А.А. Качественное моделирование технических систем на основе методов распространения ограничений / А.А. Зуенко. // *Открытые семантические технологии проектирования интеллектуальных систем // Open Semantic Technologies for Intelligent Systems (OSTIS-2016): мат. VI Междунар. научн.-техн. конф., г. Минск, 18-20 февраля 2016 г. – Минск: БГУИР, 2016. – С.573-578.*
21. Зуенко А.А. Применение методов распространения ограничений для ускорения обработки запросов к онтологиям/ А.А. Зуенко, П.А. Ломов, А.Г. Олейник // *Труды СПИИРАН.* 2017. № 1(50) - С. 112-136.
22. Zuenko, A. *Programming of Algorithms of Matrix-Represented Constraints Satisfaction by Means of Choco Library /* A. Zuenko, Y. Oleynik // *Proceedings of the Third International Scientific Conference “Intelligent Information Technologies for Industry” (ITI’18), Advances in Intelligent Systems and Computing.* Springer Nature Switzerland AG. Abraham et al. (Eds.): ITI 2018, - P. 1–10.
23. *Choco Documentation* – Режим доступа: <http://www.choco-solver.org>

Сведения об авторе

Олейник Андрей Григорьевич – д.т.н., главный научный сотрудник

e-mail: oleynik@iimm.ru

Andrey G. Oleynik – D. of Sci. (Eng.), chief researcher

УДК 004.9
DOI: 10.25702/KSC.2307-5252.2018.10.144-150

С.Ю. Яковлев

Институт информатики и математического моделирования ФИЦ КНЦ РАН

ОСОБЕННОСТИ РАЗВИТИЯ АРКТИЧЕСКОЙ ЗОНЫ РФ (ПО МАТЕРИАЛАМ КОНФЕРЕНЦИИ OMR 2018 И КОНКУРСА АРКТИКА 2018)*

Аннотация

В статье дан обзор состояния дел в сфере комплексного развития Арктической зоны Российской Федерации. Особое внимание уделено вопросам освоения Северного морского пути и добычи углеводородного сырья. Обзор выполнен на основе опыта участия в конференции Offshore Marintec Russia 2018 и конкурсе Арктика 2018.

Ключевые слова:

Арктическая зона Российской Федерации, Северный морской путь, углеводородное сырьё

S.Yu. Yakovlev

DEVELOPMENT FEATURES OF THE RUSSIAN FEDERATION ARCTIC ZONE (ON THE MATERIALS OF THE OMR 2018 CONFERENCE AND THE ARCTIC 2018 COMPETITION)

Abstract

The article provides an overview of the current status in the field of integrated development of the Russian Federation Arctic zone. Special attention is paid to the Northern Sea Route development and the hydrocarbon feedstock production. The review is based on the experience of participation in the Offshore Marintec Russia 2018 conference and the Arctic 2018 competition.

Keywords:

Russian Federation Arctic zone, Northern Sea Route, hydrocarbon feedstock

Введение

В начале октября 2018г. в Санкт-Петербурге состоялась Международная выставка - конференция по судостроению и разработке высокотехнологичного оборудования для освоения континентального шельфа Offshore Marintec Russia (OMR 2018). В рамках конференции на пленарном заседании были подведены итоги Международного конкурса научных, научно-технических и инновационных разработок, направленных на развитие и освоение Арктики и континентального шельфа (Арктика 2018). По итогам конкурса выпущен сборник [1].

Выступления на конференции, материалы конкурса позволяют получить обзорное представление о современном этапе, планах и перспективах развития Арктической зоны РФ (АЗРФ).

* Работа выполнена при финансовой поддержке РФФИ (проект № 18-07-00167-а).

1. Севморпуть – основа комплексного развития арктической зоны Российской Федерации (пленарное заседание OMR 2018)

Во многих выступлениях подчеркивалось ключевое значение арктического шельфа для развития нефтегазовой и транспортной отраслей РФ.

В докладе заместителя Министра природных ресурсов и экологии РФ дана характеристика минерально-сырьевым ресурсам региона и грузообороту Северного морского пути (СМП). Основные разведанные запасы приходятся на газ, далее по объёму идут нефть, конденсат, уголь, медь, никель. Отмечалось перспективное значение Ямала, Гыданского полуострова, Обской губы, Таймыра.

Во многих докладах повторялось целевое положение из майских указов В.В. Путина о необходимости доведения к 2024 году грузопотока по СМП до 80 млн. тонн в год (сейчас он составляет около 7 млн. тонн). Также неоднократно акцентировалась возрастающая роль сжиженного природного газа (СПГ), констатировалась необходимость кооперации усилий различных отраслей и организаций.

В выступлении представителя Министерства энергетики РФ дан прогноз: к 2035 году доля углеводородного сырья в минерально-сырьевых ресурсах АЗРФ останется примерно на сегодняшнем уровне, при этом доля газа будет расти. Намечена большая программа по геологоразведке и бурению, в основном это относится к Баренцеву и Карскому морям. Основными компаниями в Арктической зоне являются Роснефть и Газпром, они имеют свои лицензионные участки. Ранее была принята программа развития шельфа, предусматривающая налоговые льготы для работ по освоению Арктики. В перспективе, к 2040 году, планируется довести грузопоток по СМП до 200 млн. тонн. Также намечена обширная программа создания и развития инфраструктуры, построения заводов, терминалов, в том числе в Мурманской области (компании Роснефть, НОВАТЭК).

Неоднократно подчёркивалась необходимость создания и обновления нормативной и законодательной базы по всем направлениям развития российской Арктики. В настоящее время появляется примерно 5-6 документов в год, что явно недостаточно.

В докладе вице-президента ПАО «НК «Роснефть» большое внимание было уделено судостроительному комплексу «Звезда» (г. Большой Камень Приморского края) как ключевому звену развития СМП. Комплекс создаётся в период 2012 – 2024 гг. и ориентирован на самые современные цифровые технологии крупноблочного, крупнотоннажного арктического судостроения. В задачи суперверфи войдет строительство судов любой сложности, характеристик и назначений, включая ранее не производимые (в связи с отсутствием необходимого оборудования) в РФ модели. Основу продуктовой линейки «Звезды» составят средне- и крупнотоннажные коммерческие суда для транспортировки грузов, морские буровые разведочные и добычные платформы, суда обслуживающего флота, в первую очередь ледового класса. Важнейшими элементами производственной программы будут специальные танкеры и суда-газовозы для транспортировки углеводородов. На «Звезде» уже сейчас функционируют и вводятся в строй уникальные сооружения: кран «Голиаф» грузоподъёмностью 1200 тонн - один из крупнейших в мире по своим размерам (высота 139 метров, длина поперечной балки 230 метров), сейчас максимальная грузоподъёмность кранов на существующих российских верфях не превышает

500 тонн; транспортно-передаточный док грузоподъемностью 40000 тонн, длиной более 300 м, шириной до 50 м; тяжёлый достроечный стапель; самый большой сухой док в России размерами 485x114x14,5 м, в настоящее время в России имеются два сухих дока – Петровский в Кронштадте и док в Гремихе, к 2020 году в Мурманске планируется построить ещё два дока.

Отмечались суровые климатические особенности эксплуатации СМП (9 месяцев в году акватория СМП покрыта льдом). До развала СССР максимальный грузопоток по СМП был достигнут в 1987 году – 6.6 млн тонн.

В настоящее время большое внимание уделяется вопросам разведки новых месторождений, охраны окружающей среды. Самая северная скважина – «Университетская-1», 74°с.ш., 250 км от материка. Требуются новые суда различного назначения, перспективен переход на газомоторное топливо.

В выступлении представителя дирекции СМП Госкорпорации «Росатом» дан исторический обзор освоения СМП (первое упоминание о СМП датируется 1525 годом), обзор текущего состояния ледокольного обеспечения. Отмечено, что дизель-электроходы уже не обеспечивают потребностей. Грузопоток по СМП рос до 1986 года, потом упал, сейчас вновь растёт и дошёл до уровня 1986 года. Достраивается ледокол «Арктика», запущен проект «Лидер» оснащения ледокольного флота, ставится задача выхода на рынки азиатско-тихоокеанского региона. Принята целевая расстановка арктического ледокольного флота. Необходимо продлить ресурсы действующих ледоколов. Первоочередными задачами являются: освоение восточной части СМП, создание базы в Певеке, более активное привлечение компании Росатом. Также необходимо ускорить проекты, связанные с добычей и транспортировкой СПГ.

АО «ОСК» представило общую картину состояния арктического судостроения. Основные проектные организации: ЦКБ «Айсберг», ФГУП «Крыловский ГИИ», ЦКБ МТ «Рубин». Дана характеристика строящихся и проектируемых ледоколов, судов снабжения, плавучих энергоблоков, подлёдных (подводных) сооружений. Отмечена необходимость развития полярного экспедиционного туризма. Так, в Аляске этот бизнес приносит около 3 млрд. долларов в год, в Канаде – 6.5 млрд. долларов. Функционируют круизные суда ледового класса. В РФ это направление находится на этапе проектирования.

Освещена деятельность ПАО «Газпром». Представлено распределение суммарных минеральных ресурсов по акваториям: 44% - Карское море, 26% - Баренцево, более 80% приходится на северные моря. Дана карта участков суши и моря, на которые Газпрому выданы лицензии. Вводятся в эксплуатацию скважины. Сейчас флот, который эксплуатируется Газпромом, насчитывает 35 единиц (включая фрахт). Необходимо привлекать ледокольный флот. В 2015 году заключено соглашение с ФГУП «Атомфлот» о безопасной и эффективной транспортной деятельности. В 2012 году танкер-газовоз «Река Обь» успешно завершил первую в мире перевозку СПГ по СМП (из Норвегии в Японию). Рейс сопровождался тремя атомными ледоколами, занял менее 10 суток, всего было пройдено более 2600 морских миль. Отмечена важность совершенствования космических систем и технологий для СМП, наращивания вертолётного и аварийно-спасательного обеспечения.

ПАО «Газпром нефть» представило проекты «Новый порт», «Ворота Арктики».

Представители Архангельской и Мурманской областей дали характеристику территорий с точки зрения развития СМП. Описывались как созданные, так и проектируемые объекты, промышленность, инфраструктура, образование и наука.

2. Сборник работ лауреатов конкурса Арктика 2018

Предваряет сборник статья Заместителя Министра энергетики РФ Сорокина П.Ю. «Место углеводородных ресурсов континентального шельфа Арктики и Дальнего Востока в энергетической стратегии РФ» [1, с.3-8].

Указана роль топливно-энергетического комплекса для РФ, приведена структура потребления первичной энергии в мире в 2015 году по видам топлива, при этом использование нефти (31.5%) и газа (22%) превышает суммарный вклад остальных источников. Прогноз до 2040 года предполагает сокращение потребности в нефти на 4% и увеличение потребности в газе на 2%, что сохраняет ведущую роль нефтегазового топлива, причём наилучшими перспективами обладает газовая отрасль. Использование нефтепродуктов различными видами транспорта составляет около 95%, прогнозируется сохранение такой тенденции до 2050 года.

В настоящее время истощённость континентальных нефтегазовых месторождений составляет 30-50%, месторождений континентального шельфа – менее 1%, при этом основные ресурсы сосредоточены в Арктическом регионе (58%), особенно в российском секторе Арктики (газ – 70%, нефть – 40%). Общее соотношение ресурсов углеводородов сухопутной и шельфовых частей составляет 56% и 44%. В РФ более 87% сырьевых ресурсов нефти и газа находятся в морях арктического региона – Баренцево, Печорское, Карское, Лаптевых, Восточно-Сибирское, Чукотское.

Перечислены крупные реализованные на шельфе РФ проекты: ледостойкая платформа «Приразломная», подводно-подлёдный газовый промысел «Кириновское», «Сахалин-1» и «Сахалин-2». Отмечается необходимость комплексного подхода к разработке месторождений, а также создания сооружений в ледостойком исполнении. Это требует постройки соответствующих заводов. Целесообразно использовать потенциал ВПК. Толчком к развитию работ на шельфе является открытие супергигантских месторождений (Штокмановское, Русановское, Ленинградское). Необходима единая система транспортировки продукции. Для освоения глубоководных прогнозных месторождений перспективно создание высокоавтоматизированных подводно-подлёдных промыслов.

Подчёркивается необходимость разработки полной нормативно-методической базы для основных объектов континентального шельфа РФ.

Работы лауреатов конкурса Арктика 2018 охватывают широкий спектр научно-прикладных задач, связанных с развитием АЗРФ. Рассмотрим работы, удостоенные первой премии.

Авторский коллектив ООО «Газпром добыча Надым» представил работу [1, с.23-25], посвящённую управлению мерзлотно-геологическими условиями нефтегазоконденсатных месторождений при наличии многолетней тенденции повышения среднегодовых температур. Так, для полуострова Ямал величина повышения среднегодовых температур многолетнемёрзлых грунтов за последние 25-30 лет составила до 4°C. Для резервирования надёжности фундаментов используются парожидкостные охлаждающие устройства, обеспечивающие

проектный тепловой режим в основаниях зданий и сооружений. Подход апробирован на Бованенковском месторождении.

В работе [1, с.24-33] ООО «Башнефть-Полюс» рассматриваются вопросы использования попутного нефтяного газа в осложняющих условиях арктических месторождений. Представлен комплекс технологий по обеспечению эффективного водогазового воздействия для повышения нефтеотдачи без нанесения ущерба экологии региона.

В работе [1, с.34-38] АО «ЦНИИАГ» (совместно с МГТУ им. Н.Э. Баумана) описывается автономная радионавигационная система, обеспечивающая требуемую точность позиционирования судна и устойчивая к помехам. Апробация в Обской губе показала, что система, в дополнение к существующим приёмникам ГЛОНАСС/GPS, повышает безопасность мореплавания при движении по морскому каналу.

В ОАО «МАГЭ» разработана и реализована технология [1, с.39-41] одновременного выполнения ряда геофизических методов за один проход судна, что позволяет в два раза сократить временные затраты и способствует минимизации влияния на экологическую систему Арктических морей.

АО «НТЦ ВСП «Супертел ДАЛС» разработало отечественный проект «Северное Сияние» [1, с.42-45], представляющий собой комплекс оборудования для подводных волоконно-оптических линий связи. Предполагаемая трасса прокладки проходит по СМП от Мурманска до Владивостока и будет способствовать повышению информационной безопасности РФ.

Проект электросетевого комплекса Арктического региона РФ представлен в работе АО «Тюменьэнерго» [1, с.46-48]. Проект разработан с учётом климатических, геотехнических и геофизиологических условий региона с применением ряда инновационных разработок, испытанных в ЯНАО и ХМАО-Югра.

В работе ООО «Газпром нефть шельф» [1, с.49-51] описан комплекс мер и систем обеспечения безопасности Приразломного нефтяного месторождения.

Методический подход к оценке технологических рисков при эксплуатации подводных компрессорных установок развивается в работе ООО «Газпром ВНИИГАЗ» [1, с.52-59]. Применение подхода позволяет уже на этапе проектирования минимизировать риски.

Сотрудниками ИИММ КНЦ РАН предложен проект [1, с.60-63] информационно-аналитической системы поддержки планирования борьбы с типовыми промышленно-природными авариями в АЗРФ (на примере Мурманской области).

Группой предприятий (АО «НПП ПТ «Океанос» и др.) разработана концепция [1, с.64-86] роботизированной подводной сейсморазведки в подлёдных акваториях. Обоснована принципиальная возможность создания подводно-надводного аппарата повышенной автономности «Тень» с изменяемой геометрией корпуса, превышающего по своей эффективности зарубежные аналоги.

В работе ФГАОУ ВО «МГИМО МИД РФ» [1, с.87-92] выполнен анализ энергетической политики США в Арктике, начиная с 80-х годов XX века до наших дней, приведён прогноз на перспективу.

Отметим тематику работ лауреатов второй и третьей премии. Это строительство подводных (подлёдных) скважин, геологическая модель участка

шельфа, управление ледовыми операциями при строительстве морских скважин, система комплексной безопасности и мониторинга объектов и акватории СМП, арктические возобновляемые источники энергии – вихревые ветрогенераторы, перспективы применения СПГ на объектах Арктического региона, геофизическая разведка с ледовой поверхности, защита органов дыхания в условиях климата Арктической зоны, ликвидации пожаров на объектах Арктического бассейна, определение коэффициента пористости для ядра шельфовых пород, подбор ледостойкого покрытия на отгрузочном причале, защитные покрытия трубопроводов, варианты освоения Долгинского месторождения, гидропосев многолетних трав в Арктической зоне, гидроакустическое картирование дна арктического шельфа, защита опор качения машин от низких температур и водородного износа, проблемы качества жизни населения Арктики, разработка эластомерных материалов для подшипниковых узлов, экологическая безопасность горных работ на россыпных месторождениях Арктической зоны, энергосберегающий проект наружного освещения Тикси, балльно-рейтинговая оценка экологической безопасности объектов Арктики и континентального шельфа, мобильный фитотехнический комплекс для выращивания зелени и овощей в условиях Крайнего Севера.

Заключение

Таким образом, арктическая тематика представляет собой многомерный междисциплинарный «клубок» проектов, проблем, организаций, технологий. Можно наметить такие возможные признаки классификации тематики:

- разбиение территории и акватории на участки по различным признакам (месторождения, лицензионные участки, промышленные объекты, порты и объекты инфраструктуры, особые природные объекты, транспортные сети и коммуникации, энергетические сети, военные объекты, природно-климатические условия);
- иерархическая совокупность корпораций и организаций, взаимодействующих в Арктической зоне РФ (проектные, судостроительные, транспортные, добывающие, перерабатывающие, охраняющие, военные);
- совокупность проблем, свойственных арктическим регионам (геополитические, природно-климатические, социально-экономические, законодательные, обеспечение безопасности);
- совокупность проектов освоения и развития Арктики и континентального шельфа (стадия, масштаб, риски, обеспеченность).

Очевидно, выбор тех или иных элементов классификации, объединение их в структуры определяются решаемой задачей.

К числу ключевых, системообразующих факторов (проектов, задач, проблем), определяющих приоритеты развития АЗРФ, можно отнести:

- развитие СМП;
- разведка, добыча, переработка, транспортировка углеводородного сырья;
- природно-климатические факторы;
- социально-экономические факторы;
- проблемы безопасности;
- научно-технологические проблемы.

Литература

1. Сборник работ лауреатов Международного конкурса научных, научно-технических и инновационных разработок, направленных на развитие и освоение Арктики и континентального шельфа 2018г. – М.: Министерство энергетики РФ, ООО «Технологии развития», 2018. – 224с. ISBN 978-5-7688-1149-5.

Сведения об авторе

Яковлев Сергей Юрьевич - к.т.н, старший научный сотрудник, доцент

e-mail: yakovlev@iimm.ru

Sergey Yu. Yakovlev - Ph.D., senior researcher, associate professor

УДК 004.02, 004.9

DOI: 10.25702/KSC.2307-5252.2018.10.151-160

И.О. Датьев¹, М.Г. Шишаев^{1,2}, А.М. Федоров¹

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² ФГБОУ ВО «МАГУ»

МОДЕЛЬ КОЛЛЕКТИВНОЙ ИДЕНТИЧНОСТИ ДЛЯ ИЗУЧЕНИЯ МЕСТНЫХ СООБЩЕСТВ

Аннотация

Социетальная безопасность является важной частью личной безопасности человека, что особенно важно для местных сообществ Крайнего Севера. Идентичность является важным объектом изучения и мониторинга как ключевой критерий социетальной безопасности. В настоящее время социальные объекты широко представлены в виртуальном пространстве онлайн-социальных сетей. В работе предлагается концептуальная модель идентичности, а также обсуждаются возможные методы анализа идентичности местных сообществ в контексте проблем социетальной безопасности с использованием данных онлайн-социальных сетей.

Ключевые слова:

Социетальная безопасность, местные сообщества, модель коллективной идентичности

I.O. Datyev, M.G. Shishaev, A.M. Fedorov

THE COLLECTIVE IDENTITY MODEL FOR THE LOCAL COMMUNITIES STUDY

Abstract

Societal security is a crucial part of the human security what is especially important for the local communities of the High North. Since identity is an ultimate criterion of societal security the former is the essential object for studying and monitoring. At the same time social entities are now widely represented in the virtual space of social networking services. In this work, we propose a formal model of identity and discuss possible ways of analyzing the identity of local communities in the context of social security problems using data from social networking services.

Keywords:

Societal security, local communities, collective identity model

Введение

Проникновение современных телекоммуникационных сетей в нашу повседневную жизнь привело к повышению роли онлайн-социальных сетей не только в виртуальном пространстве, но и в реальных социальных процессах и явлениях. Одним из перспективных направлений изучения виртуальных коммуникаций в контексте социальных вопросов является мониторинг и анализ социетальной безопасности (societal security). Анализ социальных сетей можно рассматривать как новый эффективный инструмент для социологических исследований, связанных с социетальной безопасностью.

В современном мире миграционных потоков и столкновения культур социетальная безопасность рассматривается как чрезвычайно важный вопрос. Поскольку социетальная безопасность определяется как «способность общества сохранять свой сущностный характер при меняющихся условиях и возможных или

реальных угрозах» [26], личностная безопасность (human security) и социетальная безопасность тесно взаимосвязаны, несмотря на критику концепции социетальной безопасности [23]. Поэтому, представляются важными задачи идентификации и предотвращения угроз социетальной безопасности, связанных с миграцией, «горизонтальной» и «вертикальной» конкуренцией сообществ, а также другими причинами. К решению таких задач необходимо подходить системно, обладая информацией о сущности и внутренней структуре социетальной безопасности и источниках ее угроз.

В данной работе рассматривается проблема анализа идентичности местных сообществ как ключевого критерия социетальной безопасности на основе современных инструментов и методов анализа социальных сетей. Первый вопрос, рассматриваемый в статье: что представляет собой концептуальная структура или модель, достаточно репрезентативная с одной стороны и обладающая достаточной сложностью в контексте рассматриваемых экспериментальных инструментов, с другой? Второй вопрос: существуют ли адекватные подходы к анализу идентичности местных сообществ на основе социальных сетей? Адекватность, в данном случае, предполагает наличие достаточно надежных методов мониторинга реальной социальной сущности (местного сообщества, в нашем примере) на основании анализа ее отражения в онлайн-социальных сетях.

Для исключения неправильного толкования необходимо уточнить значения используемых в работе терминов: социальная сеть, онлайн-социальная сеть, анализ социальных сетей, местное сообщество, идентичность (значение термина обсуждается в следующем разделе).

Наиболее часто используемое определение термина «социальная сеть» дается в [27]: «Социальная сеть — это социальная структура, состоящая из множества социальных сущностей (таких как отдельные лица или организации), множества диадических связей и социальных взаимодействий между участниками этой структуры».

Анализ социальных сетей (Social Network Analysis - SNA) – научное направление, в котором изучается структура социальных сетей для выявления локальных и глобальных шаблонов (моделей), нахождения влиятельных объектов и изучения динамики сети.

Онлайн-социальные сети (сайты социальных сетей, social networking services – SNS, или социальные медиа – social media) обычно определяются как веб-сервисы, которые позволяют отдельным пользователям создавать и совместно использовать индивидуальный профиль, устанавливать связи и взаимодействовать с другими пользователями на сайте [2].

Определение термина «местное сообщество», близкое по смыслу к данной работе, приведено в [19]: «Местное сообщество — это группа взаимодействующих организмов, совместно использующих среду. В человеческих сообществах на личность участников и степень их сплоченности могут влиять намерения, убеждения, ресурсы, предпочтения, потребности, риски и другие факторы».

Анализ социальных сетей — это междисциплинарная научная область, которая возникла на основе социальной психологии, социологии и статистики. В рамках анализа социальных сетей используются теория графов, теория равновесия, теория социального сравнения и, в последнее время, концепция социальной идентичности [14]. В целом, анализ социальных сетей в настоящее

время используется в различных науках и является одной из основных парадигм современной социологии.

Модель идентичности как объект социетальной безопасности

Объектом социетальной безопасности является коллективная идентичность, поэтому именно ей следует уделить особое внимание. Понимание механизмов формирования и поддержания коллективной (групповой) идентичности является ключом к пониманию процессов, связанных с обеспечением социетальной безопасности. В данной статье идентичность рассматривается в ее социальном аспекте, т.е. как мы идентифицируем себя по отношению к другим в соответствии с тем, что у нас есть. Альберто Мелуччи [16] определил коллективную идентичность как «интерактивное и совместное определение, созданное несколькими людьми (или группами на более сложном уровне) и касающееся ориентации действий, области возможностей и ограничений, в которых происходит действие». Исходя из практических соображений, авторами данной статьи не делается принципиальных различий между терминами «социальной идентичности», введенными Генри Тайфелем и Джоном Тернером [25] и «коллективной идентичностью», в то время как некоторые авторы находят небольшие различия между социальной и коллективной идентичностью.

В области социологии, психологии, политологии и философии существует много исследований идентичности. Широко распространено мнение, что Зигмунд Фрейд начал «психологическую линию» в исследованиях идентичности, которая сосредоточена главным образом на исследовании самоидентификации или формирования самооценки. Большой вклад в развитие концепции идентичности с позиции психологической науки внес Эрик Эриксон, который считал идентичность сложной сущностью с многоуровневой структурой [10]. Многие современные исследования идентичности в психологии спровоцированы работами Генри Тайфеля, который предложил теорию социальной идентичности, объяснив формирование самоидентификации посредством соотнесения индивида и социальных групп [22].

Общей особенностью «социологической линии» в исследованиях идентичности является рассмотрение социальной среды как источника культурных норм и символов, а также моделей поведения, которые служат основой для самоопределения человека как члена данной социальной группы. В социологических теориях, применяемых для изучения идентичности, выделяются несколько основных подходов: функционализм, структурный функционализм, символический интеракционизм, феноменологическая социология, интегративный подход. Основания упомянутых подходов формируются работами Дэвида Эмиля Дюркгейма, Толкотта Парсонса, Ральфа Густава Дарендорфа, Чарльза Хортон Кули, Джорджа Герберта Мида, Ирвинга Гоффмана, Юргена Хабермаса, Питера Людвиг Бергера, Томаса Лукмана, Пьера Бурдьё, Энтони Гидденса, Сэмюэля Филлипса Хантингтона.

Несмотря на различные взгляды на концепцию идентичности, все подходы имеют некоторые общие особенности, важные для практического моделирования. Одна ценная мысль в контексте понимания и моделирования явлений идентичности — это значимость контактов между людьми, которые определяют социальную дистанцию и, следовательно, сплоченность социальной

группы. С точки зрения символического интеракционизма, идентичность возникает только в том случае, если индивид включен в социальную группу и, если имеет место связь с членами этой группы [1]. Хантингтон представляет интегративный подход в исследованиях идентичности и утверждает, что идентичности, определяемые индивидом, в то же время являются результатом взаимодействия человека или группы с другими людьми или группами [12].

Другая точка зрения, в большей или меньшей степени разделяемая большинством исследователей, заключается в том, что содержание идентичности определяется в некоторых компонентах культуры, ценностей, действий и т. д., которые разделяются или отвергаются индивидом или социальной группой. Кроме того, эти компоненты должны быть не только сформулированы, но и оценены. Итак, предполагается, что идентичность определяется в некотором метрическом пространстве. Современные исследователи заявляют, что и индивидуумы, и группы обладают идентичностью, что дает основание рассматривать личную и групповую идентичность в общем пространстве признаков [12]. С другой стороны, многие авторы концептуализируют социальный мир непосредственно как пространство - пространство Блау [15] или социальное пространство Бурдые [5]. Определив некоторую метрику над этим пространством, мы получаем возможность идентифицировать эффекты кластеризации, оценивать расстояния между точками и распознавать факт и направления «движений». В общем, такой формализм вполне адекватен рассмотренной в данной статье проблеме при условии корректировки размерности пространства и его метрики.

Третьим значимым аспектом, объединяющим разные взгляды на групповую идентичность, является «антагонистическое» определение идентичности. Фрейд рассматривал проблему процесса построения коллективных идентичностей «мы» через определение «они», отмечая огромную роль антагонистических отношений для установления идентичности среди членов сообщества, связывающих сообщество с одной коллективной идентичностью [17]. В работах политологов Зижека, Муффа, Лакло рассматривается дискурсивный механизм построения идентичности, основанный на социальном антагонизме [24]. В социологии и социальной психологии Тайфель и коллеги отмечают, что межгрупповая дискриминация и, в частности, внутригрупповой фаворитизм играют ключевую роль в структурировании социального пространства [7, 22].

Подводя итог сказанному, можно сделать вывод, что наиболее адекватным общим подходом к формализации идентичности социальных групп в целях анализа и мониторинга является его представление в форме многомерного пространства. Идентичность имеет множество измерений самого разнообразного характера - территориальных, этнических, культурных и других. Хотя некоторые компоненты идентичности могут быть количественно определены - возраст, благосостояние и т. д., другие могут быть измерены только с использованием качественных, и не всегда порядковых шкал. Таким образом, пространство идентичностей является многомерным и неоднородным, а оценки компонентов идентичности, чаще всего, являются интервальными. В соответствии с этим представлением, пространство идентичности для двумерного случая будет выглядеть так, как показано на рисунке 1. Точка пространства обозначает идентичность, и в общем случае это будет область с несферической

конфигурацией, что отражает неопределенность в формировании компонентов идентичности. Высокое сходство между точками в терминах определенной метрики указывает на возможную групповую идентичность. Необходимыми условиями формирования и сохранения групповой идентичности в дополнение к сходству являются регулярная и достаточно интенсивная связь между потенциальными членами группы и наличие значительной «анти-идентичности».

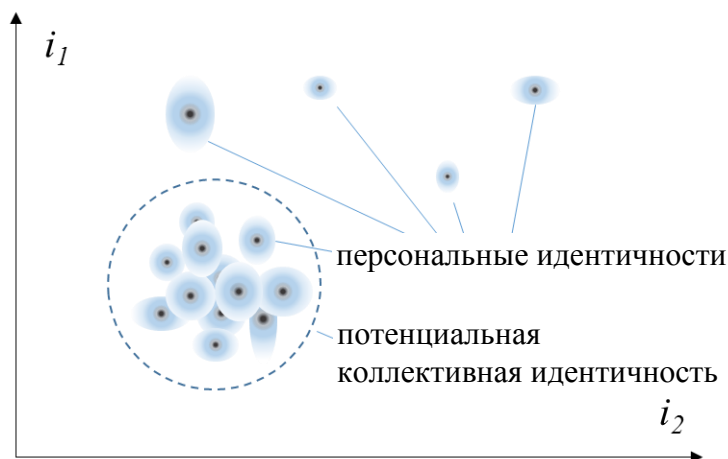


Рис. 1. Двумерное пространство идентичности

Согласно предложенной структуре, существует ряд вопросов, касающихся формализации идентичности: Как представить точки пространства идентичности с учетом неопределенности? Как определить антагонистические кластеры в пространстве? Как оценить устойчивость кластеров?

Можно предположить некоторые идеи и реализацию существующих методов для решения перечисленных проблем. В частности, представляется перспективным учитывать при оценке устойчивости кластера такие факторы, как его «диаметр», количество включенных точек, наличие и силу связей между точками кластера (вызванные персональной идентичностью). Можно также предположить, что устойчивость кластера будет зависеть от соответствующих характеристик антагонистического кластера как «угрозы», которая провоцирует формирование групповой идентичности. В качестве возможных формальных инструментов можно было бы рассмотреть нечеткие проективные пространства, грубые, мягкие, мульти- и другие атипичные множества, мягкие вычисления [20, 31].

Какой из инструментов формализации будет наиболее эффективен, во многом зависит от характера исходных данных и способа их получения. В этой статье оцениваются возможности использования информации из социальных сетей в качестве источника данных.

Анализ социальных сетей: методы и приложения

Одной из первых работ, рассматривающих изучение компьютерных сетей как особых социальных явлений, является «Компьютерные сети как социальные сети» Барри Уэллмана [28]. Автор утверждает, что «когда компьютерные сети и машины связывают людей, компьютерные сети становятся социальными сетями». Успех

онлайн-социальных сетей прослеживается в их доминировании в современном обществе, так, например, Facebook имеет 2,13 миллиарда активных пользователей в месяц и в среднем 1,4 миллиарда активных пользователей в 2017 году [11]. Существует множество методов и приложений, основанных на анализе социальных сетей: системы рекомендаций, анализ связей, экспертная идентификация, модели распространения влияния, предсказание уровня надежности участника, анализ тональности текста, поиск сообществ [3, 30, 33].

Любая крупная онлайн-социальная сеть генерирует большие данные (создаются профили, ссылки и контент). Анализируя эти данные, можно получить много полезной информации о разных сообществах и дискуссиях, а также о конкретном пользователе по отдельности [4, 13].

Поскольку опубликованные тексты сообщений не подвергаются какой-либо цензуре, они являются очень ценным материалом для контент-анализа, чтобы получить выводы о текущих социальных процессах, возникающих тенденциях, отношении людей к определенным предметам и явлениям, их участии в политической жизни. В настоящее время, эволюционирует направление, называемое Semantic Social Network Analysis [9, 18], объединяющее технологии семантической сети и классический анализ социальных сетей для улучшения анализа онлайн-социальных сетей.

Наряду с изучением самих онлайн-социальных сетей как особого вида виртуальной социальной среды, интерес исследователей также фокусируется на взаимосвязи между онлайн (существующими в виртуальном мире) и автономными (существующими в реальном мире) социальными процессами, сущностями и системами. Такие задачи наиболее важны в контексте сформулированной в данной статье проблемы. В качестве примера исследования, направленного на установление соответствия между виртуальными и реальными социальными сущностями, можно рассмотреть работу [21]. Еще одна проблема, связанная с установлением соответствия онлайн- и оффлайн-миров, — это определение географических координат виртуальных объектов. В [32] авторы сосредотачиваются на перспективах практического использования данных проекта «Виртуальное население Российской Федерации». Ряд работ, посвященных различным методам и приложениям геовизуализации данных социальных сетей, в том числе объединению геовизуализации и анализа тональности текста, рассмотрен в [8].

Основное различие между сформулированной в статье проблемой и методами анализа социальных сетей заключается в том, что в последнем случае сеть является объектом исследования, тогда как в данной работе она является инструментом изучения реальных социальных структур и процессов. В то же время, в обоих случаях возникает ряд подобных задач, связанных с идентификацией и анализом структур реальных и виртуальных сообществ. В случае онлайн-социальных сетей такие задачи рассматриваются в методах поиска сообществ (Community Mining). Сетевое сообщество представляется как группа абстрактных сетевых узлов, внутри которых связи достаточно плотные, а между которыми связи - разреженные. Проблема поиска сетевых сообществ заключается в нахождении всех сообществ внутри данной сети. Большое количество задач можно свести к проблеме поиска сетевых сообществ. В [29] авторы сформулировали основные проблемы поиска сообществ и проанализировали различные подходы к их решению.

Двумя широко рассматриваемыми проблемами в рамках поиска сообществ являются динамика (или эволюция) сообщества и пересечение сообществ. Эволюция сообщества появляется, например, когда система рекомендаций анализирует имеющиеся данные и предлагает что-то, интересующее пользователя, в результате в сети появляется новая связь (вершина, ребро). Пересекающиеся сообщества возможны, если узел является членом более чем одного сообщества. В статье [6] обсуждаются различные подходы, основанные на модульности, для обнаружения пересекающихся сообществ в социальных сетях.

Изучая идентичность сообществ, такие методы могут использоваться для мониторинга эволюции идентичности.

Заключение

В работе рассмотрены возможности и проблемы изучения идентичности местных сообществ на основе анализа социальных сетей. Задача изучения идентичности рассмотрена в контексте вопросов социетальной безопасности, которые являются важными компонентами личностной безопасности.

Идентичность — это сложный объект со множеством измерений (параметров), который представлен общими ценностями и действиями. Сложность моделирования и практического анализа идентичности усугубляется невозможностью получить точные количественные оценки параметров. Соответствующим способом представления идентичности является многомерное нечеткое пространство, размеры которого соответствуют параметрам идентичности, а точки (области) — это отношения индивида к соответствующим ценностям и действиям. Близость точек в пространстве является одним из необходимых условий возникновения и существования групповой идентичности. Кроме того, необходимыми условиями являются наличие достаточно надежных связей и интенсивного взаимодействия между членами сообщества, а также наличие достаточно значимой и компактной (в модельном пространстве) «анти-идентичности». Таким образом, формальная модель идентичности, ориентированная на использование в контексте социетальной безопасности, должна определять тройку: (групповая идентичность, внутригрупповые коммуникации, анти-идентичность).

Территориально определенный характер идентичности местных общин является важной особенностью в контексте рассматриваемой проблемы. Для Крайнего Севера важность этого фактора еще выше вследствие слабого развития транспортной инфраструктуры. Все это определяет важность таких конкретных практических задач анализа социальных сетей, как установление соответствия между виртуальными и реальными социальными сущностями, а также определение региональной идентичности (домашнего региона) сущностей виртуального пространства.

В настоящее время онлайн-социальные сети активно изучаются как независимое социальное явление. Существует большое количество методов, технологий и приложений, используемых для анализа различных объектов и процессов в онлайн-социальных сетях. Особое внимание уделено проблемам установления соответствия между виртуальными и реальными социальными сущностями, а также проблеме определения географического расположения сущностей онлайн-социальных сетей. Многие авторы относят анализ онлайн-социальных сетей к анализу больших данных. Это говорит о том, что решение

конкретных проблем в значительной степени основано на эвристике, связанной с комбинированием различных источников данных и методов их анализа. Современные онлайн-социальные сети предоставляют инструментарий (API) для углубленного анализа разнородных данных, хранящихся в сети. В идеальном случае, необходимо обрабатывать всю информацию, доступную в онлайн-социальных сетях. Представляется перспективным использование сентимент-анализа, технологий «семантик веб» и других методов, основанных на обработке формализованных знаний. Таким образом, можно сделать вывод, что анализ онлайн-социальных сетей является многообещающим и эффективным средством изучения реальных социальных явлений, включая идентичность местных сообществ. Однако конкретные технологии и алгоритмы для решения этих проблем требуют специальных исследований для оценки корректности и точности результатов, полученных на их основе.

Литература

1. Abels H. Interaction, identity, presentation. Introduction to interpretative sociology. / Transl. from German. SPb., 2000.
2. Amichai-Hamburger, Y; Hayat, T (2017). "Social Networking". In Rössler, P. The International Encyclopedia of Media Effects. 2. John Wiley & Sons, Inc. pp. 1–12. doi:10.1002/9781118783764.wbieme0170
3. Anjaria, Malhar; Mohana, Ram; and Guddeti, Reddy. A Novel Sentiment Analysis of Social Networks Using Supervised Learning, Social Network Analysis Mining, vol. 4, no. 181, March (2014)
4. Bakaev, V.A.; Blessing, A.V. Analysis of profiles in social networks. Information Technologies and Nanotechnologies, 2017. The Science of Data, pp. 1860-1863
5. Bourdieu P. (1989) Social Space and Symbolic Power / Sociological Theory Vol. 7, No. 1 (Spring, 1989), pp. 14-25
6. Chitra Devi, J.; Poovammal, E. An Analysis of Overlapping Community Detection Algorithms in Social Networks / Procedia Computer Science, Volume 89, 2016, Pages 349-358
7. Cosmides L.; Tooby J.; Kurzban R. (April 1, 2003). "Perceptions of race". Trends in Cognitive Sciences. 7 (4): 173–179
8. Croitoru, Arie; Crooks, Andrew; Radzikowski, Jacek, Stefanidis, Anthony. Geovisualization of social media. The International Encyclopedia of Geography, 2017. Pages 1-17, DOI: 10.1002/9781118786352.wbieg0605
9. Erétéo G., Buffà M., Gandon F., Corby O. Analysis of a Real Online Social Network Using Semantic Web Frameworks. The Semantic Web - ISWC 2009. Lecture Notes in Computer Science, vol 5823. Springer, Berlin, Heidelberg DOI https://doi.org/10.1007/978-3-642-04930-9_12
10. Erikson, Erik H. Identity: Youth and Crisis. Norton, New York, 1968. 336 p.
11. Facebook Newsroom: Company Info. Access mode: <https://newsroom.fb.com/company-info> (Retrieved 2018-04-12)
12. Huntington, S.P. (2004) Who are We?: The Challenges to America's National Identity. Simon and Schuster, 2004 – 428 p.

13. Khotilin, M.I.; Blagov, A.V. Visualization and Cluster Analysis of Social Networks. CEUR Workshop Proceedings, 2016. Vol.1638.– P.843-850
14. Kilduff, M.; Tsai, W. Social networks and organizations. Sage Publications, 2003, 172 pages.
15. McPherson, Miller (1983). "Ecology of Affiliation". American Sociological Review. 48: 519–532.
16. Melucci, Alberto. Nomad of the Present. Temple University Press, 1989
17. Mouffe Ch. On the political. London: Routledge, 2005.
18. Nakatsuji M., Zhang Q., Lu X., Makni B. and Hendler J. A., "Semantic Social Network Analysis by Cross-Domain Tensor Factorization," in IEEE Transactions on Computational Social Systems, vol. 4, no. 4, pp. 207-217, Dec. 2017. doi: 10.1109/TCSS.2017.2732685
19. Nkala Busi. Community Education in New HIV Prevention Technologies Research. Chapter in Handbook of Research on Technoethics, 2009, Pages: 11. DOI: 10.4018/978-1-60566-022-6.ch022
20. Petrovsky A.B. Spaces of sets and multisets. - M.: Editorial URSS, 2003. - 248 p.
21. Smirnov, I.; Sivak, E.; Kozmina, Y. In Search of Lost Profiles: The Reliability of VKontakte Data and its Importance in Educational Research. Voprosy obrazovaniya / Educational Studies Moscow. 2016. No 4. P. 106–122
22. Tajfel, H., & Turner, J. C. (1986) [1979]. "The social identity theory of intergroup behaviour (pp. 7–24)". In Austin, William G.; Worchel, Stephen. Psychology of Intergroup Relations (2nd ed.). Chicago: Nelson-Hall.
23. Theiler, T. (2003), Societal security and social psychology, Review of International Studies, 29: 249-268
24. Torfing J. New Theories of Discourse: Laclau, Mouffe and Žižek. Oxford: Blackwell, 1999.
25. Turner, J. C. & Reynolds, K. J. (2010). "The story of social identity". In T. Postmes & N. Branscombe. Rediscovering Social Identity: Core Sources. Psychology Press
26. Wæver, O., Buzan, B., Kelstrup, M., Lemaitre, P. with Carlton, D et al. (1993) Identity, Migration and the New Security Agenda in Europe (London: Pinter).
27. Wasserman, Stanley; Faust, Katherine (1994). "Social Network Analysis in the Social and Behavioral Sciences". Social Network Analysis: Methods and Applications. Cambridge University Press. pp. 1–27. ISBN 9780521387071
28. Wellman, B.; Salaff, J.; Dimitrova, D.; Garton, L., Gulia, M.; and Haythornthwaite, C. Computer Networks as Social Networks: Collaborative Work, Telework, and Virtual Community. Annual Review of Sociology Volume 22, 1996 Wellman, pp 213-238
29. Yang B., Liu D., Liu J. (2010) Discovering Communities from Social Networks: Methodologies and Applications. In: Furht B. (eds) Handbook of Social Network Technologies and Applications. Springer, Boston, MA, pp 331-346 DOI https://doi.org/10.1007/978-1-4419-7142-5_16
30. Yuan, Ting; Cheng, Jian; Zhang, Xi; Liu, Qingshan; and Lu, Hanqing. How Friends Affect User Behaviors? an Exploration of Social Relation Analysis for Recommendation, Knowledge Based Systems, vol. 88, pp. 70–84, (2015)

31. Zadeh LA, Abbasov AM, Shahbazova SN (2015) Fuzzy-based techniques in human-like processing of social network data. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*. 23(Suppl. 1):1–14 DOI 10.1142/S0218488515400012
32. Zamyatina, N. Yu.; Yashunsky, A. D. Virtual geography of virtual population. *Monitoring of Public Opinion: Economic and Social Changes*. 2018. № 1. P. 117—137. DOI: 10.14515/monitoring.2018.1.07.
33. Zhang, Jing; Fang, Zhanpeng; Chen, Wei; and Tang, Jie. Diffusion of Following Links in Microblogging Network, *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 8, pp. 2093–2106, 2015

Сведения об авторах

Федоров Андрей Михайлович – к.т.н., ведущий научный сотрудник

e-mail: fedorov@iimm.ru

Andrey M. Fedorov – PhD (Tech. Sci.), leading researcher

Шишаев Максим Геннадьевич – д.т.н, доцент, профессор РАН, главный научный сотрудник ИИММ КНИЦ РАН

e-mail: shishaev@iimm.ru

Maxim G. Shishaev – Dr.Sci. (Tech.), associate professor, professor of RAS, lead researcher

Датьев Игорь Олегович – к.т.н., старший научный сотрудник

e-mail: datyev@iimm.ru

Igor O. Datyev – PhD (Tech. Sci.), senior researcher

УДК 621.039.577:624.1 (985)
DOI: 10.25702/KSC.2307-5252.2018.10.161-169

А.О. Орлов, Ю.Г. Смирнов, В.В. Бирюков
Горный институт ФИЦ КНЦ РАН

АНАЛИЗ ПРОЕКТНЫХ РЕШЕНИЙ ПОДЗЕМНЫХ КОМПЛЕКСОВ АТОМНЫХ СТАНЦИЙ МАЛОЙ МОЩНОСТИ В УСЛОВИЯХ АРКТИКИ*

Аннотация

В статье рассмотрены основные предпосылки использования подземных атомных станций малой мощности (АСММ) и особенности строительства в арктических регионах России. Рассмотрены основные факторы по выбору конструктивно-компоновочных решений и обоснованию глубины заложения, способов доступа и параметров выработок. Приведены результаты анализа проектных решений подземных АСММ на основе ориентации подземных сооружений по функционально-технологическому принципу работы.

Ключевые слова:

арктические регионы, атомные станции малой мощности, подземное размещение, глубина заложения, способ доступа, конструктивно-компоновочные решения

A. O. Orlov, Y. G. Smirnov, V. V. Birykov

ANALYSIS OF DESIGN DECISIONS FOR UNDERGROUND COMPLEXES OF SMALL NUCLEAR POWER PLANTS IN ARCTIC

Abstract

The article considers the main prerequisites for the use of underground small nuclear power plants (SNPP) and features of construction in the Russian Arctic regions. The main factors on the choice of constructive-layout solutions and substantiation of the depth of location, methods of access and parameters of the excavations are described. Results of the analysis of design decisions of underground SNPP on the basis of the orientation of underground structures on the functional and technological work principle are presented.

Keywords:

Arctic regions, small nuclear power plants, underground disposition, depth of location, access method, constructive layout solutions

Введение

Освоение арктических регионов России является одной из главных приоритетных задач государственной политики в Арктике и ставит вопрос развития инженерной инфраструктуры в регионах многолетней мерзлоты, где комфортное проживание человека в суровых условиях весьма затруднительно.

Создание альтернативных источников энергии имеет большое политическое и экономическое значение для данных регионов, которые характеризуются отсутствием централизованного энергоснабжения, сложными природно-климатическими условиями, большими затратами на генерацию и транспортировку электроэнергии.

* Работа выполнена в рамках темы 0226-2018-0008 государственного задания Горного института КНЦ РАН.

Подземное размещение атомных станций малой мощности (АСММ) может стать недорогим альтернативным источником энергообеспечения в районах со сложными климатическими и транспортными условиями. Поэтому использование подземного пространства в условиях многолетнемерзлых пород является актуальной и важной задачей.

Общие положения. Основная идея сооружения АСММ заключается в создании безопасных подземных комплексов способных противостоять любому негативному воздействию. В существующем правовом поле законодательно закреплена приоритетность вопросов обеспечения безопасности атомных станций. В настоящее время, технически строительство подземных АСММ возможно, практически, в любых инженерно-геологических условиях.

Одно из главных мест в создании подземных комплексов отводится задачам, направленным на определение подземной архитектуры сооружения и рациональным конструктивно-компоновочным решениям. Среди основных факторов, влияющих на выбор компоновки подземного комплекса, следует отметить горно-геологические условия, технологические параметры строительства, технические показатели комплекса и ряд других.

Строительство в арктической зоне. В районах с многолетней мерзлотой, строительство подземных сооружений может сопровождаться оттаиванием или промерзанием пород. Температура в породном массиве и ее колебания так же изменяются с глубиной. Особенностью мерзлоты является то, что сезонные колебания температуры затухают на глубине 10–15 м, а на более низких глубинных отметках температура мерзлоты постоянна. Мерзлота практически не проницаема для жидкостей и газов.

При строительстве подземных АСММ в арктических условиях необходимо принимать во внимание целый ряд специфических факторов, связанных как с климатическими особенностями, так и с особенностями эксплуатации самого подземного объекта, основными из которых являются:

- техногенные воздействия, способные изменять физико-механические свойства вмещающего массива и строительных конструкций;
- повышенный температурный режим внутри камерных выработок, влияющий на напряженно-деформированное состояние горных пород;
- знакопеременный температурный режим на контактах строительных конструкций и породного массива;
- краткосрочное воздействие высоких давлений и температур при возможных аварийных ситуациях способные вызвать снижение несущей способности и противодиффузионной стойкости окружающего породного массива.

Технические характеристики. Подземная атомная станция включает в себя комплекс различного оборудования, систем защиты и управления. Технические характеристики комплекса связаны с используемым типом оборудования и параметрами горных выработок для них, к которым можно отнести: мощность, тип и целевое назначение реактора, габариты и компоновка основного оборудования, параметры главных вспомогательных выработок. Так же к важным параметрам, влияющим на конструкцию и компоновку подземного комплекса, следует отнести рельеф местности, глубину заложения основного горизонта, способ доступа к подземному комплексу с поверхности.

Рельеф местности. Рельеф для строительства подземного комплекса определяет возможность размещения поверхностного комплекса и организацию транспортной инфраструктуры. Главное влияние рельефа связано с определением предпочтительной схемы вскрытия (доступа в подземный комплекс), что напрямую затрагивает вопросы экономики, сроков строительства и обеспечения безопасности.

Способ доступа. Начало строительства связано со способом доступа к подземному комплексу через главную вскрывающую выработку. Определяющими в выборе способа доступа (вскрытия) являются горно-геологические условия и рельеф местности. К основным способам доступа относятся: при помощи *вертикальных* (шахтный ствол), *наклонных* (уклон, спиральный съезд) и *горизонтальных* выработок (штольня, тоннель), а также возможны различные их комбинации.

Вскрытие вертикальной выработкой применяется, в основном, при равнинном рельефе местности и является одним из сложных и трудоемких в горном строительстве. При гористом рельефе предпочтительнее использование горизонтальных выработок. Вскрытие наклонными выработками используется практически при любом рельефе. Экономически целесообразно проводить вскрытие до глубины 100 м, так как при большой протяжённости вскрывающих выработок, существенно повышается объём работ и стоимость строительства [1].

Глубина заложения. Обосновывается на основании параметров безопасности, с учетом приемлемых экономических затрат. Определяется толщиной горной породы, которая позволяет исключить процесс распространения радиоактивных выбросов в атмосферу при запроектной аварии и других возможных внутренних инцидентах.

Горные породы выдерживают динамические нагрузки на порядок больше, чем защитные конструкции атомных станций, расположенных на дневной поверхности.

С увеличением глубины расположения АСММ возрастают изолирующие свойства породного массива, повышается безопасность эксплуатации подземного комплекса, однако, наряду с этим увеличивается горное давление и затраты на поддержание горных выработок. По геомеханическим и экономическим параметрам оптимальным можно считать размещение АСММ на глубине 50-100 м.

Конструктивно-компоновочные решения. Все выработки подземного комплекса связаны с целевым назначением и ориентированы по функционально-технологическому принципу работы. Компоновка и выбор оптимальных параметров горных выработок подземной АСММ позволяет обеспечить длительную устойчивость подземного сооружения в течение всего срока эксплуатации[2].

Форма и размеры выработок. Форма и размеры поперечного сечения выработки должны обеспечивать необходимую технологическую возможность размещения в ней применяемого оборудования и инженерных коммуникаций, а также соблюдение необходимых требований техники безопасности при транспортировке оборудования и материалов, вентиляции и водоотлива. При расчете проектного сечения выработки в проходке кроме толщины крепи необходимо учитывать горно-геологические характеристики вмещающих пород.

Форму поперечного сечения выработки выбирается в зависимости от устойчивости пород, срока службы и назначения выработки.

Для горизонтальных выработок, как правило, принимается сводчатая форма поперечного сечения с вертикальными боковыми стенками. На рис. 1 показана принципиальная схема формирования геометрических параметров для всех типов горизонтальных выработок и камер.

Размеры поперечного сечения выработки определяются габаритами основного (L), подъемного (C) и транспортного оборудования (A), необходимыми зазорами между крепью и указанным оборудованием (D), а также количеством воздуха, которое должно проходить по выработке. Необходимо, чтобы площадь поперечного сечения выработок обеспечивала перемещение по ним воздуха с регламентируемой скоростью.

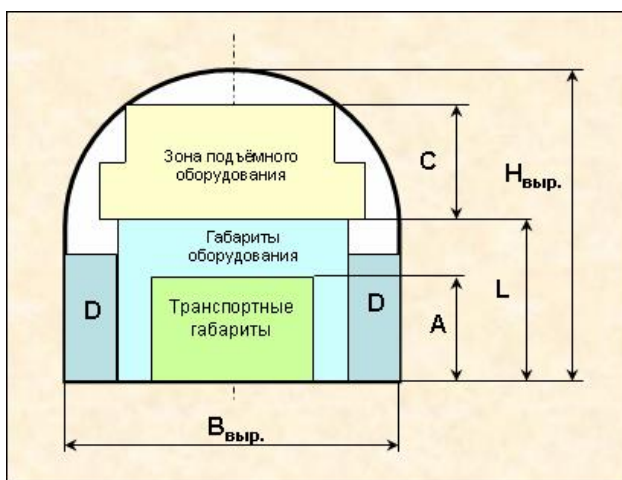


Рис. 1. Принципиальная схема формирования геометрических параметров выработки

$H_{\text{выр}}$ – высота выработки; $B_{\text{выр}}$ – ширина выработки

Методический подход. Основная идея сооружения АСММ заключается в создании безопасных подземных комплексов способных противостоять любому негативному воздействию. Научный подход к проектированию перспективных подземных комплексов, как сложных систем, основан на применении методов структурно-параметрических исследований, под которым понимается определение такой структуры проектных параметров, которая соответствовала техническому заданию и обеспечивала экстремум выбранного критерия качества [3].

Наиболее перспективным представляется использование специализированных автоматизированных систем проектирования на всех стадиях разработки от технического задания до выпуска проектной документации. Решение поставленной проблемы требует разработки системной модели АСММ как объекта проектирования. [4].

Вызывает интерес протекание процесса теплопередачи в массиве горных пород вследствие знакопеременных температур на контуре выработки. Стенка горной выработки нагревается теплым воздухом с температурой $15^0 - 20^0$ С от действия промышленного оборудования и формирует тепловой поток в глубину многолетнемерзлого горного массива.

На основе уравнения переноса энергии (1) была разработана математическая модель процесса теплопередачи

$$\frac{\partial}{\partial t}(\rho E) + \nabla \cdot (\vec{V}(\rho E + p)) = \nabla \cdot \left(k_{eff} \nabla T - \sum_j h_j J_j + (\vec{\tau}_{eff} \cdot \vec{V}) \right) + S_h$$

где

$\frac{\partial}{\partial t}(\rho E)$ - изменение полной энергии E единицы массы ρ со временем;

$\nabla \cdot (\vec{V}(\rho E + p))$ - компонент описывающий конвективный перенос энергии E единицей массы ρ в среде, движущейся со скоростью \vec{V} , и под давлением p ;

$\nabla \cdot (k_{eff} \nabla T)$ - компонент описывающий кондуктивный перенос энергии, связанный с градиентом температур. В случае анизотропной среды используется коэффициент теплопередачи k_{ij} ;

$\nabla \cdot (\vec{\tau}_{eff} \cdot \vec{V})$ - источник энергии, связанный с вязким трением в потоке жидкости;

$\nabla \cdot \left(\sum_j h_j J_j \right)$ - источник энергии, связанный с появлением в системе новой

массы вследствие химических реакций или фазовых переходов,

S_h - поступление энергии из за границы расчетного объема

В массиве горных пород перенос энергии описывается упрощенным уравнением кондуктивного теплообмена.

$$\frac{\partial}{\partial t}(\rho h) + \nabla \cdot (\vec{V} \rho h) = \nabla \cdot (k_{eff} \nabla T) + S_h$$

Разработанная математическая модель теплопереноса реализована в программном комплексе ANSYS Fluent (лицензия Горного института ФИЦ КНЦ РАН), имеющем возможность реализации модели переноса энергии в твердых телах в сочетании с конвективным теплообменом в жидкости и выделения энергии вследствие фазовых переходов и химических реакций.

В CAD редакторе Gambit на основе эскиза (рис 1) была разработана расчетная геометрия задачи (рис. 2). Площадь поперечного сечения выработки составляет 180 м^2 . Протяженность расчетной геометрии по вертикали и горизонтали равна 100м. В задаче такие размеры выбраны для компенсации погрешностей расчета, вносимых граничными условиями.

Разработанная математическая модель процесса переноса энергии в массиве горных пород выполнена в виде виртуального стенда, позволяющего проводить вычислительные эксперименты с различными граничными (рис. 3) и начальными условиями и различными физическими свойствами горных пород.

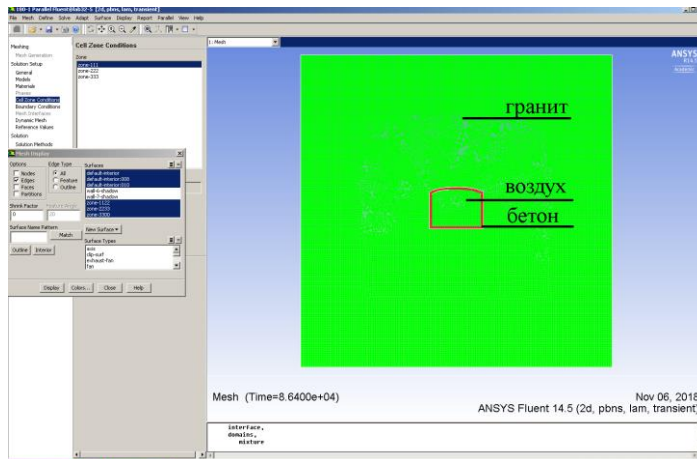


Рис. 2. Расчетная сетка плоской задачи передачи теплоты в горном массиве.

В таблице 1 приведены физические свойства твердых тел для вычислительного эксперимента, задаваемые в вычислительном эксперименте через диалоговое окно программы (рис. 4).

Таблица 1. Физические свойства материалов

	Плотность, кг/м ³	Теплопроводность Вт/м*К	Теплоемкость Дж/кг*К
Воздух	1.225	0.0242	1006.43
Бетон	2300	1.69	840
Гранит	2450	2.4	880

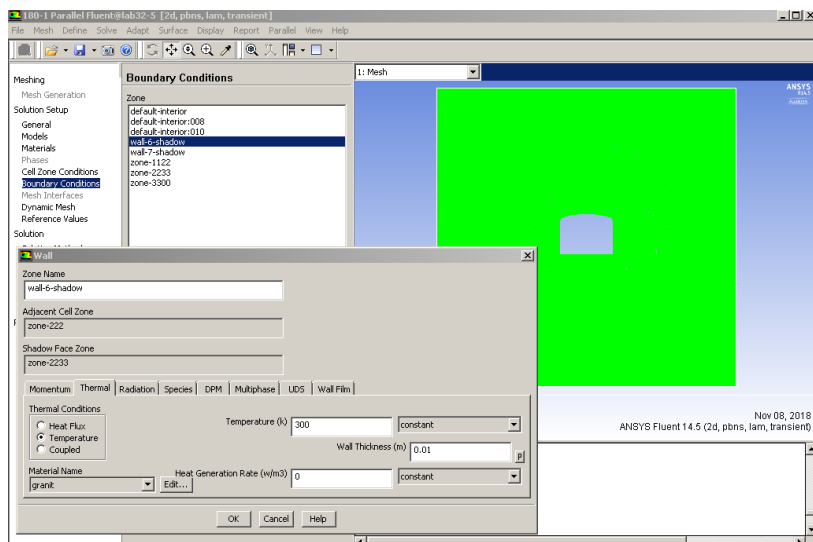


Рис. 3. Диалоговое окно задания граничных условий модели теплопереноса в горном массиве.

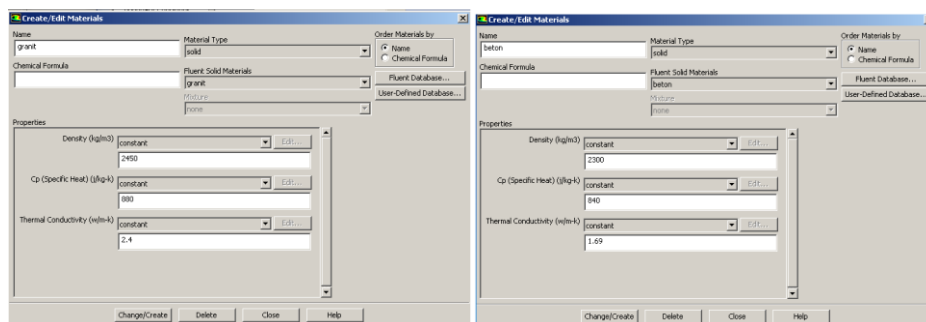


Рис. 4. Диалоговое окно задания свойств материалов

В результате проведения серии вычислительных экспериментов получена информация по тепловому состоянию горных пород с течением времени (рис.5, 6, 7).

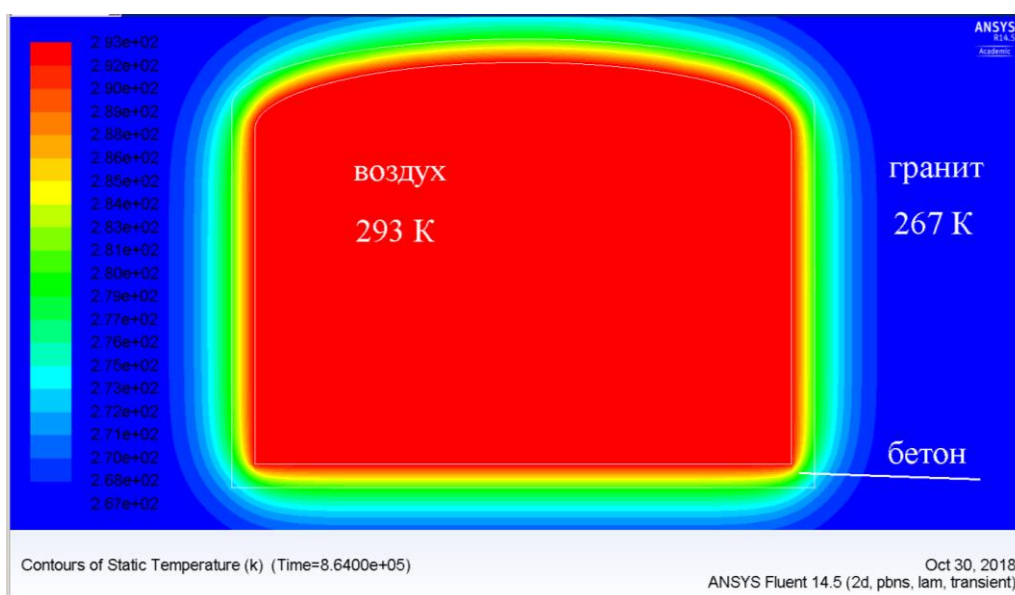


Рис. 5. Распределение температур на поперечном сечении выработки через 10 суток от начала процесса.

Заключение

В статье рассмотрены основные принципы создания подземных атомных станций малой мощности в районах многолетней мерзлоты.

Создание единого архитектурного решения подземного комплекса АСММ на основе отдельных выработок должно учитывать необходимость стабилизации основных температурных и физико-механических характеристик горных пород. Тепловая нагрузка на границе выработки приводит к изменению температуры слагающих горных пород, и, соответственно, изменению их физических свойств. Использование методов математического моделирования позволяет исследовать процессы, происходящие в горных породах с течением времени. Информация, полученная в результате проведения вычислительных экспериментов позволяет

прогнозировать напряженно деформированное состояние горных массивов, меняющееся в процессе строительства и последующей эксплуатации выработок.

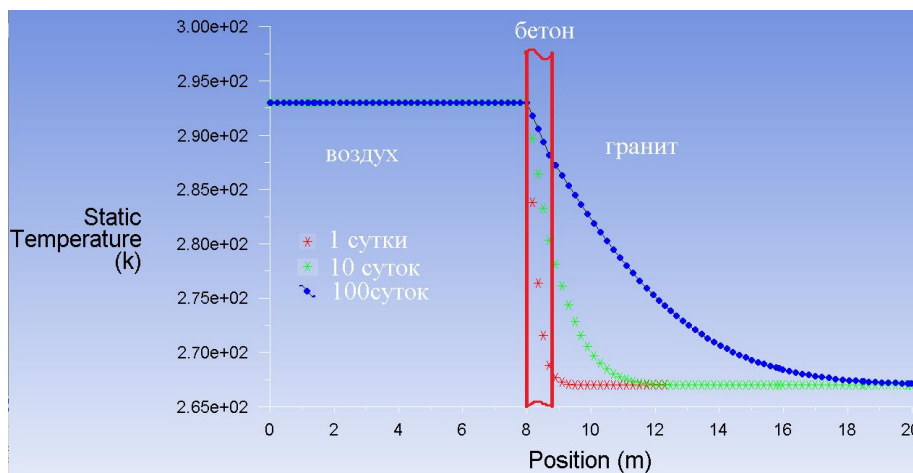


Рис. 6. Распределение температур в массиве горной породы по проществу 1, 10, 100 суток

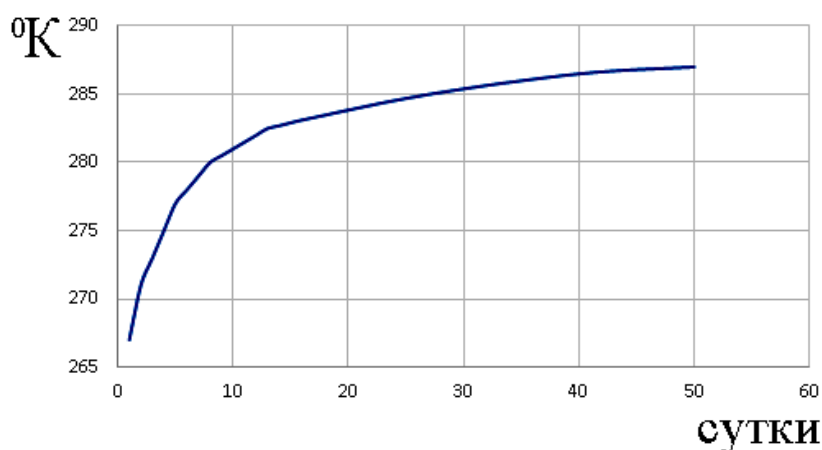


Рис. 7. График изменения температуры на границе горной выработки с течением времени

Литература

1. Орлов, А.О. Обоснование эффективных способов размещения подземных комплексов атомных станций малой мощности в российской Арктике / А.О. Орлов, Ю.Г. Смирнов. // Изв. вузов. Горн. журнал. - 2016. - №4. - С. 18-23.
2. Смирнов, Ю.Г. Анализ мирового опыта строительства подземных атомных станций и оценка возможности его использования для арктических регионов России / Ю.Г. Смирнов, А.О. Орлов. // Вестник МГТУ. - 2016. - Т.19. - № 1/1. - С. 47-52.
3. Лазарев, И.А. Композиционное проектирование сложных агрегатных систем / И.А. Лазарев - М.: Радио и связь - 1986. - 312 с.

4. Бешелев С.Д., Гурвич Ф.Г. Экспертные оценки в принятии плановых решений.
– М.: Экономика, 2008. – 256 с.

Сведения об авторах

Орлов Александр Орестович – научный сотрудник

e-mail: aleor@goi.kolasc.net.ru

Alexander O. Orlov – researcher

Смирнов Юрий Геннадьевич – научный сотрудник

e-mail: smirnov@goi.kolasc.net.ru

Yuriy G. Smirnov – researcher

Бирюков Валерий Валентинович – научный сотрудник

e-mail: birukov@goi.kolasc.net.ru

Valeriy V. Birukov – researcher

УДК 004.9

DOI: 10.25702/KSC.2307-5252.2018.10.171-178

А.С. Шемякин¹, Ю.В. Федотова²

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² Горный институт ФИЦ КНЦ РАН

АЛГОРИТМ ВИЗУАЛИЗАЦИИ РАСПРЕДЕЛЕНИЯ ЗОН ГЕОДИНАМИЧЕСКОГО РИСКА*

Аннотация

В качестве исходных данных для предлагаемого алгоритма визуализации распределения риска используются планы горных выработок с нанесёнными на них потенциальными источниками геодинамической опасности и математические модели распределения геодинамического риска. Результатом работы алгоритма является электронная карта с нанесёнными на ней зонами повышенной опасности ведения горных работ и выделенными зонами геодинамического риска.

Ключевые слова:

ГИС, QGIS, геодинамические риски, горные работы

A.S. Shemyakin, Iu.V. Fedotova

ALGORITHM FOR VISUALIZATION OF THE DISTRIBUTION OF GEODYNAMIC RISK ZONES

Abstract

As the initial data for the proposed risk distribution visualization algorithm, mine work plans with potential sources of geodynamic hazard and mathematical models of geodynamic risk distribution can be used. The result of the algorithm is an electronic map with marked areas of increased danger of mining and selected areas of geodynamic risk.

Keywords:

GIS, QGIS, geodynamic risk, mining

Введение

Исследование проводилось на примере проблем безопасности, специфичных для Мурманской области – региона, входящего в состав Арктической зоны России, и направлено на развитие информационных технологий и систем сетецентрического управления региональной безопасностью в Арктике [1, 2].

Современное состояние горных работ характеризуется наличием острейшей проблемы - техногенной сейсмичности, проявляющейся на рудниках и шахтах в виде горных ударов и техногенных землетрясений, которые определяют геодинамические риски. Эта проблема актуальна как для ряда регионов России, так и для многих горнорудных провинций мира [3].

Принятые в сфере геодинамической безопасности разрабатываемых месторождений полезных ископаемых классификации и показатели риска носят

* При поддержке НИР №АААА-А17-117040610316-8 «Модели и методы конфигурирования адаптивных многоуровневых сетецентрических систем управления региональной безопасностью в Арктической зоне Российской Федерации (АЗРФ)»

в основном качественный характер. Это объясняется сложностью физического процесса формирования очага и реализации динамического явления, большим числом взаимно влияющих факторов и трудностью математической формализации для определения количественных параметров уровня опасности. Поэтому достаточно широко применяются в промышленности, в частности в горной, вербальные оценки уровня риска [4].

Следует отметить, что геодинамическое явление может и не приводить к разрушениям горных выработок, но регистрация таких событий системами мониторинга однозначно свидетельствует о формировании геодинамических зон при ведении горных работ. Совершенно очевидно, что последствия геодинамического явления для рудника будут определяться не только энергией события, но и координатами очага. При этом, чем дальше зона геодинамической активности будет удалена от зоны влияния разрывного нарушения (зоны опасности реализации сейсмического события), тем меньше вероятность разрушения выработок, находящихся в этой зоне. В случае приближения горных работ к зоне влияния разрывного нарушения создаются условия концентрации напряжений на данном участке, что может привести к реализации динамических событий вблизи контура выработки. В связи с этим необходим учёт влияния этой зоны, поскольку она при приближении фронта работ переходит в зону геодинамической опасности (зону геодинамического риска разрушения массива горных пород) [5].

Выявление зон геодинамического риска имеет ряд особенностей, обусловленных геомеханическими процессами, протекающими в иерархично-блочном массиве и активно меняющемся полем напряжений при ведении горных работ. Среди этих особенностей необходимо выделить следующие:

1. Наличие неопределённости в описании текущей обстановки, обусловленной неполнотой информации о процессах, протекающих в горной среде.
2. Наличие структурной неопределённости, обусловленной неполнотой информации о структуре горного массива.

Понятия неопределённости о текущей обстановке и структурной неопределённости даны в [6]. Там же для учёта неопределённости предлагается использовать принципы и подходы, применяемые для определения/оптимизации уровня риска.

Обзор инструментария для визуализации пространственных данных

Очевидно, что в силу особенностей, перечисленных выше, для визуализации зон геодинамического риска необходимо так или иначе обрабатывать пространственные данные, полученные в результате мониторинга состояния массива горных пород. Такую возможность дают геоинформационные системы. Современные ГИС-решения имеют большие возможности по аналитической обработке пространственных данных, что позволяет использовать их в задачах мониторинга и анализа состояния природных ресурсов и процессов [7]. Одним из примеров такого применения ГИС-решений является [8], где использовалась ГИС ГЕОМИКС компании ВИОГЕМ для компьютерной обработки информации, полученной при выполнении геолого-структурных исследований. ГЕОМИКС является специализированным ГИС-решением, предназначенным для решения задач горного производства.

Ещё одним специализированным решением является МАЙНФРЕЙМ [9]. Данный программный комплекс можно использовать при решении технологических задач на разных стадиях производства. Программный комплекс позволяет автоматизировать процесс инженерного обеспечения при ведении горных работ и создать единое информационное поле предприятия. Это достигается за счет входящих в его состав специальных модулей и формирования на их основе единого геоинформационного пространства, в рамках которого выполняется создание и хранение моделей объектов горной технологии.

ГЕОМИКС и МАЙНФРЕЙМ являются специализированными ГИС-решениями, предназначенными для горнодобывающей отрасли, однако могут использоваться и «стандартные» ГИС-решения. Примером является работа [10]. ГИС ArcGIS использовалась совместно со сканирующей лазерной системой Leica HDS8800. При помощи сканирующей системы были получены исходные массивы точек, имеющие пространственные координаты и дополнительные атрибуты, а с помощью ArcGIS этот массив данных был систематизирован, и была построена цифровая модель рельефа.

Однако не всегда функционал используемой ГИС позволяет решать все задачи мониторинга и контроля, иногда требуется использовать комплексы программ. Так, например, в [11] используется «связка» программ: AutoCAD, Surfer 8, Microsoft Excel, где AutoCAD использовался для визуализации, Surfer 8 – для построения карт поверхности, а собственно расчёты осуществлялись в Excel.

Из приведённого краткого обзора видно, что в задачах визуализации результатов геомониторинга может использоваться разнообразное программное обеспечение: как специализированные ГИС-решения, так и геоинформационные системы общего назначения. Причем ГИС могут использоваться как самостоятельно, так и в связке с другими программными комплексами. Использование того или иного программного комплекса зависит главным образом от финансовых и технических возможностей.

Алгоритм визуализации распределения зон геодинамического риска

В данной работе показаны результаты использования открытой ГИС QGIS [12]. Из функционала, предоставляемого QGIS, наиболее интересной в данный момент функцией является возможность построения буферных зон для объектов на электронной карте [13]. Эту возможность можно использовать для визуализации опасных зон в пределах шахтных полей при ведении горных работ.

Стандартный модуль построения буферных зон (буферов) в QGIS позволяет строить зоны на основе выражений и атрибутивной информации объектов слоя. При необходимости пользователь может разработать собственную функцию вычисления размера буферной зоны на языке Python и использовать её в модуле. Поскольку пользователь разрабатывает собственную функцию на языке Python, то он не ограничен только атрибутивной информацией слоя и встроенными функциями QGIS, можно создать сколь угодно алгоритмически сложную функцию, насколько это позволят возможности языка Python.

Далее будет показано, как можно использовать возможности QGIS для визуализации распределения зон геодинамического риска с учётом неопределённостей. В общих чертах алгоритм визуализации состоит из четырёх этапов:

1. подготовка электронной карты горных выработок;
2. программирование моделей распределения зон влияния разрывных нарушений;
3. построение зон влияния разрывных нарушений в соответствии с запрограммированными моделями;
4. собственно построение геодинамических зон.

Далее каждый этап будет описан более детально.

Первым этапом является формирование электронной карты в QGIS и создание атрибутивной информации для каждого объекта электронной карты. На Рис. 1 изображён план участка горизонта подземного рудника, совмещенного с разрывными нарушениями (разломами). Каждая из этих геологических структур имеет свою зону влияния в зависимости от мощности, протяженности разлома и его минералогического заполнения. При ведении горных работ на горизонте, данные зоны являются источником геодинамических опасностей (динамических разрушений массива).

Представленная электронная карта состоит из двух слоёв: «Рудники» и «Разломы». В нашем случае можно ограничиться заданием атрибутивной информации для каждого объекта слоя «Разломы», т.к. на слое «Рудники» присутствует только один объект, и в данном случае никакие атрибуты этого объекта для расчётов не потребуются.

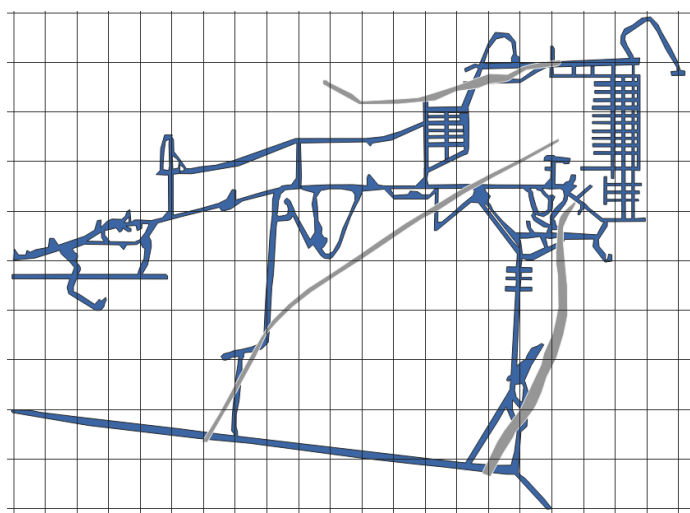


Рис. 1. Электронная карта рудника

Вторым этапом является программирование моделей распределения зон влияния разрывных нарушений на языке Python во встроенном редакторе функций QGIS. Желательно, чтобы исходные модели учитывали структурную неопределённость и неопределённость информации о текущей обстановке.

Поскольку разработка моделей распределения зон геодинамического риска является темой отдельного исследования, в данной работе ограничимся простейшей моделью для того, чтобы проиллюстрировать работу алгоритма визуализации этих зон.

Возьмем простую модель распределения зон влияния разрывных нарушений, в которой ширина буферной зоны равна средней мощности разлома.

Для реализации этой простейшей модели необходимо, чтобы среди атрибутов разлома был атрибут со средней мощностью разлома. В таком случае потребуется программирование единственной функции, которая будет возвращать значение атрибута средней ширины разлома в качестве результата. Причём, в качестве входного параметра функция будет принимать значение в метрах, а возвращать будет – в градусах дуги меридиана. Специфика представления данных в градусах связана с назначением QGIS - созданием электронных карт, привязанных к географическим координатам и, вследствие этого, качественного отображения рельефа поверхности.

После того, как все нужные функции запрограммированы, требуется загрузить их в редактор, и они станут доступны для редактора функций. В нашем случае выражение для построения буферных зон будет выглядеть так, как показано на Рис. 2, где `my_buf` – запрограммированная функция по переводу метров в градусы дуги меридиана, `"br_width"` – атрибут объекта в слое «Разломы». Этот атрибут хранит среднюю ширину разлома в метрах.

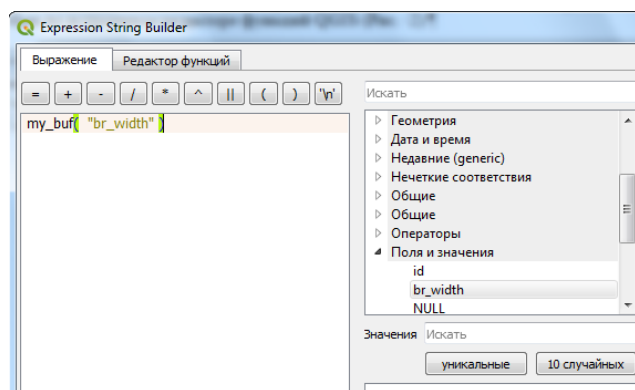


Рис. 2. Выражение в редакторе функций

Третьим этапом является построение зон влияния разрывных нарушений на электронной карте.

Встроенный в QGIS модуль построения буферов может самостоятельно перебрать все объекты, для которых необходимо построить буфер. Для каждого перебираемого объекта будет рассчитываться значение выражения, заданного на предыдущем этапе. В итоге все рассчитанные буферные зоны будут отрисованы на новом слое, причём для каждого исходного объекта будет создан отдельный объект буферной зоны (см. Рис. 3). Буферные зоны, отображённые на электронной карте, представляют собой зоны влияния разрывных нарушений при ведении горных работ.

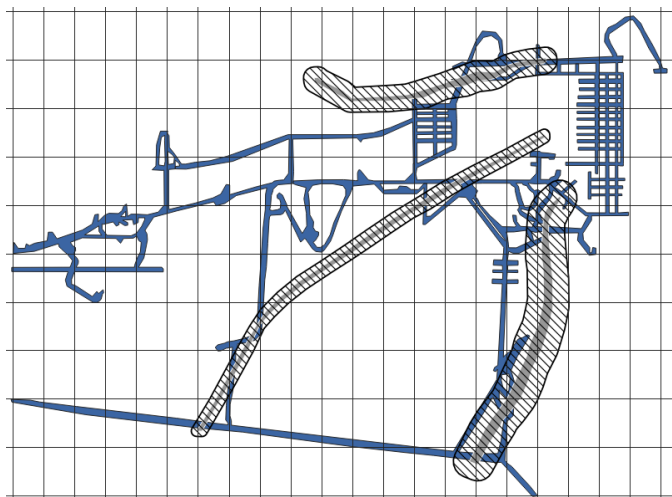


Рис. 3. Буферные зоны (зоны влияния разрывных нарушений)

Заключительным этапом является собственно построение зон геодинамического риска. Для этого необходимо выделить места пересечения зон влияния разрывных нарушений с выработками, которые собственно, и являются зонами геодинамического риска разрушения горных выработок при ведении горных работ в динамической форме (Рис. 4). Именно для этих зон необходимо проводить профилактические мероприятия по минимизации последствий геодинамических явлений при обеспечении безопасности ведения горных работ.

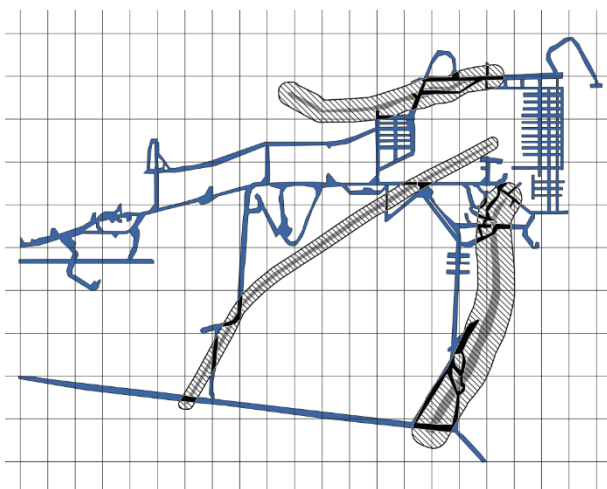


Рис. 4. Зоны геодинамического риска

Заключение

Геодинамические риски всегда имеют место при ведении горных работ в высоконапряженных массивах скальных пород. В связи с этим на горных предприятиях необходимо получение оперативной информации о распределении

зон геодинамического риска для принятия профилактических мероприятий по минимизации последствий реализации геодинамических явлений.

В данной работе представлен простой алгоритм визуализации распределения зон геодинамических рисков с учётом неопределённостей. Показано, что результат достигается в четыре этапа.

Недостатком данного алгоритма является то, что он разработан для однородного массива, что не соответствует действительности. Для неоднородного массива алгоритм будет более сложным, как и модели распределения зон геодинамического риска.

Достоинством алгоритма является относительная простота достижения результата, который может уже сейчас применяться для визуализации зон геодинамического риска на руднике, на горизонте, на участке. При этом:

- объём трудозатрат на создание электронной карты горных выработок варьируется в зависимости от детальности и формата исходных планов горных работ;
- программирование моделей не обязательно потребует значительных трудозатрат, в то время как разработка и проверка адекватности таких моделей занимает значительное время.

В дальнейшем предполагается совершенствование предложенного алгоритма путем добавления различных параметров массива с целью более адекватной визуализации геодинамических зон и, соответственно, выделении в них зон геодинамического риска.

Использование модулей QGIS позволит визуализировать результаты мониторинга не только для локальных моделей отдельных участков или горизонтов разрабатываемых месторождений, но и в целом шахтных полей близко расположенных горнодобывающих предприятий.

Литература

1. Маслобоев, А.В. Информационное измерение региональной безопасности в Арктике. / Маслобоев, А.В., Путилов В.А. - Апатиты: КНЦ РАН, 2016. - 222 с.
2. Яковлев, С.Ю. Информационная технология поддержки управления промышленно-экологической безопасностью Арктического региона / Яковлев С.Ю., Путилов В.А., Маслобоев А.В. // Надежность и качество: Труды международного симпозиума (г.Пенза, 21-31 мая 2018 г.). – Пенза: Изд-во ПГУ, 2018. – Т. 1. - С. 53-59.
3. Мельников, Н.Н. Техногенная сейсмичность – опасный антропогенный фактор при ведении горных работ в высоконапряженных массивах / Н.Н. Мельников, А.А. Козырев, В.И. Панин // Геоэкология. Инженерная геология. Гидрогеология. Геокриология. – 2015. - №5. – С. 425-433.
4. Федотова, Ю.В. Оценка уровня геодинамической безопасности на апатитовых рудниках / Федотова Ю.В., Панин В.И. // ГИАБ. - 2016. - №4. - С.344-353
5. Козырев, А.А. Мониторинг природных и техногенных процессов при ведении горных работ в высоконапряженных скальных массивах: Монография / Козырев, А.А., Каспарьян Э.В., Федотова Ю.В. - LAP Lambert Academic Publishing RU, Saarbrucken, Deutschland, 2016. - 161с. ISBN 978-3-659-97138-9

6. Яковлев, С.Ю. Проблема учёта неопределённости при управлении региональной безопасностью / С.Ю. Яковлев // Труды Кольского научного центра РАН. - 3/2017(8). Информационные технологии. – Вып. 8. – Апатиты: Изд-во КНЦ РАН, 2017. – С. 54 – 61.
7. Кадочников, А.А. Программное обеспечение информационно-аналитических систем на основе геоинформационного интернет-сервера / А.А. Кадочников // Вычислительные технологии. Т.12. – 2007. - №S2. – С. 70-78.
8. Овсянников, А.Н. Районирование карьерного поля по взрываемости горных пород (на примере рудника «Железный» ОАО «Ковдорский ГОК») / А.Н. Овсянников, И.М. Игнатенко, В.А. Дунаев // Научные ведомости Белгородского Государственного Университета. серия: естественные науки. – 2015. – №15(212). – С. 172-177.
9. Варванович, Н.Н. Автоматизированное ведение горных работ в программном комплексе МАЙНФРЭЙМ / Н.Н. Варванович, Л.С. Ломако //Маршейдерский вестник. – 2009. - №3(71). – С. 30-35.
10. Ожигина, С.Б. Создание пространственной модели прибортового массива карьера при помощи ГИС ArcGIS / Ожигина С.Б., Жумабеков Т.А., Сиренко Г.В., Ожигин Д.С. // Интерэкспо Гео-Сибирь. Т.1. – Новосибирск: СГУГиТ. – 2008. - №2. – С. 129 – 134
11. Писаренко, М.В. Алгоритм расчета ожидаемых сдвижений и деформаций земной поверхности с использованием ГИС технологий / М.В. Писаренко, И.Л. Борисов // Вестник Кузбасского Государственного Технического Университета. – 2008. - №2(66). – С. 127-130.
12. QGIS – свободная географическая информационная система с открытым кодом. – Режим доступа: <https://qgis.org/ru/site/>
13. Пространственный анализ векторных данных (Буфер). – Режим доступа: https://docs.qgis.org/2.18/ru/docs/gentle_gis_introduction/vector_spatial_analysis_buffers.html

Сведения об авторах

Шемякин Алексей Сергеевич – младший научный сотрудник

e-mail: shemyakin@iimm.ru

Alexey S. Shemyakin – junior researcher

Федотова Юлия Викторовна – к.т.н., старший научный сотрудник

e-mail: julia@goi.kolasc.net.ru

Iuliia V. Fedotova – PhD (Eng.), senior researcher

