



Российская Академия Наук

ТРУДЫ

Кольского научного центра РАН

9/2019 (10)

**ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ**

выпуск 9

Апатиты
2019

0+

9/2019 (10)

Российская Академия Наук

0+

УДК 551.46
ISSN 2307-5252

ТРУДЫ

Кольского научного центра

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

выпуск 9

DOI: 10.25702/KSC.2307-5252.2019.9

Научно-информационный журнал

Основан в 2010 году
Выходит 11 раз в год

Учредитель — Федеральное государственное бюджетное учреждение науки Федеральный исследовательский центр «Кольский научный центр Российской академии наук»

Свидетельство о регистрации СМИ
ПИ № ФС77-58457 от 25.06.2014

выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций.

Главный редактор, председатель Редакционного совета
С. В. Кривовичев, чл.-корр. РАН, д. г.-м. н., проф.

Заместитель главного редактора
В. К. Жиров, чл. - корр.

Редакционный совет:
академик Г. Г. Матишов,
чл. - корр. А. И. Николаев,
д. э. н. Ф. Д. Ларичкин
д. т. н. В. А. Путилов,
д. ф. - м. н. Е. Д. Терещенко,
к. т. н. А. С. Карпов (отв. секретарь)

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

Научное издание

Технический редактор В. Ю. Жиганов
Подписано к печати 10.12.2019.
Дата выхода в свет 20.12.2019.
Формат бумаги 70×108 1/16.
Усл. печ. л. 17.24. Заказ № 45. Тираж 300 экз.
Свободная цена.

Адрес учредителя, издателя и типографии:
Федеральное государственное бюджетное учреждение науки
Федеральный исследовательский центр
«Кольский научный центр РАН»
184209, г. Апатиты, Мурманская обл.,
ул. Ферсмана, 14
Тел.: (81555) 7-53-50; 79-5-95, факс: (81555) 76425
E-mail: ksc@ksc.ru. www.ksc.ru, naukaprint.ru

СОДЕРЖАНИЕ

	Стр.	
Введение.....	7	
Маслобоев А. В. Путилов В. А.	Моделирование многоуровневых распределенных систем сетцентрического управления региональной безопасностью.....	9
Быстров В. В. Халиуллина Д. Н.	Разработка системы информационного мониторинга и индикаторной оценки угроз экологической безопасности арктических коммуникаций.....	32
Вицентий А. В. Шишаев М. Г. Гогоберидзе Г. Г.	Архитектура мультипредметной системы информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности.....	42
Вицентий А. В. Шишаев М. Г.	Разработка подхода к модификации и редуцированию геоизображений с учетом их информативности.....	52
Шестаков А. В. Шемякин А. С. Яковлев С. Ю.	Трёхмерное моделирование в задачах расчёта и визуализации региональных техногенно-экологических рисков.....	59
Быстров В. В. Халиуллина Д. Н. Малыгина С. Н.	Разработка формальных моделей и средств сетцентрического управления кадровой безопасностью региона.....	69
Малоземова М. Л. Ломов П. А.	Технология пополнения тезауруса на основе применения дистрибутивного подхода к анализу естественно-языковых текстов.....	84
Диковицкий В. В. Шишаев М. Г. Пимешков В. К.	Метод автоматизированного извлечения понятий и парадигматических отношений тезауруса из текстов на естественном языке на базе лексико-синтаксических шаблонов.....	91
Фридман О. В.	Логические нейронные сети: методы автоматического конструирования, редукции, извлечения правил.....	97
Олейник Ю. А. Зуенко А. А.	Анализ возможностей совмещения парадигм программирования в ограничениях и объектно-ориентированного программирования.....	109

Зуенко А. А. Зуенко О. Н.	Применение методов декларативного программирования в задачах constrained clustering.	116
Македонов Р. А. Зуенко А. А.	Декларативный подход к планированию закупок малым предприятием.....	126
Ломов П. А.	Применение графовых субд в задачах анализа данных.....	137
Шишаев М. Г. Вицентий А. В. Куприков Н. М.	Концепция национальной системы управления данными: современный контекст реализации.....	146
Щур А. Л. Федоров А. М. Датьев И. О.	Цифровизация в государственном и муниципальном управлении: мировой опыт, проблемы и перспективы использования онлайн-социальных сетей.....	158
Каржавин В. К.	Математическое моделирование геохимических систем и перспективных схем переработки минералов.....	172
Нифантов В. М.	Диагностика, оценка и прогнозирование технического состояния технологического оборудования при помощи нечеткой экспертной системы в централизованных системах технического обслуживания и ремонта.....	187

9/2019 (10)

UDC 551.46
ISSN 2307-5252

Russian Academy of Sciences

0+

TRANSACTIONS

Kola Science Centre

Editor-in-Chief: S. V. Krivovichev,
Corr. Member of the RAS, Prof.

Deputy Editor-in-Chief:
V. A. Masloboev, Dr. Sc. (Engineering)

Editorial Council:

G. G. Matishov, Acad. of RAS,
A. I. Nikolaev, Cor. Member of RAS,
F. D. Larichkin, Dr. Sc. (Economics),
V. A. Putilov, Dr. Sc. (Engineering),
E. D. Tereshchenko, Dr. Sc.
(Physics and Mathematics),
A. S. Karpov PhD (Engineering) –
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.2307-5252.2019.9

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>

CONTENTS

	Page
Introduction	7
Masloboev A. V. Putilov V. A.	9
Modeling of multi-level distributed systems for regional security network-centric control.....	
Bystrov V. V. Khaliullina D. N.	32
Development of information monitoring system and indicator assessment of threats to environmental safety of arctic communications.....	
Vicentiy A. V. Shishaev M. G. Gogoberidze G. G.	42
Architecture of a multi-subject system of information support of the solution of regional management problems taking into account contradictions of spatial planning of economic activity.....	
Vicentiy A. V. Shishaev M. G.	52
Development of the modification and reduction of geoimages approach taking into account its informativeness.....	
Shestakov A. V. Shemyakin A. S. Yakovlev S. Yu.	59
3D modeling in tasks of calculation and visualization of regional technogenical-ecological risks.....	
Bystrov V. V. Khaliullina D. N. Malygina S. N.	69
Development of formal models and means for network-centric management of personnel security in the region.....	
Malozemova M. L. Lomov P. A.	84
Thesaurus replenishment technology based on the application of a distributive approach to analysis of natural language texts.....	
Dikovitskiy V. V. Shishaev M. G. Pimeshkov V. K.	91
Method of automated extraction of concepts and paradigmatic relations of thesaurus from texts in natural language on the basis of lexico-syntactic templates.....	
Fridman O. V.	97
Logic neural networks: methods of automatic design, reduction, rule extraction.....	
Oleynik Y. A. Zouenko A. A.	109
Possibility analysis of constraint and object-oriented programming paradigms combining.....	
Zuenko A. A. Zuenko O. N.	116
Application of declarative programming methods in the constrained clustering problems.....	

Makedonov R. A. Zuenko A. A.	Declarative approach to procurement planning by a small enterprise.....	126
Lomov P. A.	Application of graph data bases for data analysis.....	137
Shishaev M. G. Vicentiy A. V. Kuprikov N. M.	Concept of a national data management system: modern implementation context.....	146
Shchur A. L. Fedorov A. M. Datyev I. O.	Digitalization in state and municipal governance: world experience, problems and prospects for using online social networks.....	158
Karzhavin V. K.	Mathematical modeling of geochemical systems and perspective schemes for processing minerals.....	172
Nifantov V. M.	Diagnostics, assessment and forecasting of technical condition of technological equipment by using fuzzy expert system in centralized maintenance and repair systems.....	187

INTRODUCTION

The research and development results presented in the tenth issue of the "Information Technology" series of Proceedings of the Kola Science Center of the Russian Academy of Sciences, cover a wide range of urgent problems in the field of modeling complex systems and decision support - from theoretical aspects to applied technologies.

The practical tasks to which the studies of the authors of the papers are oriented are largely determined by the specifics of systems and objects localized in the Arctic zone of the Russian Federation (AZRF).

Interdisciplinary consideration of problems, coordination of activities and decisions of various management entities are needs for sound decisions in the field of managing the functioning and development of various systems in the Russian Arctic. At the same time, management entities can have different departmental affiliations, are often geographically distributed, and their "personal" interests can vary significantly. A number of authors propose solutions based on a network-centric approach with combined use of various modeling technologies, first of all, multi-agent ones to provide informational and analytical support for management under such conditions.

The complexity of the objects of control and modeling requires the inclusion in to consideration of factors and characteristics of various nature, the joint processing of quantitative and qualitative data, and the increase in cognitive presentation of the results of computer information processing. To do this, as a rule, methods, models and technologies are used that relate to various areas of artificial intelligence. A number of them are presented in this issue. These are models of neural networks, methods for the automated knowledge extraction from heterogeneous sources, including texts in natural language, methods of cognitive visualization that allow automated generation of problem-oriented geomaps taking into account the "interests" of users. The papers devoted to the development and use of constraints programming methods are also in the issue.

In the context of the "Digital Economy of the Russian Federation" program, in this issue presents the main provisions of the national system for managing data and features of the Arctic zone of the Russian Federation as a test region for implementing practical solutions to this system. The process of developing ideas of e-democracy in the world is considered in one of the articles. The specific features of the use of online social networks at various levels of government highlights in the article

Traditionally, the issue includes articles on the use of modeling and information technology in the development and support of the functioning of technological processes.

The "Information Technologies" series is addressed to specialists in the field of modeling, as well as the development, creation and practical use of systems and technologies for information support of decision-making in various areas of management and production, to lecturers and students of relevant specialties.

ВВЕДЕНИЕ

Результаты исследований и разработок, представленные в десятом выпуске серии «Информационные технологии» Трудов Кольского научного центра РАН, охватывают достаточно широкий спектр актуальных задач в области моделирования сложных систем и поддержки принятия решений — от теоретических аспектов до прикладных технологий.

Практические задачи, на решение которых ориентированы исследования авторов работ, во многом определяются спецификой систем и объектов, локализованных в Арктической зоне Российской Федерации (АЗРФ).

Обоснованные решения по управлению функционированием и развитием систем различной природы в АЗРФ требуют использования междисциплинарного рассмотрения проблем, координации деятельности и согласования решений различных субъектов управления. При этом субъекты управления могут иметь различную ведомственную принадлежность, зачастую территориально распределены и их «личные» интересы могут существенно различаться. Для информационно-аналитической поддержки управления в таких условиях ряд авторов предлагает решения, основанные на сетцентрическом подходе с комбинированным использованием различных технологий моделирования и, в первую очередь, мультиагентных.

Сложность объектов управления и моделирования требует включения в рассмотрение факторов и характеристик различной природы, совместной обработки количественных и качественных данных, повышения когнитивности представления результатов компьютерной обработки информации. Для этого, как правило, используются методы, модели и технологии, относящиеся к различным направлениям искусственного интеллекта. Ряд из них представлен в настоящем выпуске. Это модели нейронных сетей, методы автоматизированного извлечения знаний из разнородных источников, включая тексты на естественном языке, методы когнитивной визуализации, позволяющие в автоматизированном режиме формировать проблемно-ориентированные геоизображения с учетом «интересов» пользователей. Также в выпуске представлены работы, посвященные развитию и использованию методов программирования в ограничениях. В контексте программы «Цифровая экономика Российской Федерации» в настоящем выпуске представлены основные положения концепции Национальной системы управления данными и особенности Арктической зоны РФ как тестового региона реализации практических решений этой системы. В одной из работ рассмотрен процесс развития идей электронной демократии в мире, выделены специфические особенности применения онлайн-социальных сетей на различных уровнях управления.

Традиционно в сборник включены работы по использованию моделирования и информационных технологий в области разработки и поддержке функционирования технологических процессов.

Выпуски серии «Информационные технологии» адресованы специалистам в области моделирования, а также разработки, создания и практического использования систем и технологий информационной поддержки принятия решений в различных сферах управленческой и производственной деятельности, преподавателям и студентам вузов соответствующих специальностей.

А. В. Маслобоев, В. А. Путилов

Институт информатики и математического моделирования ФИЦ КНЦ РАН

МОДЕЛИРОВАНИЕ МНОГОУРОВНЕВЫХ РАСПРЕДЕЛЕННЫХ СИСТЕМ СЕТЕЦЕНТРИЧЕСКОГО УПРАВЛЕНИЯ РЕГИОНАЛЬНОЙ БЕЗОПАСНОСТЬЮ*

Аннотация

В работе представлены результаты математического моделирования систем сетецентрического управления региональной безопасностью как многоуровневых распределенных систем. Для решения задач координации децентрализованного принятия решений в сетецентрической системе управления региональной безопасностью разработана и исследована агентная многоуровневая рекуррентная иерархическая модель управления безопасностью региона. В модели совмещаются координация путем развязывания взаимодействий и создание коалиций между агентами на разных уровнях управления. Этим обеспечивается сокращение времени на выработку и реализацию решений системой организационного управления региональной безопасностью. Рассмотрены приложения созданной модели на примере решения задач координации управления экологической безопасностью Северного морского пути.

Ключевые слова:

координация, сетецентрическое управление, региональная безопасность, поддержка принятия решений, многоуровневая распределенная система, рекуррентная модель.

A.V. Masloboev, V.A. Putilov

MODELING OF MULTI-LEVEL DISTRIBUTED SYSTEMS FOR REGIONAL SECURITY NETWORK-CENTRIC CONTROL

Abstract

The paper represents work-out and research results of mathematical modeling of the regional security network-centric control systems as the multi-level distributed systems. For decentralized decision-making coordination problem-solving within the regional security network-centric control system an agent-based multi-level recurrent hierarchical model for security management of the region has been developed and analyzed. The model implements combined coordination techniques by unleashing interactions and by coalition formation between the agents at the different control levels of the system. That provides time reduction of decision-making and realization by the regional security organizational management system. Model applications by the example of Northern Sea Route environmental safety management coordination problem-solving are considered.

Keywords:

coordination, network-centric control, regional security, decision-making support, multi-level distributed system, recurrent model.

Введение

Важной задачей в управлении безопасностью региональных социально-экономических систем является определение эффективности существующей системы обеспечения региональной безопасности. На практике эффективность этой системы зачастую определяется согласованностью принимаемых

*Работа выполнена при поддержке Министерства науки и высшего образования РФ (тема НИР № 0226-2019-0035) и Российского фонда фундаментальных исследований — проект № 18-07-00167-а.

управленческих решений на всех уровнях управления развитием региона. При этом задача координации процессов принятия решений в сфере региональной безопасности осложняется не только децентрализованным характером организационного управления безопасностью, но и тем, что аспектов региональной безопасности, собственно как и субъектов ее обеспечения, много, и они очень разноплановые.

В условиях динамичности и распределенности элементов социально-экономической среды региона задача координации заключается в достижении согласованности в работе всех звеньев системы управления региональной безопасностью путем установления рациональных связей между ними, что обеспечивает получение оптимального решения общей задачи системы при оптимизации подзадач, решаемых подсистемами. Если в многоуровневой системе обеспечения региональной безопасности не поддерживается требуемый (допустимый для каждой конкретной ситуации) уровень согласованности (координируемости) управлений, то выявляются недостатки этой системы и формируется перечень рекомендаций по их устранению. Реализация этих рекомендаций сопровождается реконfigurацией системы и ее адаптацией под текущую обстановку в регионе.

Задача координации децентрализованного принятия решений по управлению социально-экономическими системами широко рассматривалась в самых разных аспектах, как в отечественных, так и в зарубежных исследованиях. Однако, отметим, что для области региональной безопасности это относительно новая задача, общая идея решения которой предлагалась авторами ранее в монографии [1]. Несогласованность в распределенной системе организационного управления региональной безопасностью снижает оперативность и качество принимаемых управленческих решений. Централизованное управление безопасностью региона в нынешних условиях является малоэффективным и не обеспечивает нужного эффекта. Выходом из этой проблемной ситуации является переход на модель сетевидного управления региональной безопасностью. Такой подход наиболее адекватно отражает реальную природу управления социально-экономическими системами и учитывает децентрализованный характер процессов обеспечения региональной безопасности, как по функциональной структуре, так и по составу участников. Сетевидное управление региональной безопасностью заключается в реализации сетевой структуры организационного управления с выделенными управляющими центрами, взаимодействие между которыми осуществляется на базе их интеграции в единое региональное информационное пространство. Для регионов России это предполагает отказ от закрытых иерархических структур управления безопасностью с жесткими организационными связями и централизованным управлением и переход к открытым сетевым виртуальным организационным структурам с гибкими связями и децентрализованным управлением.

Показатели региональной безопасности различны по природе и содержанию и, как правило, оптимизируются различными элементами многоуровневой системы управления региональной безопасностью. При этом эти показатели являются параметрами локальных целевых функций субъектов управления безопасностью и определяют их предпочтения и сферы интересов. В условиях децентрализованного управления безопасностью региона, многокритериальности решаемых задач и различий в целеполагании разнородных

субъектов управления это обуславливает необходимость координации процессов принятия управленческих решений по оптимизации и согласованию этих показателей на разных уровнях управления.

Для решения этой задачи в работе предлагается агентная многоуровневая рекуррентная иерархическая модель региональной безопасности. Специфика модели заключается в использовании функционально-целевой технологии [2] и математического аппарата теории иерархических многоуровневых систем [3] для реализации процедур согласования локальных решений сетецентрического управления. Модель обеспечивает координацию сетецентрического управления региональной безопасностью за счет удовлетворения требований взаимосвязи между целевыми функциями элементов многоуровневой системы обеспечения региональной безопасности на разных уровнях управления.

1. Многоуровневая рекуррентная иерархическая модель управления региональной безопасностью

Рекуррентная модель управления региональной безопасностью, построенная на базе функционально-целевого подхода [2], является основой формализации задач структурно-алгоритмической организации средств информационно-аналитической поддержки в этой предметной области и методов решения этих задач. Специфика рекуррентной модели определяется, во-первых, тем, что при построении модели целевого управления использована иерархия двухоперационных алгебр цепочек целей и совершенно аналогичных по структуре цепочек действий, обеспечивающих достижение этих целей. Во-вторых, иерархия целей в модели непосредственно порождает модель иерархии действий, что обеспечивает использование иерархии целей не только в качестве средства описания задачи, но и как средства проектирования системы информационной поддержки управления.

В многоуровневых распределенных системах такое понятие, как «цель», тесно связано с понятиями «принятие решений» [4] и «системы принятия решений» [5]. Целенаправленное поведение, в сущности, представляет собой последовательность принимаемых и реализуемых решений. Цели определяются через решаемые задачи.

Разработанная модель основана на иерархической структуре задач управления региональной безопасностью и использует последовательно-параллельные композиции целей управления и действий по достижению этих целей. Иерархическая структура заложена в самом понятии региональной безопасности, заведомо образуемой различными по характеру, но взаимосвязанными составляющими: экономической, экологической, социальной и другими. Каждая из составляющих региональной безопасности образуется набором объектов, субъектов, процессов и методов обеспечения безопасности, угроз и источников опасностей. Такая детализация проводится до нижнего уровня — уровня примитивов.

Процесс последовательной детализации задач управления региональной безопасностью представляется деревом декомпозиции целей управления. Доказанная в работе [2] теорема о подсистемах многоуровневой системы показывает, что система в целом должна строиться из таких подсистем, которые обеспечивают покрытие соответствующих подзадач основной целевой задачи

многоуровневой системы. Из теоремы следует, что синтез структуры системы должен проводиться изоморфно построению глобальной цели из некоторой совокупности подцелей.

Иерархическое представление систем используется в разных приложениях, в том числе и для многоуровневых систем управления. Это объясняется простотой и наглядностью иерархических моделей, хорошо отражающих реальные взаимосвязи в окружающем нас мире, включая организации людей. Как показывает практика, практически любую сложную распределенную систему управления можно свести к иерархической. Примеры синтеза иерархических структур управления представлены, например, в работе [4]. Существуют и другие доводы в пользу иерархических многоуровневых систем [2, 3, 6]:

- для решения общей задачи системы могут эффективно использоваться ограниченные возможности подсистем;
- эти системы появляются при интеграции уже созданных систем;
- системы лучше адаптируются к изменениям и усложнениям задач и обладают хорошими показателями надежности (неисправности в работе какой-либо подсистемы не всегда распространяются на всю систему).

В применении к задачам синтеза комплексных автоматизированных систем управления региональной безопасностью на практике, как правило, используются модели в виде графа с произвольной сетевой структурой [7], нечеткие [8], когнитивные [9] и имитационные [1] модели опять же на базе таких графов. Для многоуровневых распределенных систем такие модели уже малопригодны, так как приводят к сложным моделям в виде системы вложенных графов произвольной структуры. Идеологически близкий к настоящей работе программно-целевой подход [10] предполагает использование моделей не древовидной структуры, что объясняется спецификой предметной области, где непременно должны присутствовать связи между элементами одного уровня. При решении задач в программно-целевом подходе используется достаточно сложный аппарат траекторной оптимизации.

В отличие от известных типов графовых моделей созданная многоуровневая рекуррентная модель формируется регулярным рекуррентным применением к процессу декомпозиции основной целевой задачи двухуровневой иерархической структуры, имеющей один элемент на верхнем уровне и заданное моделью предметной области число элементов нижнего уровня. Этим обеспечивается учет особенностей структуры предметной области и решаемых задач. Такой простой вид этой (элементарной) модели и регулярные правила построения модели на базе элементарной обеспечивают как формальную постановку и решение, так и практическую реализацию задач синтеза структуры автоматизированной системы и алгоритмов управления региональной безопасностью, оптимальных в смысле заданных критериев эффективности функционирования этой системы.

Перейдем к рассмотрению формальной рекуррентной иерархической модели управления региональной безопасностью. Макроструктура многоуровневой системы управления региональной безопасностью, построенная изоморфно декомпозиции основной целевой задачи обеспечения региональной безопасности, представляется в виде дерева. Корню дерева ставится

в соответствие подсистема верхнего уровня (собственно система), вершинам дерева, отстоящим от корня на одно ребро, – подсистемы, реализующие классы безопасности, на три ребра, — подсистемы, реализующие методы и средства обеспечения безопасности, и т.д.

Отображение структуры задач предметной области на структуру программно-аппаратных средств позволяет определить набор элементарных компонентов для структурно-алгоритмического синтеза системы информационно-аналитической поддержки управления региональной безопасностью. Для этих целей в рамках модели определен алфавит функциональных операций (множество атомарных действий системы) и разработаны процедуры вывода последовательно-параллельных комбинаций этих действий, обеспечивающих решение задач вышележащих уровней, вплоть до корня дерева.

Синтез адекватной системы - трудоемкая задача, связанная с необходимостью удовлетворения условий изоморфизма на всех соответствующих уровнях декомпозиции задачи и организации системы, причем эти требования должны удовлетворяться для любой задачи рассматриваемой предметной области, что находится в противоречии с требованием гибкости системы по отношению к описанию предметной области. Поэтому для практических приложений актуальна задача синтеза покрывающих систем [2], обеспечивающих решение задач субъекта управления на всех уровнях организации системы с удовлетворяющими его значениями параметров качества цепочек действий. При этом по известным параметрам атомарных элементов нижнего уровня макроструктуры строятся отображения алгебры цепочек на алгебры соответствующих параметров «снизу вверх» до уровня иерархии системы, на котором субъект управления может принять решение либо о целесообразности использования синтезированной системы, либо о необходимости изменения постановки задачи или коррекции программно-аппаратного обеспечения системы с целью изменения параметров атомарных элементов.

Любой элемент M макроструктуры системы характеризуется состоянием S , управляющим воздействием U для задания режима работы элемента и его состояния, входной информацией W . Поскольку элементы макроструктуры — это программы, ориентированные на получение, обработку и передачу информации, то результатом работы элемента M является некоторая выходная информация V . Рассмотрим результирующую информацию V как некоторую функцию от состояния элемента макроструктуры M , входной информации и управляющего воздействия:

$$V = M(U, S, W). \quad (1)$$

Под элементарной неделимой единицей алгоритма управления безопасностью условимся понимать функциональную операцию L – некоторую совокупность действий исполнительской системы, зависящих от управляющего воздействия, состояния элемента макроструктуры и его внутренней структуры:

$$L = M(U, S). \quad (2)$$

Функциональные операции выполняют преобразование входной информации W в выходную следующим образом:

$$L_{MUS} : V = L_{MUS}(W). \quad (3)$$

Считывание состояния элемента макроструктуры достигается подачей специального управляющего сигнала U_0 :

$$S = L_{MUS_0} = (0). \quad (4)$$

Таким образом, определен алфавит функциональных операций (множество атомарных действий системы):

$$L = M \times U \times S. \quad (5)$$

Представим содержательную информацию рассматриваемой предметной области (региональной безопасности) в виде формальных высказываний. Построим алгебраическую систему $A = \langle A, Q, R \rangle$, состоящую из непустого множества A , семейства алгебраических операций Q и семейства отношений R . Для задания такой системы определим некоторые исходные объекты, которые будем рассматривать как неделимые; перечислим способы комбинирования исходных объектов между собой; укажем условие, которому удовлетворяют те и только те комбинации исходных объектов, которые считаются элементами системы; сформулируем условие, при котором два элемента системы считаются равными.

Отождествим семейство Q алгебраических операций с организацией достижений совокупной цели управления из известных атомарных целей, достижение которых реализуется атомарными действиями, заданными алфавитом L . Такие совокупные цели обеспечиваются комбинациями последовательного и параллельного (одновременного) достижения атомарных целей, то есть композициями элементов функционального алфавита целей, построенными с использованием двух обобщенных операций:

1. операция \otimes : достичь атомарной цели a_2 после достижения атомарной цели a_1 ;
2. операция \oplus : достичь атомарной цели a_2 одновременно с атомарной целью a_1

Использование принципа управления через целеполагание обеспечивает организацию всего многообразия вариантов обеспечения региональной безопасности через композиции элементов функционального алфавита, построенные с использованием двух введенных обобщенных операций. Совокупная цель достигается последовательно-параллельной комбинацией подцелей нижнего уровня.

Аналогичные операции вводятся для атомарных действий – элементов функционального алфавита L . Зададим операцию \otimes как последовательное применение следующей функциональной операции к результату предыдущей:

$$\otimes : L_j = L_{j-2} \otimes L_{j-1} \rightarrow L_j(W) = L_{j-1}(L_{j-2}(W)). \quad (6)$$

Операцию \oplus зададим как одновременное выполнение двух атомарных воздействий:

$$\oplus : L_j = L_i \oplus L_k \rightarrow L_j(W) = \begin{cases} L_j(W_i) \\ L_j(W_k) \end{cases}. \quad (7)$$

Операция \otimes производит последовательный запуск и исполнение выбранных атомарных элементов вычислительного процесса. Операция \oplus производит параллельный запуск выбранных атомов.

В работе [2] проведены исследования полученной алгебры строк (цепочек), определены свойства замкнутости, ассоциативности, коммутативности относительно введенных операций \oplus и \otimes , а также установлено наличие нулевого, единичного и обратных элементов. Не теряя общности, ограничим рассмотрение алгеброй действий, в которой нагляден физический смысл введенных обобщенных операций. Полученные результаты справедливы и для алгебры целей.

Зададим на алгебре цепочек A некоторое отношение эквивалентности R . Отношение эквивалентности может задаваться как совпадение параметров цепочек (например, длины или используемых операций), либо как совпадение параметров результата, то есть при одинаковой входной информации в результате выполнения двух разных цепочек получаем результирующую информацию, принадлежащую в обоих случаях к одному некоторому множеству.

Известно, что заданное некоторым образом отношение эквивалентности R разбивает все множество цепочек на множество непересекающихся классов эквивалентности. Исходя из этого, все семантически одинаковые цепочки находятся в пределах одного класса эквивалентности. Классы эквивалентности $\{z_i\}$ характеризуются следующими соотношениями:

$$\{z_i\}: \begin{cases} 1) \forall a_1, a_2 \in z_i, a_1 R a_2 \\ 2) \bigcup_i z_i = A \\ 3) z_i \cap z_j = \begin{cases} z_i, & i = j \\ \emptyset, & i \neq j \end{cases} \end{cases} . \quad (8)$$

В каждом классе эквивалентности задается новое отношение эквивалентности, разбивающее каждый класс эквивалентности на подклассы, и т.д. В результате получается семейство алгебр классов эквивалентности

$$A^k = \langle \Sigma^k, \{\otimes, \oplus\} \rangle, \quad (9)$$

где Σ^k — множество цепочек над алфавитом $\{z_{j^k}\}$.

Таким образом, строятся модели декомпозиции целей управления на комплексной предметной области и декомпозиции действий соответствующей автоматизированной системы управления, обеспечивающих достижение этих целей. Они получены абстрагированием от конкретного содержания составляющих предметных областей и заменой их понятием классов эквивалентности функций (целей или действий в зависимости от приложения модели), то есть множеств функций, эквивалентных в смысле их предметной направленности. В каждом классе эквивалентности задано новое отношение эквивалентности, относящее функции к разным поднаправлениям и разбивающее каждый класс эквивалентности на подклассы. Рекуррентный процесс детализации исходной функции продолжается вплоть до достижения уровня «примитивов» — элементарных функций, неделимых с точки зрения субъекта управления. Задание множества отношений эквивалентности функций определяет топологию на множестве функций. Базой этой топологии является множество примитивов.

Полученная декомпозиция предметной области представляется древовидным графом иерархии классов, в котором узлы – имена классов, ребра — отношения включения, корень — имя функции на комплексной предметной области, листья – примитивы:

$$\left\{ z_{\vec{j}^k}^k \right\}_{k=1}^K : z_{\vec{j}^k}^k = \bigcup_{j_{k+1}} z_{\vec{j}^{k+1}}^{k+1}$$

$$z_{\vec{j}^k}^k = \bigcap z_{\vec{j}^k}^k = \begin{cases} z_{\vec{j}^k}^k, \vec{i}^k = \vec{j}^k \\ \emptyset, \vec{i}^k \neq \vec{j}^k \end{cases}, \quad (10)$$

$$\vec{j}^{k+1} = (\vec{j}^k, j_{k+1}), \vec{j}^1 = 1, k = \overline{1, K}$$

где $\left\{ z_{\vec{j}^k}^k \right\}_{k=1}^K$ — множество классов эквивалентности; K — число уровней декомпозиции; k — индекс уровня декомпозиции; $\vec{j}^k = \{j_i\}, i = \overline{1, K}$ — вектор-индекс длиной k класса эквивалентности на k -ом уровне декомпозиции; $j_i, i = \overline{1, K}$ — i -й компонент вектор-индекса; $z_{\vec{j}^k}^k$ — имя класса на k -ом уровне декомпозиции с вектор-индексом \vec{j}^k .

Система (10) порождается системой отношений эквивалентности:

$$\left\{ R_{\vec{j}^k}^k \right\}_{k=1}^{K-1} : \forall j_{k+1}, \forall x, y, x R_{\vec{j}^k}^{k+1} y \Rightarrow x R_{\vec{j}^k}^k y$$

$$\forall \vec{j}^k \exists j_{k+1} : x R_{\vec{j}^k}^k y \Rightarrow x R_{\vec{j}^{k+1}}^{k+1} y, \quad (11)$$

$$\forall \vec{j}^k, \forall \vec{i}^k : \vec{i}^k \neq \vec{j}^k : x R_{\vec{i}^k}^k y \Rightarrow \neg x R_{\vec{j}^k}^k y$$

где $R_{\vec{j}^k}^k$ отношение эквивалентности, разбивающее $z_{\vec{j}^k}^k$ на $\left\{ z_{\vec{j}^{k+1}}^{k+1} \right\}, x, y \in z_{\vec{j}^k}^k$.

Построенная алгебраическая система A , состоящая из множества элементов, двух алгебраических операций и семейства отношений эквивалентности, является формальной моделью постановки и решения задач организации процесса управления региональной безопасностью, поскольку одинаковым образом описывает цели управления и действия по достижению этих целей на любом уровне декомпозиции исходной задачи.

В общем случае имеется множество классов эквивалентности

$$Z = \left\{ z_{\vec{j}^k}^k \right\}_{k=1}^K, \quad (12)$$

где K — число уровней декомпозиции.

Общая рекуррентная модель представляет собой иерархию алгебр

$$A^k = \langle \Sigma^k, \{ \otimes, \oplus \} \rangle, \quad (13)$$

гомоморфно отображенных друг на друга «снизу вверх»:

$$\gamma_k : A^{k+1} \rightarrow A^k, \quad (14)$$

где γ_k есть совокупность отношений $\left\{ R_{\vec{j}^k}^k \right\}$.

Таким образом, рекуррентная модель предметной области (региональной безопасности) получена на основе рекуррентной декомпозиции целей управления

и обеспечивает синтез структуры соответствующей автоматизированной системы управления в этой предметной области на основе адекватного декомпозиции целей управления процесса детализации действий по достижению этих целей.

Процедуры синтеза и анализа многоуровневых иерархических систем предполагают, что составляющие систему элементы обладают ограниченными возможностями по решению задач, стоящих перед системой. В связи с этим глобальная задача, отражающая назначение системы в целом, разбивается на совокупность подзадач таким образом, что решение глобальной задачи эквивалентно решению этой совокупности. Такой подход применяется как при проектировании структур многоуровневых систем, так и при организации решения системой задач [2, 3]. В многоуровневых системах, как иерархической, так и сетевой структуры при этом возникают специфические проблемы управления, связанные с координацией децентрализованного принятия решений на разных уровнях организации системы.

Разработанная модель обеспечивает формальную основу для синтеза программной исполнительской среды информационно-аналитической поддержки и координации сетевидного управления региональной безопасностью.

2. Координация сетевидного управления региональной безопасностью на базе рекуррентной модели

Система организационного управления региональной безопасностью, построенная на базе сети распределенных ситуационных центров, относится к классу многоуровневых сетевидных систем управления. Такие системы управления распределенными объектами (отдельными системами), согласно работам [1, 11], характеризуются свойствами открытости, самоорганизации, децентрализации функций управления и принятия решений, слабой иерархии в контуре принятия решений и способностью порождать цели внутри себя. Центральной задачей для данного класса систем является согласование и координация взаимодействия управляющих элементов и подсистем в условиях децентрализованного принятия решений. Главная задача координации [12] — достижение согласованности в работе всех звеньев системы путем установления рациональных связей (коммуникаций) между ними. Характер этих связей может быть различным, так как зависит от координируемых процессов.

Под *координацией* понимается свойство системы находить оптимальные решения общей задачи управления при оптимизации подзадач управления, решаемых подсистемами. Другими словами, координация означает такое воздействие элемента вышестоящего уровня на элементы нижестоящего уровня, которое заставляет нижестоящие элементы действовать согласованно. Для обеспечения координации необходимо определить ограничения на взаимосвязи между подсистемами. Выделяют два базовых способа координации [3]: координация путем изменения целей и координация путем изменения ограничений.

Известны различные методы координации управления в многоуровневых распределенных системах. К ним относятся игровые и градиентные методы [12], основанные на предложенных в [3] необходимых и достаточных условиях координируемости локально организованной иерархии динамических систем; методы координации, основанные на принципах самоорганизации на базе моделей вычислительных полей [13]; методы нечеткой параметрической

координации в многоуровневых иерархических системах [14]; методы координации путем прогнозирования, развязывания и оценки взаимодействий [3]; методы координации на основе моделей коалиционных рефлексивных игр [15] и другие [2, 16, 17]. Для конкретных приложений эти методы могут дополнительно комбинироваться.

В работе [18] рассматриваются методы координации группового поведения субъектов управления (агентов) в сложных системах, основанные на моделях командной работы, процедурах конструирования общих правил взаимодействия агентов и механизмах обмена информацией на метауровне. Эти методы обеспечивают достижение ситуации равновесия с заранее заданными свойствами группового поведения.

Предлагаемый подход к решению задачи координации сетецентрического управления региональной безопасностью основан на применении разработанной многоуровневой рекуррентной иерархической модели. В модели совместно реализуются процедуры координации путем развязывания взаимодействий и самоорганизации агентов в коалиции на разных уровнях управления. Другой отличительной особенностью подхода является комбинированное использование методов анализа совместимости решаемых задач, оценивания деловой репутации (показателей качества работы) субъектов управления, функционирующих на разных уровнях иерархии системы, оценки меры связности синтезируемых организационных структур управления и агентных механизмов самоорганизации в открытых децентрализованных системах. Эти методы подробно рассмотрены в работе [1].

Разработанная рекуррентная модель (12)–(14) получена регулярным рекуррентным применением к процессу декомпозиции основной целевой задачи двухуровневой структуры [3], имеющей один элемент на верхнем уровне и заданное моделью предметной области число элементов нижнего уровня. Вариант организации управления на основе такой модели и способ координации зависят от степени распределенности общей задачи системы между уровнями. Наиболее простое решение заключается в том, что элемент верхнего уровня (координатор) имеет точное описание поведения элементов нижнего уровня. Такая постановка приводит к обычным задачам дискретной оптимизации. Второй подход состоит в том, что задача координации решается с учетом взаимодействия семейства взаимосвязанных подсистем (элементов) нижнего уровня. Нижележащие элементы в иерархии управления рассматриваются как локально-организованные, автономные самостоятельные сущности (агенты). При этом полагается, что каждая из подсистем решает свою задачу и преследует свои цели, а формализация задачи координатора основывается на информации о том, каким образом элементы нижнего уровня при выборе своих решений учитывают эти взаимодействия. Этот вариант позволяет для решения задачи, стоящей перед всей системой в целом, использовать совокупность решающих элементов, расположенных на различных уровнях организации системы, даже если каждый элемент в отдельности (включая и координатора) не в состоянии решить общую задачу. Для решения общая задача разбивается на подзадачи, решение которых производится групповыми усилиями решающих элементов.

При этом возникает ряд проблем, связанных с необходимостью синхронизации времени решения задач нижестоящими элементами и существования некоторого плана достижения целевого состояния этими

элементами. Для решения этих проблем могут использоваться, например, методы сетевого планирования [19].

Основные этапы алгоритма решения задачи координации:

- 1) выбор принципа координации;
- 2) модифицирование функций и взаимосвязей (изменение целей (функций качества) элементов нижестоящего уровня или изменение ограничений, налагаемых на принимаемые ими решения);
- 3) выбор координирующих воздействий.

При этом важной проблемой, которую необходимо решать в процессе координации, является декомпозиция глобальной цели (задачи) системы на подцели (подзадачи) и их распределение между вышестоящим и нижестоящими решающими элементами так, чтобы они были согласованными, то есть должен выполняться *постулат совместимости*: если задача вышестоящего элемента координируется по отношению к нижестоящему, то она координируется по отношению к глобальной задаче системы.

В работе реализован способ координации путем развязывания взаимодействий. В этом случае координация реализуется путем изменения целей (функций качества) элементов нижестоящего уровня, так как связующие входы выбираются произвольно. При таком способе координации должны выполняться принципы согласования взаимодействий и функций качества, а также постулат совместимости целей (задач), решаемых элементами нижестоящих уровней, по отношению к глобальной цели (задаче) системы. В работе [3] доказано, что принцип согласования взаимодействий применим, если глобально оптимальное управляющее воздействие обеспечивается локальными решениями всякий раз, когда связующие входы согласованы. Принцип согласования функций качества применим, если глобально оптимальное управляющее воздействие складывается из локальных решений всякий раз, когда согласованы ожидаемые значения и фактические значения локальных функций.

Таким образом, система координируема на основе принципов согласования взаимодействий и функций качества, если принципы применимы и существует координирующий сигнал, обеспечивающий согласованность связующих входов, а также ожидаемых и фактических значений локальных функций соответственно.

Координатор на метуровне может влиять на работу децентрализованных элементов нижнего уровня посредством изменения коэффициентов в функциях качества функционирования нижележащих элементов системы (например, изменение тарифов, стоимостных показателей, законодательства, нормативных регламентов) или иным способом. Это делается для того, чтобы элементы нижележащего уровня функционировали согласованно. При этом предполагается, что связующие сигналы (входы и выходы) между элементами нижнего уровня совпадают.

Для формализации понятия координации в [3] вводится предикат $P(x, D)$:

$(\forall x, \forall D), [P(x, D) \equiv x \text{ есть решение } D]$,

где D — решаемая системой задача. Предикат $P(x, D)$ является истинным тогда и только тогда, когда D — решаемая задача, а x — одно из ее решений.

Пусть задачи, решаемые нижестоящими элементами, параметризуются только координирующими сигналами элемента вышестоящего уровня (рис. 1).

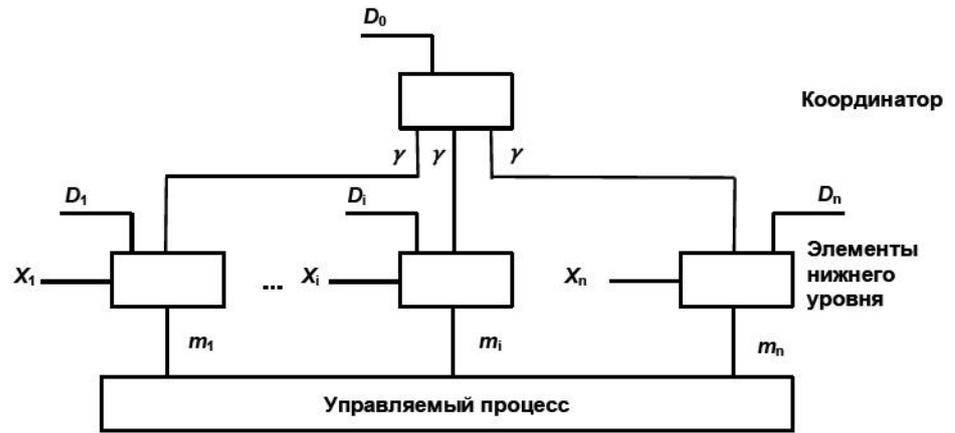


Рис. 1. Структура двухуровневой иерархической модели управления

Пусть D_0 — конкретная задача вышестоящего элемента. Каждый координирующий сигнал $\gamma \in \Gamma$ (Γ — множество координирующих сигналов) вышестоящего элемента конкретизирует задачу $D_i(\gamma)$, которую будет решать i -й элемент нижестоящего уровня. Пусть $\bar{D}(\gamma) = \{D_1(\gamma), \dots, D_n(\gamma)\}$ — совокупность таких задач (n — число элементов нижестоящего уровня). Задачи, решаемые элементами нижестоящего уровня, координируемы по отношению к задаче D_0 вышестоящего уровня тогда и только тогда, когда справедливо предложение:

$$(\exists \gamma), (\exists x), [P(x, \bar{D}(\gamma)) \& P(\gamma, D_0)], \gamma \in \Gamma, x \in X, \quad (15)$$

то есть когда имеет решение задача D_0 вышестоящего уровня и для координирующего сигнала γ , решающего данную задачу, имеется множество решений $\bar{D}_i(\gamma)$ задач нижестоящих элементов.

Предложению (15) эквивалентно следующее:

$$(\exists \gamma), (\exists x), [P(x, \bar{D}(\gamma)) \& Q_0(\gamma, x)], \quad (16)$$

вследствие того, что справедливо утверждение $P(\gamma, D_0) \Leftrightarrow (\exists x)[Q_0(\gamma, x)]$, где Q_0 — заданный предикат, определенный для всех пар:

$$(\gamma, x): (\gamma, x) \in \Gamma \times X; x = X_1 \times \dots \times X_n$$

Общая (глобальная) задача системы определяется по отношению ко всему управляемому системой процессу в целом, поэтому множество решений этой задачи есть множество глобальных управлений M . Будем рассматривать случай, когда глобальные управляющие сигналы из множества M , воздействующие на управляемый процесс в целом, исходят только из элементов нижнего уровня. Тогда эти сигналы представляются отображением: $\pi_M: X \rightarrow M$. Задачи, решаемые элементами нижестоящего уровня, координируемы по отношению к данной глобальной задаче D тогда и только тогда, когда справедливо предложение:

$$(\exists \gamma), (\exists x), [P(x, \bar{D}(\gamma)) \& P(\pi_M(x), D)], \gamma \in \Gamma, x \in X, \quad (17)$$

то есть элемент вышестоящего уровня имеет возможность влиять на элементы нижестоящего уровня таким образом, что их результирующее воздействие на управляемый процесс в целом дает решение глобальной задачи системы.

Рассмотрим совместимость в многоуровневых системах на примере двухуровневой системы [20], в которой имеются задачи трех типов: глобальная, для элемента вышестоящего уровня, для элементов нижестоящего уровня. Эти задачи должны быть определенным образом согласованы между собой (совместимы). Для этого необходимо, чтобы координация задач $D_i, i = \overline{1, n}$ элементов нижестоящего уровня относительно задачи D_0 вышестоящего уровня была соответствующим образом связана с глобальной задачей D . Постулат совместимости [3]:

$$(\forall \gamma), (\forall x), \{ [P(x, \overline{D}(\gamma)) \& Q_0(\gamma, x)] \Rightarrow [P(x, \overline{D}(\gamma)) \& P(\pi_M(x), D)] \}, \quad (18)$$

$$x \in X, \gamma \in \Gamma, D = \{D_i\}_{i=1}^n,$$

то есть решаемые элементами нижестоящего уровня задачи $D_i, i = \overline{1, n}$ скоординированы относительно глобальной задачи D всякий раз, когда задачи $D_i, i = \overline{1, n}$ скоординированы относительно задачи D_0 , решаемой элементом вышестоящего уровня. Если постулат (18) выполняется, задачи совместимы. Если задачи совместимы, то решение глобальной задачи D достигается тогда, когда элемент вышестоящего уровня координирует элементы нижестоящего уровня по отношению к решению собственной задачи.

При использовании способа координации путем развязывания взаимодействий успех в координации элементов нижестоящего уровня можно оценить, исходя из степени рассогласованности между фактическими взаимодействиями элементов нижестоящего уровня и теми, которые были бы желательны с точки зрения этих элементов. Связующие сигналы, выбираемые элементами нижестоящего уровня, задаются отображением $\pi_U : X \rightarrow U$, то есть $u \in U$ — это часть решения $x \in X$. Принцип согласования взаимодействий дается предложением [3]:

$$(\forall \gamma) (\forall x), \{ [P(x, D(\gamma)) \& K(\pi_M(x)) = \pi_U(x)] \Rightarrow P(\pi_M(x), D) \}, \quad (19)$$

$$\gamma \in \Gamma, x \in X, u \in U, D = \{D_i\}_{i=1}^n.$$

Принцип (19) утверждает, что управляющее воздействие $m = \pi_M(x)$ на управляемый процесс решает поставленную глобальную задачу D тогда, когда x является решением задач $D_i, i = \overline{1, n}$ элементов нижестоящего уровня и желаемые связующие сигналы $u^y = \pi_U(x)$ совпадают (согласованы) с фактическими связующими сигналами $u = K(m), K : U \rightarrow M$, имеющими место тогда, когда к процессу приложено управляющее воздействие $m = \pi_M(x)$. Другая форма записи принципа согласования взаимодействий [3]:

$$(\forall \gamma)(\forall x), \{ [P(x, D(\gamma)) \& q(\gamma, x) = \tilde{q}(\gamma, x)] \Rightarrow P(\pi_M(x), D) \}, \quad (20)$$

$$\gamma \in \Gamma, x \in X, D = \{D_i\}_{i=1}^n,$$

где \tilde{q} и q — заданные функции, отображающие множество $\Gamma \otimes X$ на числовую ось и используемые для оценки точности согласования между фактическими и желаемыми связующими сигналами элементов нижестоящего уровня.

Введем некоторые дополнительные обозначения. Для многоуровневой системы (12)–(14) не будем специально выделять множество управляющих воздействий M , а будем предполагать, что на каждом уровне k системы (12)–(14) для каждого класса эквивалентности z_j^k (j -го элемента k -го уровня) имеется множество F_j^k . Соответственно, для всей модели (12)–(14) имеется множество

$$\Gamma = \left\{ \left\{ \Gamma_j^k \right\}_{j=1}^{N_k} \right\}_{k=1}^K.$$

Элементами множества Γ_j^k являются координирующие воздействия γ_j^k .

Пусть $\Gamma^k = \left\{ \Gamma_j^k \right\}_{j=1}^{N_k}$. При этом на нижнем уровне ($k = K$): $\Gamma^k = \left\{ \Gamma_j^k \right\}_{j=1}^{N_k} \leftrightarrow M$.

В соответствии с приведенной формализацией уровень $k=2$ порождает N_2 двухуровневых систем, уровень $k=i - N_i$ двухуровневых систем и, соответственно, уровень $k=K-1 - N_{K-1}$ двухуровневых систем (рис. 2).

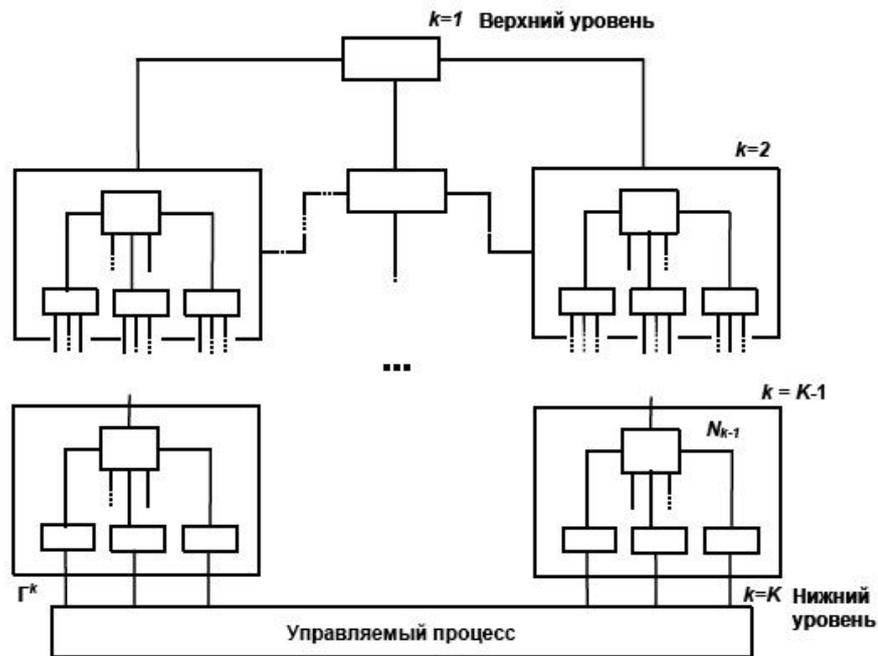


Рис. 2. Структура многоуровневой рекуррентной иерархической модели

Пусть множество управляющих воздействий представлено в виде:

$$\Gamma^k = \Gamma_1^k \times \dots \times \Gamma_{N_k}^k.$$

Предположим, что j -й класс эквивалентности k -го уровня системы (12)–(14) связан с определенным набором классов эквивалентности $(k+1)$ -го уровня, то

есть на $(k+1)$ -ом уровне N_{k+1} классов эквивалентности собраны в N_k групп, связанных отношениями эквивалентности с N_k классами эквивалентности уровня k . Каждая такая группа представляет собой двухуровневую систему S_j^k , порождаемую каждым классом эквивалентности z_j^k :

$$\left(\forall j, j = \overline{1, N_k}\right) \left(\forall k, k = \overline{1, K-1}\right) z_j^k \rightarrow S_j^k. \quad (21)$$

Общее число таких двухуровневых систем:

$$N_S = \sum_{k=1}^{K-1} N_k.$$

Группирование (21) по элементам вышестоящих уровней отражено в [1] введением вектор-индекса. Далее для упрощения записи вектор-индексы опускаются. Предполагается, что все координирующие воздействия, согласующие сигналы, локальные функции качества и т.д., рассматриваемые для определенного уровня двухуровневой системы (12)–(14), группируются по элементам этого уровня в соответствии с заданными отношениями эквивалентности.

Пусть связующие сигналы на k -ом уровне системы (12)–(14) определяются посредством отображения $Q_k: \Gamma^k \rightarrow U^k, k = \overline{1, K}$, где U^k — множество связующих сигналов на уровне k .

Глобальная задача оптимизации D отражает глобальную цель многоуровневой системы и определяется парой (g, Γ^k) , где g — заданная целевая функция. Решением задачи D является такое воздействие $\hat{\gamma}^k$ на нижнем уровне K системы, что:

$$g(\hat{\gamma}^k) = \min_{\Gamma^k} g(\gamma^k), \hat{\gamma}^k \in \Gamma^k.$$

Пусть D_j^k — задача, решаемая j -м элементом k -го уровня системы (12)–(14). Задачи, решаемые на этом уровне, также будут оптимизационными. Локальная оптимизационная задача $D_j^k(\gamma), (\gamma \in \Gamma_j^{k-1}, j = \overline{1, N_{k-1}}, k = \overline{2, K})$ определяется парой $(g_{j\gamma}^k, X_{j\gamma}^k)$, где $g_{j\gamma}^k$ — заданная локальная целевая функция, определенная на множестве решений X_j^k , а $X_{j\gamma}^k$ — заданное подмножество X_j^k , причем $X_{j\gamma}^k = \Gamma_j^k \times U_j^k, j = \overline{1, N_k}, k = \overline{2, K}$.

Решением локальной задачи $D_j^k(\gamma)$ является элемент $x_j^{k\gamma} \in X_{j\gamma}^k$, такой, что:

$$g_{j\gamma}^k(x_j^{k\gamma}) = \min_{X_{j\gamma}^k} g_{j\gamma}^k(x_j^k), x_j^k \in X_{j\gamma}^k, j = \overline{1, N_k}, k = \overline{2, K}.$$

Имеются два способа воздействия на локальные задачи оптимизации [3]: через локальные целевые функции $g_{j\gamma}^k$ или локальные функции качества $G_{j\gamma}^k$ (координация путем изменения целей) и через множество допустимых решений $X_{j\gamma}^k$ (координация путем изменения ограничений). При координации путем изменения целей задаются функции

$$G_{j\gamma}^k: \Gamma_j^k \times U_j^k \times V^k \rightarrow V, j = \overline{1, N_k}, k = \overline{2, K},$$

где B — заданное множество; V — множество платежей. Пусть каждому $\gamma_j^{k-1}, \gamma_j^{k-1} \in \Gamma^{k-1}$ приписано свое (единственное) $\beta_j^{k\gamma} \in B^k$. Тогда из G_{jB}^k получается локальная функция качества

$$G_{j\gamma}^k(\gamma_j^k, u_j^k) = G_{jB}^k(\gamma_j^k, u_j^k, \beta_j^{k\gamma}), \quad j = \overline{1, N_k}, \quad k = \overline{2, K}.$$

При координации путем изменения ограничений каждый координирующий сигнал $\gamma_j^{k-1} \in \Gamma^{k-1}$ ($k-1$ -го уровня определяет для конкретного j -го локального элемента k -го уровня множество допустимых решений $X_{j\gamma}^k$, которое в нашем случае является подмножеством множества $\Gamma_j^k \times U_j^k$. Множества $X_{j\gamma}^k$ представляют собой ограничения, накладываемые на локальные решения. Пусть $X_{j\gamma}^k = \Gamma_j^k \times U_j^{k\gamma}$, где $U_j^{k\gamma}$ — заданное подмножество U_j^k . Тогда координация сводится к выбору соответствующих подмножеств связующих сигналов.

При решении задачи оптимизации на формальной рекуррентной модели нужно выбрать в классах эквивалентности представителей, имеющих определенные характеристики (связующие сигналы), влияющие на суммарные характеристики совокупностей таких представителей. В общем случае прогнозировать предварительно точные значения этих характеристик или диапазонов их изменения — сложная задача, решение которой может быть основано в некоторых частных случаях на априорных знаниях о зависимости суммарных характеристик от характеристик конкретных представителей классов эквивалентности. В связи с этим, наиболее естественным для решения поставленной задачи является именно координирование путем развязывания взаимодействий.

При этом $\forall \gamma_j^k, U_j^{k\gamma} = U_j^k, j = \overline{1, N_k}, k = \overline{2, K}$ и, следовательно, $X_{j\gamma}^k = X_j^k = \Gamma_j^k \times U_j^k, j = \overline{1, N_k}, k = \overline{2, K}$, то есть локальные оптимизационные задачи формулируются для решения независимо друг от друга, и каждый локальный элемент должен выбирать оптимальным образом не только координирующие сигналы для элементов нижележащего уровня, связанных с ним отношениями эквивалентности, но и локальные связующие сигналы.

Многоуровневая система (12)–(14) координируема, если истинно следующее предложение:

$$S_j^k, (\forall j, j = \overline{1, N_k}), (\forall k, k = \overline{2, K}), (\exists \gamma^{k-1})(\exists x^{k\gamma})(\exists \hat{\gamma}^k): [\pi(x^{k\gamma}) = \hat{\gamma}^k], \quad (22)$$

где $\hat{\gamma}^k$ — глобально оптимальное координирующее воздействие на нижнем уровне системы; $x^{k\gamma} = (\gamma^{k\gamma}, u^{k\gamma})$, так что каждая пара $(\gamma^{k\gamma}, u^{k\gamma})$ является оптимальной; $\pi: \Gamma^k \times U^k \rightarrow \Gamma^k$.

В (22) утверждается, что система координируема, если для каждой подсистемы $S_j^k, j = \overline{1, N}, k = \overline{1, K-1}$ существует координирующий сигнал $\gamma^k \in \Gamma^k$ и локальные оптимальные решения $(\gamma_j^{(k+1)\gamma}, u_j^{(k+1)\gamma})$ для элементов нижнего уровня подсистемы S_j^k такие, что координирующий сигнал $\gamma^{k\gamma} = (\gamma_1^{k\gamma}, \dots, \gamma_N^{k\gamma})$ является глобально оптимальным, то есть $\gamma^{k\gamma} = \hat{\gamma}^k$.

Задачей любого элемента вышестоящего уровня в каждой подсистеме S_j^k , $j = \overline{1, N_k}$, $k = \overline{1, K-1}$ является выработка оптимального координирующего сигнала. Каждая из этих задач должна быть сформулирована таким образом, чтобы ее решение являлось искомым оптимальным координирующим воздействием на элементы нижнего уровня подсистемы S_j^k . При формализации этих задач целесообразно использовать постулат совместимости (18) и принципы координации [3], из которых мы выбрали принцип согласования. Принцип согласования взаимодействий (19), (20) для многоуровневой системы (12)–(14) с учетом проведенной формализации выражается следующим предложением:

$$S_j^k, (\forall j, j = \overline{1, N_k}), (\forall k, k = \overline{2, K}), (\forall \gamma_j^{k-1}), (\exists x^{k\gamma}) (\exists \hat{\gamma}^k):$$

$$: \{[(\gamma^k, u^k) = x^{k\gamma} \& L^k(\gamma^k) = u^k] \Rightarrow (\gamma^k = \hat{\gamma}^k)\}$$
(23)

где $L^k = \Gamma^k \rightarrow U^k$.

Предложением (23) утверждается, что глобально оптимальное координирующее воздействие обеспечивается локальными решениями всякий раз, когда для каждой подсистемы S_j^k , $j = \overline{1, N_k}$, $k = \overline{1, K-1}$ связующие сигналы для элементов нижнего уровня согласованы.

При использовании принципа согласования функции качества сравниваются локальные затраты (функции качества), а не связующие сигналы. Пусть

$$(\forall j, j = \overline{1, N_k}), (\forall k, k = \overline{2, K}), (\forall \gamma^{k-1}, \gamma^{k-1} \in \Gamma^{k-1}),$$

$$(\exists g_{j\gamma}^k : \Gamma^k \times U^k \rightarrow V): [g_{j\gamma}^k(\gamma^k, u^k) = (g_{1\gamma}^k(\gamma_1^k, u_1^k), \dots, g_{j_k\gamma}^k(\gamma_{j_k}^k, u_{j_k}^k))]$$
(24)

где j_k — число элементов нижнего уровня в подсистеме S_j^{k-1} .

С учетом (24) принцип согласования функций качества:

$$S_j^k, (\forall j, j = \overline{1, N_k}), (\forall k, k = \overline{2, K}), (\forall \gamma_j^{k-1}), (\exists x^{k\gamma}) (\exists \hat{\gamma}^k):$$

$$: \{[(\gamma^k, u^k) = x^{k\gamma} \& \bar{g}_{j\gamma}^k(\gamma^k, L^k(\gamma^k)) = \bar{g}_{j\gamma}^k(\gamma^k, u^k)] \Rightarrow (\gamma^k = \hat{\gamma}^k)\}$$
(25)

то есть глобально оптимальное координирующее воздействие обеспечивается локальными решениями всякий раз, когда согласованы ожидаемые и фактические локальные затраты для каждой подсистемы S_j^k , $j = \overline{1, N_k}$, $k = \overline{1, K-1}$.

Выражениями (23), (25) определены принципы согласования взаимодействий и функций качества для многоуровневой системы (12)–(14).

При оптимизации элементами многоуровневой системы локальных целевых функций могут возникнуть конфликты (несогласованность) между локальными решениями [21]. Принципы координации обеспечивают отсутствие конфликтов, если при оптимизации локальных целевых функций выполняются условия согласования. Эти условия, в свою очередь, обеспечиваются, если система обладает определенными свойствами. К таким свойствам относятся, например, монотонность, безусловная, ограниченная внутриуровневая или межуровневая согласованность и другие. Указанные свойства определяются через взаимосвязи между целевыми функциями. Для определения этих взаимосвязей вводятся специальные вспомогательные функции [2]:

а) функция глобальных затрат $g : \Gamma \rightarrow V$;

б) локальные функции затрат $h_{j\gamma}^k : \Gamma \rightarrow V$.

Тогда $(\forall j, j = \overline{1, N_k}), (\forall k, k = \overline{2, K}), (\forall \gamma^{k-1} \in \Gamma^{k-1}), (\forall \gamma^k \in \Gamma^k),$
 $[h_{j\gamma}^k(\gamma^k) = g_{j\gamma}^k(\gamma_j^k, L_j^k(\gamma_j^k))],$

то есть $h_{j\gamma}^k(\gamma^k)$ учитывает затраты, которые производит j -й локальный элемент k -го уровня при выработке локального координирующего воздействия γ_j^k и фактически реализуемом связном сигнале $u_j^k = L_j^k(\gamma_j^k)$.

в) межуровневые функции качества $\psi_{j\gamma}^k : V^n \rightarrow V$ в любой подсистеме S_j^k , $j = \overline{1, N_k}, k = \overline{1, K-1}$ для любого $\gamma_j^k \in \Gamma_j^k$ существует единственное отношение

$$\psi_{j\gamma}^k \subseteq V^n \otimes V : \psi_{j\gamma}^k = [(h_{1\gamma}^{k+1}(\gamma^{k+1}), \dots, h_{j_{k+1}}^{k+1}(\gamma^{k+1})), h_j^k(\gamma_j^k)],$$

связывающее суммарные затраты элемента вышестоящего уровня с фактическими локальными затратами элементов нижнего уровня. Если $\psi_{j\gamma}^k$ — функция, то:

$$h_j^k(\gamma_j^k) = \psi_{j\gamma}^k(h_{1\gamma}^{k+1}(\gamma^{k+1}), \dots, h_{j_{k+1}}^{k+1}(\gamma^{k+1})), j = \overline{1, N_k}, k = \overline{1, K-1}, \quad (26)$$

$\psi = \left\{ \left\{ \psi_{j\gamma}^k \right\}_{j=1}^{N_k} \right\}_{k=1}^{K-1}$ — множество межуровневых функций качества.

г) кажущаяся глобальная целевая функция $g_\Gamma : \Gamma \otimes U \rightarrow V$. Функция g_Γ дает суммарные затраты, какими они представляются локальным решающим элементом. g_Γ не всегда дает истинные суммарные затраты, потому что учитывает все пары $(\gamma, u) \in \Gamma \otimes U$, хотя для некоторых из этих пар нарушается условие $u_j^k = L_j^k(\gamma_j^k)$. Однако g_Γ представляет истинные суммарные затраты $g(\gamma)$ всякий раз, когда $u = L(\gamma)$ для $\forall \gamma \in \Gamma$.

Для исследования влияния свойств целевых функций на возникновение конфликтных ситуаций между глобальной и локальными целями вводятся дополнительные ограничения на координирующие управляющие воздействия.

Необходимые и достаточные условия координируемости локальных решений сетцентрического управления в многоуровневых распределенных системах на основе принципов внутриуровневой и межуровневой согласованности доказаны в работе [2].

Таким образом, проведенные исследования подтверждают координируемость систем сетцентрического управления региональной безопасностью, представленных введенной рекуррентной моделью, при удовлетворении специальных (но довольно общих) требований взаимосвязи между показателями (целевыми функциями), оптимизируемыми различными элементами этих многоуровневых систем.

3. Применение моделей координации сетцентрического управления

Устойчивое развитие арктических регионов в значительной степени зависит от безопасности функционирования Северного морского пути (СМП) — транспортной артерии, связывающей западное и восточное полушария Земли. Использование СМП для морских перевозок, по сравнению с альтернативным маршрутом (Суэцким каналом), проходящим через южные регионы, имеет ряд

преимуществ: сокращение времени доставки грузов, снижении издержек на зарплату экипажам судов и их фрахта, экономия топлива, отсутствие пиратских нападений, но сдерживается из-за суровых природно-климатических условий и высоких рисков возникновения чрезвычайных ситуаций природного и техногенного характера. Для ряда арктических регионов СМП является единственным средством жизнеобеспечения и напрямую влияет на их продовольственную безопасность. В связи с этим, обеспечение экологической и транспортной безопасности арктических коммуникаций является актуальной задачей.

Решение данной задачи затрудняется необходимостью оперативной обработки и анализа больших объемов разноплановой информации о влиянии разнородных факторов на состояние СМП, а также координации деятельности субъектов регионального управления и стейкхолдеров в этом районе для принятия стратегических решений. Информационный мониторинг гидрометеорологической, ледовой и навигационной обстановки на акваториях СМП позволяет лишь частично решить эту проблему.

Для решения этой задачи на территории Арктической зоны РФ создается сеть комплексных аварийно-спасательных центров МЧС России с целью обеспечения экологической безопасности СМП при транспортной и хозяйственной деятельности. Некоторые из этих центров уже построены и введены в эксплуатацию, в том числе один из них в Мурманской области.

Таким образом, одним из направлений приложения предложенной в работе модели является информационное сопровождение функционирования региональных аварийно-спасательных центров в арктических регионах в части решения задач координации процессов принятия управленческих решений в сфере обеспечения экологической безопасности СМП.

Так, в качестве тестовой задачи рассматривалась наиболее типичная для СМП ситуация, связанная с аварийными разливами нефтепродуктов на его акватории. Эта ситуация относится к категории быстропротекающих, техногенных чрезвычайных ситуаций и требует оперативного согласованного принятия управленческих решений по формированию планов соответствующих антикризисных мероприятий и их реализации. Для решения этой задачи нашли применение предложенный в работе подход к координации сетецентрического управления безопасностью, основанный на многоуровневой иерархической рекуррентной модели, а также метод синтеза виртуальных организационных структур управления безопасностью, изложенный в работе [22].

На основе этих разработок моделировался процесс координации принятия решений по ликвидации рассматриваемой кризисной ситуации. Двухуровневая система сетецентрического управления экологической безопасностью СМП представлена сетью компьютеров, выполняющих функции управляющих центров. Система развернута на локальной сети со скоростью передачи данных 100 Мбит/с., состоящей из четырех рабочих станций (AMD Ryzen 7 2,2 GHz, 12 Gb RAM; Intel Core i7 1,8 GHz, 8 Gb RAM; Intel Core i5 3,1 GHz, 4 Gb RAM; Pentium Dual Core 2,5 GHz, 4 Gb RAM). На каждом компьютере установлены компоненты распределенной агентной платформы и запущены программные агенты, имитирующие работу профильных ведомств, участвующих в процессах локализации чрезвычайных ситуаций рассматриваемого типа. Агенты используются для определения входных параметров моделирования и формирования оперативного контекста анализируемой ситуации. В качестве

входных параметров рассматривались время суток, метеоданные (температура воздуха, давление, направление и скорость ветра), тип нефтепродуктов, координаты места разлива, масса разлива, тип судна, наличие возгорания и пострадавших.

Модель реализована с использованием агентной платформы JADE [23] и программного имитатора Netlogo [24], образующих в совокупности мультиагентную среду моделирования. Для тестовой задачи сгенерированы 15 программных агентов, функционирующих в рамках двухуровневой системы управления безопасностью СМП: 6 из них являются управляющими агентами верхнего уровня (координаторы), остальные агенты представлены подчиненными подразделениями на нижнем уровне. На основе применения метода синтеза виртуальных организационных структур сформировано четыре альтернативных варианта территориальных комиссий по координации управления ликвидацией рассматриваемой чрезвычайной ситуации с учетом оперативного контекста ситуации. Для расчета оптимального количества сил и средств для ликвидации симулированной на модели ситуации использованы динамические модели распределения ресурсов в условиях чрезвычайных ситуаций, предложенные в работе [25] и обеспечивающие учет времени реагирования, стратегии ликвидации, производительности сил, местоположения и затрат на проведение ликвидации (антикризисных мероприятий).

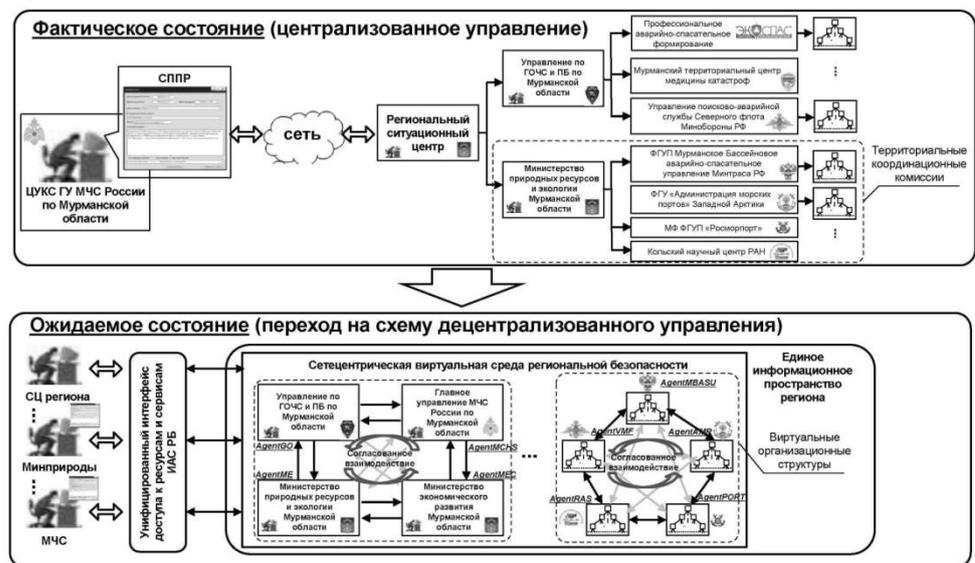


Рис. 3. Переход на модель сетецентрического управления экологической безопасностью СМП

Общее модельное время на принятие решений в моделируемой ситуации, включая время на получение информации системой от диспетчера или ЛППР о типе ситуации, время на генерацию оперативного контекста ситуации, время на поиск (доопределение) неизвестных параметров ситуации, а также время на согласование совместных действий и конфигурацию организационных структур управления безопасностью в рамках переговорного процесса между агентами, составило 17,6259 сек. Обычно принятие решений в подобной ситуации требует

значительных временных затрат, связанных с формированием территориальных координационных комиссий, ведением переговоров, как минимум, между диспетчерами трех профильных ведомств, согласованием сферы ответственности. В лучшем случае время от поступления сигнала о возникновении чрезвычайной или кризисной ситуации на акватории СМП, до принятия решения, в частности, для рассмотренного примера с аварийным разливом нефтепродуктов, занимает до 4 часов.

По экспертным оценкам это время в значительной мере может быть уменьшено в режиме децентрализованного управления безопасностью СМП (рис. 3), который является естественным для функционирования данного объекта управления, поскольку единой системы организационного управления безопасностью СМП в настоящее время не существует (организационно разнородные ведомства не подчиняются друг другу, преследуют собственные цели и, как следствие, плохо координируются). Предварительные исследования показали, что при таком подходе обеспечивается сокращение времени на принятие решений по предотвращению и ликвидации чрезвычайных ситуаций, характерных для СМП, примерно на 30-40%. Эффект достигается за счет применения созданной рекуррентной модели координации локальных решений сетцентрического управления безопасностью СМП и агентных технологий в информационно-аналитических системах региональных ситуационных центров.

Заключение

В работе представлены результаты математического моделирования систем сетцентрического управления региональной безопасностью как многоуровневых распределенных систем. Системы этого класса характеризуются слабой формализованностью, многоаспектностью происходящих в них процессов и их взаимосвязанностью, динамичностью, организационной разнородностью и сложностью координации взаимодействия управляющих элементов и подсистем.

Для решения задач координации децентрализованного принятия решений в сетцентрической системе управления региональной безопасностью создана многоуровневая рекуррентная иерархическая модель управления безопасностью региона. Модель построена на основе функционально-целевой технологии и математического аппарата теории иерархических многоуровневых систем. Модель является аналогом оператора структурно-функционального синтеза исполнительской среды информационно-аналитической поддержки управления.

Рекуррентность заключается в применении модели к самой себе при декомпозиции целей управления и детализации соответствующих действий для их достижения (результат предыдущего выполнения процедуры декомпозиции является исходным для последующего). Рекуррентный процесс детализации исходной функции продолжается до достижения уровня «примитивов» — элементарных функций, неделимых с точки зрения субъекта управления.

Координация реализуется путем контроля ограничений на управление. При таком способе координации должны выполняться принципы согласования взаимодействий и функций качества, а также постулат совместимости задач, решаемых элементами системы, по отношению к общей задаче системы.

Отличительной особенностью предложенной модели является то, что в ней совмещаются координация путем развязывания взаимодействий и создание коалиций между агентами на разных уровнях управления. Самоорганизация

агентов в коалиции и согласование локальных решений сетецентрического управления обеспечивают сокращение времени на выработку и реализацию решений системой организационного управления региональной безопасностью.

В рамках дальнейших исследований планируется расширить спектр действующих факторов внутренней и внешней среды, определить структуру взаимосвязей между ними и их формализация в предложенных моделях координации, а также учет влияния человеческого фактора в процессе принятия решений на разных управления.

Литература

1. Маслобоев, А. В. Информационное измерение региональной безопасности в Арктике / А. В. Маслобоев, В. А. Путилов.- Апатиты: КНЦ РАН, 2016. 222 с.
2. Кузьмин И. А. Распределенная обработка информации в научных исследованиях / И. А. Кузьмин, В. А. Путилов, В. В. Фильчаков. Л.: Наука, 1991. 304 с.
3. Месарович М. Теория иерархических многоуровневых систем / М. Месарович, Д. Мако, И. Такахара. М.: Мир, 1973. 343 с.
4. Цыгичко В. Н. Руководителю о принятии решений. Изд. 3-е перераб и доп. / В. Н. Цыгичко. М.: Красанд, 2010. 352 с.
5. Larichev O. I. Decision Support Systems for Illstructured Problems: Requirements and Constraints / O. I. Larichev, A. B. Petrovsky // Organizational Decision Support Systems. Amsterdam, North-Holland, 1988. Pp. 247–257.
6. Шориков А. Ф. Методология моделирования многоуровневых систем: иерархия и динамика / А.Ф. Шориков // Прикладная информатика. 2006. № 1. С. 136–141.
7. Маслобоев А. В. Концептуальная модель интегрированной информационной среды поддержки управления безопасностью развития региона / А. В. Маслобоев, В. А. Путилов // Вестник МГТУ. 2011. Т. 14. № 4. С. 842–853.
8. Палюх Б. В. Приложения метода разделения состояний к управлению технологической безопасностью на основе индекса безопасности / Б. В. Палюх, В. Н. Богатиков, В. В. Алексеев, А. Е. Пророков. Тверь: Изд-во ТвГТУ, 2009. 398 с.
9. Информационная безопасность систем организационного управления. Теоретические основы. В 2 т. / Под ред. Н. А. Кузнецова, В. В. Кульбы. М.: Наука, 2006. Т. 1. 496 с.
10. Бурков В. Н. Введение в теорию управления организационными системами / В. Н. Бурков, Н. А. Коргин, Д. А. Новиков. М.: Либроком, 2009. 264 с.
11. Sarbazi-Azad H. Large Scale Network-Centric Distributed Systems / H. Sarbazi-Azad, A. Y. Zomaya. John Wiley & Sons, 2013. 760 p.
12. Фридман А. Я. Координация иерархических организационных систем: игровой и градиентный подходы / А. Я. Фридман, О. В. Фридман, В. А. Зеленцов // Проблемы теории и практики управления. 2011. № 6. С. 14–22.
13. Городецкий В. И. Самоорганизация и многоагентные системы. Часть 1. Модели многоагентной самоорганизации / В. И. Городецкий // Известия РАН. Теория и системы управления. 2012. № 2. С. 92–120.

14. Запорожцев В. В. Нечеткая параметрическая координация в многоуровневой иерархической системе / В. В. Запорожцев, В. И. Новосельцев, А. Ю. Струков // Системы управления и информационные технологии. 2012. Т. 50. № 4. С. 142–145.
15. Новиков Д. А. Рефлексивные игры / Д. А. Новиков, А. Г. Чхартишвили. М.: СИНТЕГ, 2003. 160 с.
16. Юдицкий С. А. Триадный подход к моделированию систем сетцентрического управления / С. А. Юдицкий, П. Н. Владиславлев, Д. С. Точ // Управление большими системами. 2010. № 28. С. 24–39.
17. Stoilov T. Goal and Predictive Coordination in Two Level Hierarchical Systems / T. Stoilov, K. Stoilova // International Journal of General Systems. 2008. Vol. 37. No. 2. Pp. 181–213.
18. Городецкий В. И. Прикладные многоагентные системы группового управления / В. И. Городецкий и др. // Искусственный интеллект и принятие решений. 2009. № 2. С. 3–24.
19. Дыбова М. А. Анализ направлений развития методов сетевого планирования и управления организационными системами / М. А. Дыбова, В. М. Колыхалин, С.А. Мишин // Вестник Воронежского института ФСИИ России. 2013. № 1. С. 86–90.
20. Findeisen W. Control and Coordination in Hierarchical Systems. John Wiley & Sons, 1980. 467 p.
21. Макаров А. А. Методы и модели согласования иерархических решений. М.: Наука, 1979. 237 с.
22. Маслобоев А. В. Метод автоматизированного синтеза виртуальных организационных структур для задач управления региональной безопасностью // Программные продукты и системы. 2013. № 4(104). С. 141–149.
23. Bellifemine F. Developing Multi-Agent System with JADE / F. Bellifemine, G. Caire, D. Greenwood. John Wiley & Sons, 2007. 300 p.
24. Wilensky U. An introduction to agent-based modeling: Modeling natural, social and engineered complex systems with NetLogo / U. Wilensky, W. Rand. Cambridge, MA: MIT Press. 2015. 504 p.
25. Ямалов И. У. Моделирование процессов управления и принятия решений в условиях чрезвычайных ситуаций. М.: БИНОМ. Лаборатория знаний, 2007. 288 с.

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

В. В. Быстров, Д. Н. Халиуллина

Институт информатики и математического моделирования ФИЦ КНЦ РАН

РАЗРАБОТКА СИСТЕМЫ ИНФОРМАЦИОННОГО МОНИТОРИНГА И ИНДИКАТОРНОЙ ОЦЕНКИ УГРОЗ ЭКОЛОГИЧЕСКОЙ БЕЗОПАСНОСТИ АРКТИЧЕСКИХ КОММУНИКАЦИЙ*

Аннотация

Статья посвящена созданию компонентов информационных систем поддержки принятия решений в сфере управления экологической безопасностью макрорегионов Российской Федерации. Авторы рассматривают вопросы, связанные с разработкой системы мониторинга и оценки угроз экологической безопасности арктических коммуникаций. Описывается архитектура разрабатываемой системы, построенная на принципах сетецентрического управления и с применением мультиагентных технологий. В статье уделяется внимание формированию системы индикаторной оценки состояния экологической безопасности, а также детально рассматривается структура и функционал компонентов разрабатываемых программных агентов. Для оценки возможных сценариев развития экологической ситуации предлагается добавить в интеллектуальные агенты прогностические функции, реализуемые посредством имитационного моделирования.

Ключевые слова:

мониторинг, индикаторная система, сетецентрическое управление, мультиагентные системы, экологическая безопасность, арктические коммуникации.

V. V. Bystrov, D. N. Khaliullina

DEVELOPMENT OF INFORMATION MONITORING SYSTEM AND INDICATOR ASSESSMENT OF THREATS TO ENVIRONMENTAL SAFETY OF ARCTIC COMMUNICATIONS

Abstract

The article is devoted to the creation of components of information systems to support decision-making in the field of environmental safety management of macroregions of the Russian Federation. The authors consider the issues related to the development of a system for monitoring and assessing threats to the environmental safety of Arctic communications. The architecture of the developed system is described. The architecture is built on the principles of network-centric management and with using multi-agent technologies. The authors pay attention to the formation of the system of indicator assessment of the state of environmental safety, and also consider in detail the structure and functionality of the components of the developed software agents. To assess possible scenarios of the environmental situation development, it is proposed to add predictive functions to the intelligent agents, implemented through simulation.

Keywords:

monitoring, indicator system, network-centric management, multi-agent systems, environmental safety, Arctic communications.

* Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ и Российского фонда фундаментальных исследований — проект № 18-07-00167 А.

Введение

В последнее время все чаще поднимаются вопросы развития Северного морского пути [1], использование которого позволяет не только сократить время доставки грузов в отдаленные районы нашей страны и другие государства, но и открывает новые горизонты в освоении Арктики. Исторически к исследованию арктических территорий относят такие направления как: изучение континентального шельфа, поиск новых источников неразработанных энергоресурсов (нефти и газа), расширение знаний о биоразнообразии арктической флоры и фауны и другие.

Хозяйственное освоение арктических территорий, как и любая другая экономическая деятельность человека, ориентирована на получение практической выгоды, но и наносит определенный вред окружающей среде. Загрязнение прибрежных вод, воздуха и почв арктических районов, таяние льдов и другие негативные последствия хозяйственной деятельности человека нарушают хрупкую экосистему Арктики и приводят к снижению биологического разнообразия. Как следствие, страдает коренное население данных территорий, поскольку их натуральное хозяйство и быт напрямую зависят от доступности естественных биологических ресурсов.

Ключевым моментом при покорении Арктики должны стать меры по уменьшению экологического ущерба. Для этого необходимо разработать систему мониторинга и индикаторной оценки состояния экологической безопасности арктических территорий, в частности, арктических коммуникаций в связи с бурным развитием Северного морского пути.

Система мониторинга и индикаторной оценки состояния окружающей среды

В настоящем исследовании под мониторингом экологического состояния арктических коммуникаций будем понимать систему наблюдений, оценки и прогноза изменения окружающей среды под влиянием антропогенных воздействий в акватории Северного морского пути. В понятие арктических коммуникаций входят морские транспортные сети и инфраструктурные объекты, которые обслуживают данные сети.

На сегодняшний день существуют готовые решения (сервисы) в области мониторинга определенных наборов показателей из разных предметных областей, основной задачей которых является сбор информации от первичных источников-сенсоров и последующее отображение наблюдаемой информации в цифровом виде (например, на электронной карте [2]). Обычно в таких сервисах в роли первичных источников информации выступают автоматические идентификационные системы, спутниковые системы связи, системы дистанционного зондирования Земли, радиолокационные станции, доклады с борта судов и прочие. При этом все системы мониторинга можно условно разделить на две большие группы. К первой группе относятся системы, которые собирают данные о движении различных судов, загруженности портов и другой транспортно-логистической информации (MarineTraffic [3], ShipFinder[4], «Виктория» [5], FleetMon [6], VesselFinder [7], Vesseltracker [8], Myshiptracking [9] и др.). Вторая группа представлена программно-аппаратными комплексами, осуществляющими мониторинг состояния окружающей среды, т.е. собирают данные о толщине и концентрации льда, солёности, движении различных слоев

океана (в частности, течений), различных атмосферных явлениях и т.п. (Mercator [10], Arctic-Roos [11], Cryo [12], EUMETSAT OSI SAF[13], Polar View [14], NCEI [15], SIOS [16] и др.). Все рассмотренные выше системы достаточно эффективны, и обладают необходимым функционалом для реализации собственных задач, но ни одна из них не позволяет осуществлять комплексный мониторинг разнородных показателей арктических коммуникаций. Одним из комплексных решений в данной области является отечественный программно-аппаратный комплекс «СПИИРАН-НТБВТ», направленный на увеличение доли автоматизации в подготовке данных наблюдений по Арктическому региону [17, 18]. Для успешного функционирования данного решения необходим конкретный заказчик, в соответствии с потребностями которого разворачивается и настраивается система мониторинга.

Таким образом, проблема создания программных систем мониторинга, реализующих комплексный сбор и анализ данных из разнородных источников, является актуальной. В рамках реализации научно-исследовательского проекта «Методы и средства сетецентрического управления экологической безопасностью Северного морского пути» авторы статьи разрабатывают систему информационного мониторинга и индикаторной оценки угроз экологической безопасности арктических коммуникаций, которая бы сочетала в себе функции перечисленных систем, а также дополняла бы их возможностью комплексной оценки текущего и прогнозируемого состояний окружающей среды. Существенным отличием разрабатываемой системы является закладываемая в ней «интеллектуальность», которая подразумевает проведение мониторинга и анализа собранных данных в своеобразной виртуальной мультиагентной среде. Предполагается, что интеллектуальность разрабатываемой системы обеспечит добавление в ее структуру специальных «интеллектуальные агентов». При этом роль такого типа агентов заключается не только в сборе информации от объектов реального мира (сенсоров, датчиков), но и в последующем ее анализе, а также прогнозировании возможных сценариев развития экологической обстановки при хозяйственном использовании арктических коммуникаций. В качестве эффективного и зарекомендовавшего себя на практике средства прогнозирования планируется использовать имитационное моделирование.

При разработке системы информационного мониторинга и индикаторной оценки активно используются принципы сетецентрического управления: оперирования большими объемами данных, сетевая структура системы, динамический центр принятия решений, разнородность узлов сети и другие. В частности, представление виртуальной среды мониторинга и оценки в виде сетецентрической структуры позволит, по мнению авторов, менять в системе центр принятия решения и координации в зависимости от конкретной стадии и вида решаемой задачи.

Система индикаторной оценки экологического состояния

Подходя к вопросу создания системы мониторинга и оценки угроз экологической безопасности Северного морского пути, для начала необходимо определить структуру показателей, описывающих экологическое и связанной с ней экономическое состояние арктических коммуникаций. В рамках исследования задача формирования системы наблюдаемых показателей решалась

посредством выборки показателей из существующих систем мониторинга за состоянием окружающей среды и оценки экономических параметров природопользования и добавления к ним специфических параметров, характеризующих хозяйственное использование транспортно-логистических систем Арктики.

Все показатели, которые используются в разрабатываемой системе мониторинга экологической обстановки можно условно разделить на несколько групп:

1. Метеорологические. Содержат информацию о толщине льда, колебаниях волн, скорости и направлении ветра, атмосферном давлении, солнечной активности, радиации, снежном покрове (высота, содержание влаги), температурные показатели и др.

2. Навигационные. К данной группе относятся такие показатели, как геолокационные данные о местоположении судов, находящихся в акватории Северного морского пути, курс движения, параметры технического состояния элементов транспортно-логистических систем и др.

3. Экологические. Данная группа показателей включает в себя биоразнообразие, различные виды загрязнений — почвы (твердыми отходами, стоками и т.п.), атмосферного воздуха, водных объектов и др.

4. Экономические: объемы инвестиций в инфраструктуру, загруженность портов и судов, параметры товаропотоков, стоимость ГСМ, мировые цены на нефть, затраты на амортизацию, стоимость грузов и др.

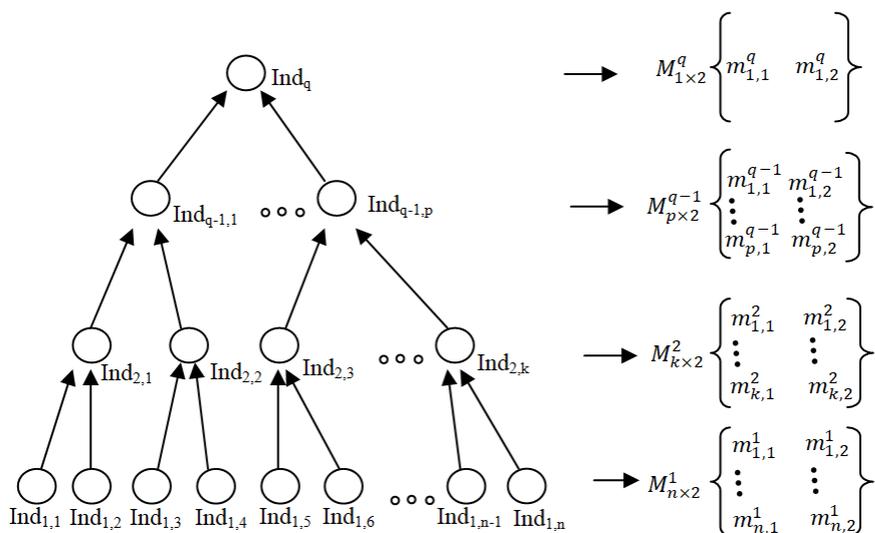


Рис. 1. Иерархическая структура системы индикаторов с матричным представлением ограничений

Для реализации оценки состояния экологической безопасности арктических коммуникаций предложена система индикаторов. Сформированная индикаторная система представляет собой иерархическую структуру, нижним уровнем которой является фиксированный набор определенных показателей, данные о которых собираются разрабатываемой системой информационного

мониторинга из различных источников. Для каждого уровня иерархии определяется матрица ограничений, в которой указываются граничные (предельно допустимые) значения для каждого из индикаторов. Введение в рассмотрение матриц ограничений позволяет выявлять критическую ситуацию в оцениваемом состоянии экологической безопасности, признаком которой является выход текущего значения хотя бы одного индикатора за пределы интервала его допустимых значений. Алгоритм индикаторной оценки предполагает выявление критической ситуации на каждом уровне иерархии построенного дерева индикаторов посредством последовательного сравнения текущего значения индикатора с соответствующими элементами матрицы ограничений. Процесс построения дерева (системы) индикаторов был выполнен по принципу «снизу-вверх»: следующий уровень иерархии формировался на основе предыдущего за счет агрегирования (например, методом вычисления взвешенной суммы параметров) показателей (рис. 1).

В результате верхний уровень иерархии построенного дерева индикаторов представлен одной вершиной, которая является, по своей сути, интегральным показателем экологической безопасности арктических коммуникаций.

Стоит отдельно отметить, что в рамках текущего исследования выявленные критические ситуации рассматриваются как признаки потенциальных угроз экологической безопасности. Подтверждение факта наличия угрозы, т.е. переход ее из категории потенциальной в фактическую, осуществляется посредством прогнозирования развития ситуации в заданных сценарных условиях с последующей оценкой результатов мониторинга и моделирования с привлечением «экспертов-предметников».

Архитектура системы мониторинга и индикаторной оценки

Предлагаемая архитектура системы мониторинга экологической безопасности состоит из двух уровней: виртуального и реального. В качестве объектов реального уровня выступают различные датчики, снимающие показания океанических вод (температуру, соленость, скорость течения и др.), метеорологические станции и воздушные зонды, отвечающие за атмосферные осадки и качество воздуха, гидролокационные и радиолокационные системы, следящие за биоразнообразием флоры и фауны арктической зоны, орбитальные спутники и др. (рис. 2).

Виртуальный уровень архитектуры представляет собой множество агентов виртуальной среды. На текущей стадии исследования для реализации всех проектируемых функций системы мониторинга и оценки предусмотрено создание трех типов агентов виртуальной среды:

«Агенты-датчики». Являются своеобразными сборщиками информации. Считывают оперативную информацию с датчиков/сенсоров и записывают ее в локальную базу данных (ЛБД) с последующей передачей в общее распределенное хранилище (ОРХ) системы, а также реализуют возможность формировать ответы на прямые запросы других объектов виртуальной среды.

«Агенты-агрегаторы». Отвечают за комплексный сбор разнородных данных от существующих сервисов мониторинга за состоянием окружающей среды и функционированием морских транспортно-логистических систем, рассмотренных в разделе «Система мониторинга и индикаторной оценки».

состояния окружающей среды». Как и агенты первого типа, они могут временно хранить получаемые данные в своей локальной базе данных, а также передавать их в общее распределенное хранилище информации. Поскольку информация, которой оперируют агенты-агрегаторы, формируется из разнородных веб-сервисов, предоставляющих ее в разных форматах, то основной функцией данного типа программных агентов является согласование собираемых данных между собой и представление их в едином внутрисистемном формате. При выявлении конфликтных ситуаций в предоставлении противоречивых данных из разных источников эти агенты агрегируют получаемую информацию с учетом заданных пользователем системы критериев (например, по минимальному, максимальному или среднему значению). Данный тип агента имеет модульную структуру, которая позволяет наиболее эффективно реализовать возможность оперативного изменения перечня веб-сервисов, с которыми может взаимодействовать агент. Каждый программный модуль описывает интерфейс взаимодействия агента с конкретным веб-сервисом, используя вызов функций собственного API данного сервиса. В случае получения сложного ответа от веб-сервиса на сформированный агентом запрос предлагается его парсерить, используя зарекомендовавший себя на практике формат JSON.

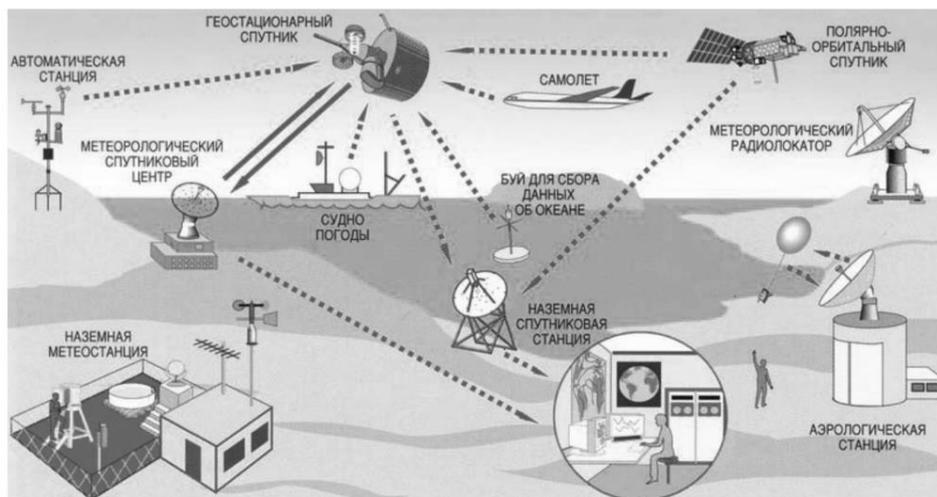


Рис. 2. Виды объектов реального уровня архитектуры [19]

«Интеллектуальные агенты». Основными функциями данного типа агентов являются: проведение оценки текущего состояния экологической обстановки в заданной геолокации; прогнозирование будущего состояния экологической обстановки в заданной геолокации; проведение индикаторной оценки состояния экологической обстановки; координация разнотипных агентов внутри виртуальной среды.

Оценка текущего состояния экологической обстановки проводится в соответствии с формируемым деревом индикаторов экологической безопасности арктических коммуникаций, описанных в «Системе индикаторной оценки экологического состояния». «Интеллектуальные агенты» анализируют значения нижнего уровня иерархии в соответствии с заданной матрицей ограничений для выявления возможных критических ситуаций. При наличии

критической ситуации (выход значений индикатора/-ов за установленные границы) данные агенты информируют об этом лицо, принимающее решение, а также других агентов сети с занесением информации в общее распределенное хранилище.

Прогнозирование будущего состояния экологической обстановки и ее индикаторная оценка производится на основе имитационного моделирования, которое позволяет многократно воспроизводить моделируемые процессы, учитывая различные внешние воздействия на изучаемую систему. Для формирования прогнозов возможных сценариев развития ситуации в экологической безопасности используются три подхода имитационного моделирования: системная динамика, дискретно-событийное и агентное моделирование. Как и в случае с текущим состоянием экологической обстановки, индикаторная оценка прогнозируемого состояния производится на основе дерева индикаторов и соответствующих матриц ограничений.

В соответствии с принципами сетевидного управления предлагается «интеллектуальные агенты» наделить координирующей функцией. В зависимости от текущей стадии решаемой задачи различные «интеллектуальные агенты» могут брать на себя роль координатора при организации распределенного (коллективного) принятия решения в процессе реагирования на возникновение критической ситуации с целью сохранения экологического равновесия.

Архитектура «интеллектуального» агента

Из всех типов агентов разрабатываемой системы информационного мониторинга и индикаторной оценки экологического состояния «интеллектуальные» агенты обладают наибольшей функциональностью. Для реализации основных функций «интеллектуального» агента предлагается блочная архитектура, представленная на рисунке 3. На рисунке стрелками отображаются потоки информации между функциональными блоками архитектуры агента, а также с внешними компонентами.

Взаимодействие агента с пользователем происходит через соответствующий блок графического интерфейса пользователя (GUI). Внешний вид и функциональное наполнение графического интерфейса зависит от категории пользователя, взаимодействующего с агентом. На данном этапе работы выделяется три категории пользователей: администратор системы (производит настройку системы и следит за работоспособностью системы), эксперт (участвует в процессах индикаторной оценки и прогнозирования) и лицо, принимающее решение (задает параметры решаемой задачи, принимает участие в оценке адекватности результатов прогнозирования).

В блоке формирования информации производится выбор формы представления получаемых результатов мониторинга и индикаторной оценки (графики, диаграммы, таблицы и др.) и непосредственное наполнение их конкретными данными. Если пользователю необходимо помимо просмотра результатов мониторинга экологической обстановки выполнить анализ данных по фиксированному набору критериев или получить прогноз на период, то это указывается в блоке постановки задачи.

Сбор и анализ показателей мониторинга производится в одноименном блоке. Причем данные могут быть получены как с внешнего распределенного

хранилища информации, так и из локальной базы агента, если они уже были записаны туда ранее. За организацию взаимодействия с локальным и внешним хранилищами данных отвечает соответствующий блок архитектуры «интеллектуального» агента. Для согласования содержимого локальной базы данных и внешнего распределенного хранилища периодически блоком согласования БД осуществляется их синхронизация, направленная на выявление разнообразных коллизий в хранимых данных. Политика разрешения выявленных коллизий данных определяется на этапе развертывания системы информационного мониторинга и индикаторной оценки, либо в результате выработки коллективного решения коалицией заинтересованных интеллектуальных агентов.

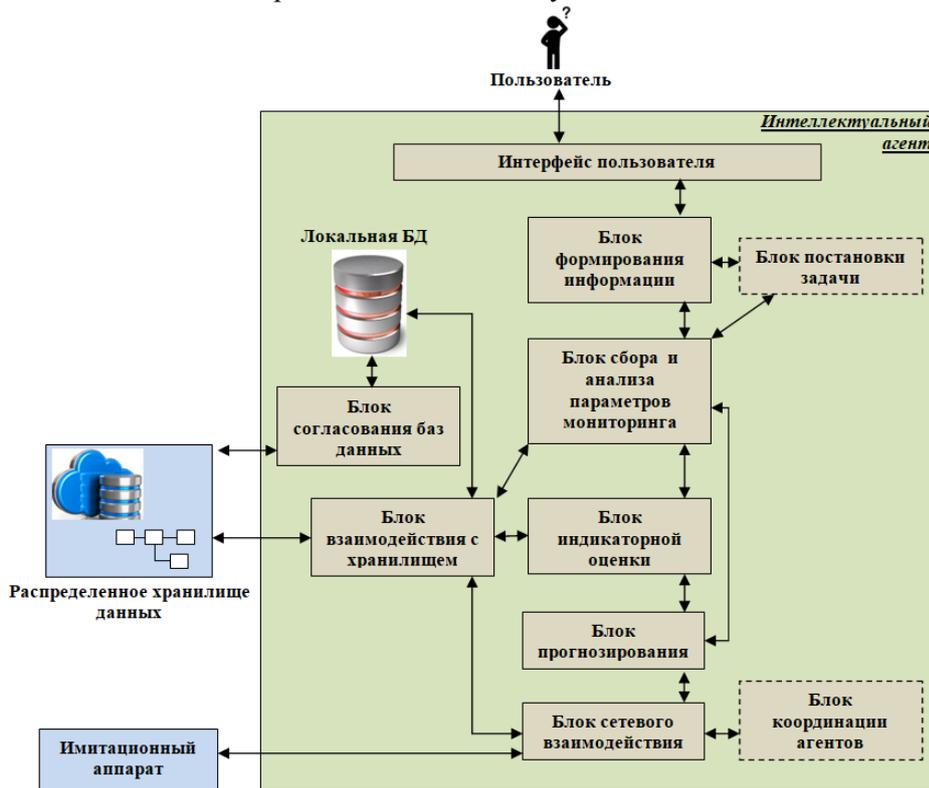


Рис. 3. Архитектура «интеллектуального агента»

Блок взаимодействия с хранилищем архитектурно связан с функционированием блоков индикаторной оценки и сетевого взаимодействия. В первом случае, алгоритм индикаторной оценки основывается на оперировании значениями показателей мониторинга и матрицами ограничений, содержащимися в соответствующих структурах данных хранилища. Информационные потоки между блоками индикаторной оценки и прогнозирования передаются систематически, а их направление определяется в зависимости от конкретной стадии решения задачи. В свою очередь, блок прогнозирования производит выбор и настройку определенного набора имитационных моделей (в зависимости от класса решаемой задачи), осуществляет обработку результатов вычислительных экспериментов. Планируется добавить в разрабатываемую

систему мониторинга и оценки средства организации распределенного вычислительного эксперимента, подразумевающего распараллеливание процесса имитации за счет выполнения вычислительных операций отдельных компонентов модели на разных узлах виртуальной среды.

Блок сетевого взаимодействия введен в архитектуру интеллектуального агента для описания совокупности протоколов и интерфейсов для организации взаимодействия между элементами виртуальной среды на более низком техническом уровне. Он выступает своеобразной платформой для других блоков интеллектуального агента, для функционирования которых требуется взаимодействовать с другими агентами.

Заключение

Задача обеспечения экологической безопасности хозяйственного использования Северного морского пути является актуальной и связана с разработкой средств мониторинга и оценки экологического состояния арктических коммуникаций. Такие программные инструменты должны выполнять следующие функции: в режиме реального времени получать информацию от различных источников; анализировать и агрегировать ее; строить прогнозы на различную временную перспективу. На сегодняшний день авторам статьи не удалось выявить программно-аппаратного решения, полностью реализующего указанные функции.

Авторы предлагают разработать программную систему, реализующую в себе мультиагентную технологию информационного мониторинга и индикаторной оценки угроз экологической безопасности арктических коммуникаций. Разрабатываемая мультиагентная система обладает сетевой структурой и позволяет формировать коалиции агентов в соответствии с их уровнем компетенции. При этом в системе реализуется возможность динамически переносить центр координации и принятия решения от одного агента к другому в соответствии с решаемой задачей.

По мнению авторов, использование имитационного моделирования в системе мониторинга и оценки позволит не только исследовать факторы, влияющие на экологическую обстановку Северного морского пути, но и строить прогнозы возможных сценариев развития арктических коммуникаций с целью выявления потенциальных угроз для их экологической безопасности. Применение системно-динамического подхода даст возможность выявить общие тенденции в изменении показателей экологической обстановки, а также имитировать развитие кризисных ситуаций. Использование агентных моделей позволит учесть влияние каждого объекта на всю систему в целом.

Литература

1. Маслобоев А. В. Информационное измерение региональной безопасности в Арктике / А. В. Маслобоев, В. А. Путилов. Апатиты: КНЦ РАН, 2016. 222 с.
2. Лаверов Н. П., Попович В. В., Ведешин Л. А., Коновалов В. Е. Создание системы мониторинга окружающей обстановки в Арктике. Современные подходы и возможности системы освещения обстановки в Арктике в интересах обеспечения безопасного судоходства Северного морского пути. Современные проблемы дистанционного зондирования Земли из космоса. 2017. Т. 14, № 3. С. 141–157.

3. MarineTraffic: Global Ship Tracking Intelligence | AIS Marine Traffic. Режим доступа: <https://www.marinetraffic.com>.
4. Ship Finder – The Live Marine Traffic Tracking App. Режим доступа: <https://shipfinder.co>.
5. Виктория. Режим доступа: https://victoria-online.ru/index_rus/index.html.
6. Live AIS Vessel Tracker with Ship and Port Database. Режим доступа: <https://www.fleetmon.com>.
7. Free AIS Ship Tracking of Marine Traffic – VesselFinder. Режим доступа: <https://www.vesselfinder.com>.
8. Terrestrial & Satellite AIS Tracking Service in Realtime – Vesseltracker. Режим доступа: <https://www.vesseltracker.com>.
9. My Ship Tracking Free Realtime AIS Vessel Tracking Vessels Finder Map. Режим доступа: <https://www.myshiptracking.com>.
10. Ocean Forecasts – Mercator Ocean. Режим доступа: <http://bulletin.mercator-ocean.fr>.
11. Arctic Regional Ocean Observing System. Режим доступа: <https://arctic-roos.org>.
12. Welcome to cryo.met.no | cryo.met.no. Режим доступа: <https://cryo.met.no/>.
13. OSI SAF | Observing oceans from space. Режим доступа: <http://www.osi-saf.org/>.
14. Polarview.org. Режим доступа: <http://polarview.org/>.
15. National Centers for Environmental Information (NCEI). Режим доступа: <https://www.ncei.noaa.gov/>.
16. The Observing System | sios.metsis.met.no. Режим доступа: <https://www.sios-svalbard.org/ObservingSystem>.
17. Научно-исследовательские разработки. Режим доступа: <http://ntb.oogis.ru/index.php/index.php?id=19>.
18. Лаверов Н. П., Попович В. В., Ведешин Л. А., Гальяно Ф. Р. Методы анализа данных дистанционного зондирования Земли // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 6. С. 145–153.
19. Тихонова И. О., Кручинина Н. Е. Основы экологического мониторинга. – ИНФРА-М Москва, – 2015. – 240 с.

Сведения об авторах

Быстров Виталий Викторович

к.т.н., старший научный сотрудник

e-mail: bystrov@iimm.ru

Vitaliy V. Bystrov

Candidate of Tech. Sciences, senior researcher

Халиуллина Дарья Николаевна

к.т.н., научный сотрудник

e-mail: khaliullina@iimm.ru

Darya N. Khaliullina

Candidate of Tech. Sciences, researcher

А. В. Вицентий¹, М. Г. Шишаев^{1,2}, Г. Г. Гогоберидзе²

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² Федеральное государственное бюджетное образовательное учреждение высшего образования «Мурманский арктический государственный университет»

АРХИТЕКТУРА МУЛЬТИПРЕДМЕТНОЙ СИСТЕМЫ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ РЕШЕНИЯ ЗАДАЧ РЕГИОНАЛЬНОГО УПРАВЛЕНИЯ С УЧЕТОМ ПРОТИВОРЕЧИЙ ПРОСТРАНСТВЕННОГО ПЛАНИРОВАНИЯ ХОЗЯЙСТВЕННОЙ ДЕЯТЕЛЬНОСТИ*

Аннотация

В данной статье рассматривается разработка обобщенной архитектуры мультимедийной системы информационной поддержки решения задач регионального управления с использованием концептуальной и семантической моделей предметной области и технологии когнитивной геовизуализации с целью выявления потенциальных противоречий связанных с деятельностью основных стейкхолдеров региона. В статье приводится постановка задачи и основные требования к разрабатываемой архитектуре. Описывается общая концепция мультимедийной системы. Подробно представлены модель, структура, основные компоненты и их взаимосвязи, а также функции, реализуемые в рамках предлагаемой архитектуры системы. В заключительной части статьи кратко изложены выводы, полученные в ходе работы и потенциальные направления дальнейших исследований.

Ключевые слова:

архитектура информационных систем, поддержка регионального управления, пространственное планирование, когнитивный интерфейс, концептуальная модель предметной области, семантическая модель предметной области, динамическая визуализация пространственных данных, Арктическая зона Российской Федерации, Мурманская область.

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

ARCHITECTURE OF A MULTI-SUBJECT SYSTEM OF INFORMATION SUPPORT OF THE SOLUTION OF REGIONAL MANAGEMENT PROBLEMS TAKING INTO ACCOUNT CONTRADICTIONS OF SPATIAL PLANNING OF ECONOMIC ACTIVITY

Abstract

This paper discusses the development of a generalized architecture of a multi-subject system for information support for solving regional management problems using conceptual and semantic models of the subject area and cognitive geovisualization technology in order to identify potential contradictions associated with the activities of the main stakeholders of the region. The paper provides the statement of the problem and the basic requirements for the architecture being developed. The general concept of a multi-subject system is described. The model, structure, main components and their interconnections, as well as functions implemented within the framework of the proposed system architecture, are presented in detail. The concluding part of the paper summarizes the conclusions obtained during the work and potential directions for further research.

Keywords:

architecture of information systems, support for regional management, spatial planning, cognitive interface, conceptual model of a subject area, semantic model of a subject area, dynamic visualization of spatial data, Arctic zone of the Russian Federation, Murmansk region.

* 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.

1. Введение

Для решения задач регионального управления и пространственного планирования, включая морское (акваториальное) пространственное планирование (МПП) разработаны различные подходы: системный, балансовый, нормативный, программно-целевой и другие. Однако в области морского пространственного планирования и рациональной организации хозяйственной и иной деятельности основных стейкхолдеров региона наибольшее распространение получили системный и экологический подходы [1], позволяющие осуществлять распределение, анализ и согласование различных видов деятельности во времени и пространстве с учетом состояния окружающих экосистем.

Применение методов пространственного планирования в масштабах региона позволяет более точно оценить комплексное влияние различных видов деятельности (хозяйственной, природоохранной, военной и т.д.) как на природно-экологические системы, так и другие виды деятельности. Кроме того, эти методы позволяют обнаруживать и предупреждать потенциальные конфликты между основными стейкхолдерами [2], осуществляющими деятельность в границах региона. В настоящее время, программно-алгоритмическая реализация методов и технологий регионального управления и пространственного планирования, включая МПП, осуществляется в рамках мультипредметных информационных систем, характерной особенностью которых является разнородность используемых данных, решаемых задач и пользователей системы. [3] Для повышения эффективности работы пользователей с мультипредметной информационной системой (МИС) необходимо осуществлять подготовку выводимой информации в соответствии с информационными потребностями конкретного пользователя или типов пользователей [4, 5]. Для этой цели в работе предлагается использовать адаптивные подходы когнитивной визуализации данных, в том числе и пространственных. [6, 7] Таким образом, разработка научно-методических основ управления пространственным планированием и когнитивных интерфейсов мультипредметных информационных систем поддержки регионального управления является актуальной научной задачей, решение которой позволит повысить эффективность информационной поддержки решения задач регионального управления.

Одной из актуальных задач для реализации планов социально-экономического развития Арктической зоны Российской Федерации (АЗРФ) в целом, и Мурманской области, в частности, является разработка четких пространственных планов развития морских акваторий и прибрежных территорий. [8] В качестве рабочего определения морского пространственного планирования в мы используем определение Межправительственной океанографической комиссии ЮНЕСКО, в котором МПП рассматривается как общественный процесс анализа и пространственно-временного распределения деятельности человека в морских районах. Несмотря на наличие Стратегии развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года [9] и Стратегии социально-экономического развития Мурманской области до 2020 года и на период до 2025 года [10], определяющих в качестве приоритетных направлений и основных мероприятий разработку и апробацию моделей комплексного управления прибрежными зонами (КУПЗ) в арктических регионах, развитие рыбного промысла и аквакультурного кластера, технологического кластера обеспечения шельфовой

добычи в Арктике, технологического и материального обеспечения морской деятельности, туристического кластера и развитие морского круизного туризма, ни для одного из регионов РФ в настоящее время не разработаны системные планы МПП и программы комплексного управления прибрежными зонами.

2. Проблема пространственной конкуренции природопользователей

Отсутствие системного подхода в области регионального управления и пространственного планирования является одной из главных причин повышения конфликтности как между различными отраслями, так и внутри них, и негативно влияет на сохранение биологического разнообразия и экосистем регионов АЗРФ, а также на темпы социально-экономического развития регионов в целом. Выявлению потенциальных конфликтов основных стейкхолдеров региона и поиску путей разрешения конфликтов в значительной степени может способствовать разработка и апробация программ и моделей комплексного управления прибрежными зонами в арктических регионах.

Для АЗРФ одними из наиболее значимых потенциальных конфликтов с экономической точки зрения являются конфликты, вызванные противоречиями, возникающими при освоении месторождений углеводородов и рыболовством (рыбоводством) на континентальном шельфе и в прибрежных зонах. [11] Кроме того, в последние годы значительно увеличилась активность использования акваторий и участков открытого моря в северных широтах в оборонных, транспортных, туристических и военно-стратегических целях, что также приводит к возникновению конфликтов, связанных с согласованием времени и возможности использования одних и тех же участков моря.

Для Мурманской области, являющейся полиресурсным, старопромышленным регионом, ориентированным на кластерное развитие, в том числе развитие морехозяйственного кластера и эксплуатацию Северного морского пути для достижения высоких темпов социально-экономического развития, проблема пространственного планирования стоит особенно остро. Появление противоречий и потенциальных конфликтов в процессе хозяйственной деятельности, связанной с использованием различных видов ресурсов, сосредоточенных на одной территории является естественным процессом, требующим адекватных мер реагирования со стороны лиц, принимающих решения. [12] Выявление противоречий связанных с деятельностью хозяйствующих субъектов, эксплуатирующих общие природные объекты и территории, является ключевой проблемой территориального планирования, решение которой позволит добиться снижения конфликтности взаимодействия основных стейкхолдеров. Эта проблема относится к числу слабоструктурированных, многокритериальных задач, требующих учета больших объемов разнородных данных различной природы. Для ее эффективного решения необходима разработка и применение специализированных методов, технологий и программных средств, основанных на использовании современных информационных технологий. В частности, программных решений, которые можно отнести к классу систем поддержки принятия решений (СППР).

3. Постановка задачи

Для повышения эффективности принятия решений в области поддержки регионального управления в условиях АЗРФ и развития морехозяйственных

комплексов на примере морехозяйственного кластера Мурманской области разработана архитектура мультипредметной системы информационной поддержки решения задач пространственного планирования. Система обеспечивает выявление потенциальных противоречий, связанных с деятельностью основных стейкхолдеров региона, эксплуатирующих общие природные объекты, ресурсы и территории, при выполнении задач пространственного планирования. Для выявления потенциальных конфликтов природопользователей применяются разработанные авторским коллективом концептуальная модель предметной области [13, 14], семантическая модель предметной области [15–17] и технология динамической визуализации пространственных данных для построения когнитивных интерфейсов пользователей системы [18, 19].

4. Описание полученных результатов

Предлагаемая обобщенная архитектура мультипредметной системы информационной поддержки решения задач пространственного планирования разделена на несколько уровней, отличающихся реализуемыми функциями и решаемыми задачами. Архитектура описывает общую концепцию, определяет модель, функциональные модули, структуру и взаимосвязи компонентов внутри системы.

Для эффективной разработки архитектуры необходимо определить набор основных функций мультипредметной системы информационной поддержки решения задач регионального управления. К таким функциям можно отнести следующие:

- сбор, первичная обработка, валидация, гармонизация, хранение, передача и первичный анализ информации, агрегируемой системой из разнородных источников;
- обеспечение работы хранилища данных (ХД) и оперативных баз данных (ОБД), содержащих выборки (информационные срезы) данных, соответствующие требованиям решаемых задач в определенный момент времени;
- обеспечение выборок данных из хранилища данных для формирования витрин данных различных пользователей (типов пользователей);
- обработка данных и знаний по запросам пользователей средствами анализа данных, в том числе интеллектуального, реализованными в подсистеме анализа данных, включая методы работы с данными и знаниями реализованные на основе разработанных ранее концептуальной модели предметной области (КМПО), семантической модели предметной области (СМПО), модели конфликтной ситуации, онтологической модели предметной области, технологии когнитивной динамической визуализации пространственных данных (КДВПД) для систем информационной поддержки задач морского пространственного планирования с учетом интересов и противоречий различных групп природопользователей, средств комплексной оценки устойчивости арктических районных муниципальных образований в модели стратегического пространственного планирования морехозяйственной деятельности при учете средне- и долгосрочных изменений, технологий автоматизированного семантического моделирования и визуального анализа текстовых документов для быстрого понимания смысла, метода деревьев классификации и других;
- представление пользователям системы обработанной информации в виде, удобном для восприятия, в том числе и с применением методов динамической когнитивной визуализации пространственных данных в рамках когнитивных интерфейсов МИС поддержки пространственного управления регионом,

и дальнейшего анализа средствами, специфичными для данного пользователя, реализованными в рамках его автоматизированного рабочего места (АРМ);

- обеспечение функций администрирования и настройки работы системы и ее отдельных компонентов, а также управления правами доступа и привилегиями, реализованных в рамках отдельного конфигуратора системы или подсистемы конфигурирования;

- прочие функции, такие как обеспечение целостности данных, информационной безопасности, оптимизации хранимых данных и т.д.

С учетом перечисленных выше функций и иных требований, была разработана обобщенная архитектура мультипредметной системы информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности (рис. 1).

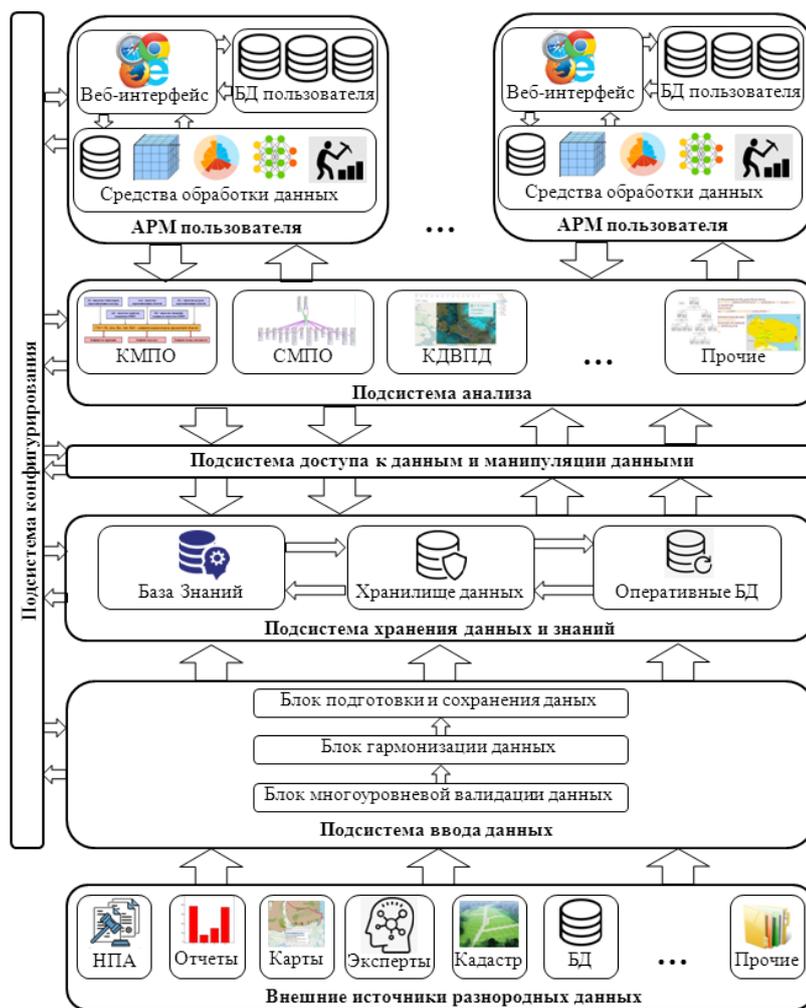


Рис. 1. Обобщенная архитектура мультипредметной системы информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности

В представленной архитектуре можно выделить несколько подсистем, реализующих различные функции. На нижнем уровне находится подсистема ввода данных, позволяющая собирать, агрегировать и хранить гетерогенные данные, полученные из различных источников в различных форматах. Разнородность получаемых данных приводит к необходимости проводить процедуры валидации и гармонизации данных, реализованные в соответствующих блоках подсистемы. Валидация данных включает в себя нескольких уровней проверки корректности поступающей информации, в зависимости от того каким именно образом определены ограничения на входящие данные. Учитывая тот факт, что валидатор проводит процедуры проверки данных на соответствие формату, синтаксису, и другим ограничениям на этапе, предшествующем этапу подготовки и сохранения данных, это позволяет предотвратить возможность поступления невалидных данных в подсистему хранения данных и знаний мультипредметной системы информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности.

Как правило, данные, прошедшие валидацию, все еще остаются достаточно разнородными по структуре, составу, форматам и другим характеристикам. Такая ситуация приводит к необходимости проведения процедур гармонизации данных, для чего в подсистему ввода данных включен блок гармонизации данных. Процесс гармонизации данных включает в себя функции по предварительной обработке данных (согласование форматов данных, вида данных, систем измерений и т.д.). Результатом этого процесса является гармонизированный массив данных, передаваемый блоку подготовки и сохранения данных для перемещения их в хранилище данных.

После проведения всех необходимых процедур подготовки данных, они передаются в подсистему хранения данных и знаний. Основой, ядром данной подсистемы является хранилище данных, содержащее всю информацию, поступающую в систему. Хранилище данных представляет собой предметно-ориентированную базу данных построенную специальным образом для эффективного обеспечения информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности. Также важными элементами подсистемы являются база знаний, которая, при реализации на практике, может состоять из нескольких специализированных баз (прикладных, общих, системных знаний) и оперативные базы данных, обеспечивающие повышение эффективности и скорости извлечения данных пользователями.

Подсистема доступа к данным и манипуляции данными является одновременно своеобразным экраном и шлюзом между уровнем данных и подсистемой анализа, использующей данные для обработки запросов пользователей. Подсистема доступа к данным и манипуляции данными реализует одновременно функции защиты данных и разрешения доступа к данным в соответствии с правами и привилегиями, указанными в поступившем запросе. Также важной функцией этой подсистемы является отделение уровня данных от уровня бизнес-логики системы. Кроме того, в случае необходимости внесения изменений в основные механизмы доступа к данным, модификации будут подвергнуты только процедуры, реализованные в подсистеме доступа к данным и манипуляции данными, а подсистема хранения данных и знаний и подсистема анализа не будут затронуты.

Такой подход позволяет существенно повысить надежность системы в целом и упростить выявление потенциальных проблем безопасности, целостности и прочих сбоев в работе, касающихся доступа к данным системы информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности.

Одной из важнейших и наиболее наукоемких подсистем, представленных в архитектуре, является подсистема анализа. Она включает в себя основной набор инструментов работы с данными и знаниями, представленных в системе и доступных пользователям для решения своих задач. Основу подсистемы анализа составляют модели, методы, технологии и инструменты анализа данных, разработанные авторским коллективом в процессе предыдущих исследований. Среди наиболее значимых для мультипредметной системы информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности компонентов можно отметить концептуальную и семантическую модели предметной области, описанные, в частности, в работах [13, 15], технологию когнитивной динамической геовизуализации для построения когнитивных интерфейсов пользователя [20], а также такие важные инструменты анализа данных, заложенные в систему, как методика комплексной оценки устойчивости арктических районных муниципальных образований в модели стратегического пространственного планирования морехозяйственной деятельности при учете средне- и долгосрочных изменений [21], метод деревьев классификации [22], технологии автоматизированного семантического моделирования и визуального анализа текстовых документов для быстрого понимания смысла, и другие.

Для взаимодействия с системой информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности пользователям предоставляется набор инструментов для построения запросов к базам данных и анализа полученной информации. Доступ ко всем основным инструментам предоставляется пользователям посредством веб-интерфейса. Такое решение позволяет снизить когнитивную нагрузку на пользователя, связанную с освоением нового для него интерфейса, так как большая часть пользователей уже имеет опыт работы с браузером и другими веб-интерфейсами, а также позволяет снизить требования к программно-аппаратной части, реализованной на стороне клиента, ввиду того, что веб-интерфейсы, как правило, не требовательны к ресурсам компьютера конечного пользователя.

Кроме того, использование веб-технологий для построения интерфейсов позволяет легко адаптировать их под потребности различных групп пользователей (аналитиков, лиц, принимающих решения (ЛПР), региональных чиновников, специалистов в области информационных технологий и других) решающих различные задачи. Также большое значение имеет возможность предоставления удаленного доступа к системе различным пользователям, в том числе находящимся за границами региона. Веб-интерфейс позволяет использовать возможности системы информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности на различных совещаниях, транслировать экран на проектор, использовать для доступа мобильные устройства и так далее.

В целом, концепция построения пользовательского интерфейса соответствует современной концепции витрин данных, позволяя обеспечивать каждого пользователя (тип пользователей) тематической, узкоспециализированной выборкой информации из хранилища данных и настраиваемым набором средств обработки данных в соответствии с решаемыми данным пользователем (типом пользователей) задачами.

Еще одним важным компонентом системы информационной поддержки решения задач регионального управления с учетом противоречий пространственного планирования хозяйственной деятельности, представленном в архитектуре, является подсистема конфигурирования или конфигуратор системы. На подсистему конфигурирования возложены функции настройки и управления взаимодействием компонентов внутри системы. Наличие единого конфигулятора позволяет существенно упростить поддержку, реализацию изменений и тонкую настройку системы на разных уровнях, что, в свою очередь, позволяет решать ряд важных задач, таких как, например, контроль и изменение прав доступа к данным, обеспечение информационной безопасности и контроля целостности данных, и других.

6. Заключение

В данной работе описано решение задачи разработки обобщенной архитектуры мультимедийной системы информационной поддержки решения задач пространственного планирования и регионального управления с использованием концептуальной и семантической моделей предметной области и технологии когнитивной геовизуализации с целью выявления потенциальных противоречий связанных с деятельностью основных стейкхолдеров региона. В соответствии с поставленной задачей, при разработке архитектуры были учтены основные функции и требования к системе, представлены модель, структура, основные компоненты системы и их взаимосвязи.

Предложенная в работе архитектура может быть использована в качестве методической основы для практической программно-аппаратной реализации мультимедийных систем поддержки регионального управления в условиях АЗРФ, в том числе и для решения задач морского (акваториального) пространственного планирования и развития морехозяйственных комплексов.

Литература

1. 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.
2. Вицентий А. В., Шишаев М. Г., Гогоберидзе Г. Г. Разработка технологии геовизуализации для систем информационной поддержки задач пространственного планирования с учетом интересов и противоречий различных групп природопользователей на основе концептуальной модели конфликта // Труды Кольского научного центра РАН. 2018. Вып. 9. Информационные технологии. С. 91–98.

3. Шишаев М. Г., Диковицкий В. В. Технология синтеза адаптивных пользовательских интерфейсов для мультипредметных информационных систем // Труды Кольского научного центра РАН. 2014. Вып. 5. Информационные технологии. С. 101–108.
4. 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). С. 69–79.
5. 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.
6. 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. 2016. Vol. 466. Pp. 483–495.
7. 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. 2017. Vol. 575. Pp. 449–459.
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. Стратегия развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2020 года URL: <http://docs.cntd.ru/document/499002465>.
10. 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.
11. 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.
12. 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.
13. Вицентий А. В., Шишаев М. Г., Гогоберидзе Г. Г., Ершова А. А., Ломов П. А. Разработка концептуальной модели предметной области для информационной поддержки морехозяйственной деятельности Мурманской области // Методологические проблемы управления макросистемами. 2018. С. 16–21.
14. Вицентий А. В., Гогоберидзе Г. Г., Шишаев М. Г., Ершова А. А., Ломов П. А., Румянцева Е. А. Концептуальная модель предметной области морехозяйственной деятельности Мурманской области как инструмент согласования деятельности морепользователей // Арктические берега: путь к

устойчивости Материалы XXVII Международной береговой конференции. 2018. С. 320–323.

15. Вицентий А. В., Шишаев М. Г., Ломов П. А., Гогоберидзе Г. Г. Разработка онтологии предметной области для формального логического вывода и визуализации в системах информационной поддержки задач морского пространственного планирования // Труды Кольского научного центра РАН. 2018. Вып. 9. Информационные технологии. С. 77–90.
16. Вицентий А. В. Разработка подхода к интеграции разнородных информационных ресурсов на основе семантического описания // Север и рынок: формирование экономического порядка. 2014. № 1 (38). С. 11–17.
17. 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. 2019. Vol. 763. Pp. 216–228.
18. Вицентий А. В. Разработка технологии динамической когнитивной геовизуализации // Труды Кольского научного центра РАН. 2015. № 3. С. 87–93.
19. Вицентий А. В., Шишаев М. Г., Диковицкий В. В. Опыт использования визуализации семантических моделей документов для экспресс контент-анализа (на примере нормативно-правовых актов, регламентирующих развитие пространственно-распределенных систем) // Труды Кольского научного центра РАН. 2018. Вып. 9. Информационные технологии. С. 99–111.
20. 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.
21. Lednova J., Chusov A., Shilin M., Gogoberidze G. Integrated indicator approach for economic-environmental assessment of coastal local municipalities // 2018 IEEE/OES Baltic International Symposium (BALTIC). 2018. Pp. 1–10.
22. 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. 2018. Vol. 662. Pp. 350–359.

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.), lead researcher of IIMM KSC RAS

Гогоберидзе Георгий Гививич

д.э.н., доцент

e-mail: gogoberidze.gg@gmail.com

Gogoberidze Georgii

Dr.of Sci., associate professor

А. В. Вицентий, М. Г. Шишаев

Институт информатики и математического моделирования ФИЦ КНЦ РАН

РАЗРАБОТКА ПОДХОДА К МОДИФИКАЦИИ И РЕДУЦИРОВАНИЮ ГЕОИЗОБРАЖЕНИЙ С УЧЕТОМ ИХ ИНФОРМАТИВНОСТИ

Аннотация

Данная статья посвящена рассмотрению проблемы разработки подхода к модификации и редуцированию синтезированных геоизображений с учетом разнородных факторов, включая вариативность целей построения геоизображения при изменении решаемой задачи. В статье рассмотрены вопросы создания обобщенной модели пользователя мультипредметной географической информационной системы и модели задачи, необходимых для оценки качества геоизображения. Предлагается оценивать качество геоизображений с точки зрения решаемой задачи на основе оценки количества информации и уровня информативности изображения.

Ключевые слова:

геоизображение, когнитивная геовизуализация, поддержка принятия решений, редуцирование геоизображения, оценка информативности геоизображения, пространственная ситуация, когнитивный интерфейс, модель пользователя ГИС, динамическая визуализация пространственных данных.

A.V. Vicentiy, M. G. Shishaev

DEVELOPMENT OF THE MODIFICATION AND REDUCTION OF GEOIMAGES APPROACH TAKING INTO ACCOUNT ITS INFORMATIVENESS

Abstract

This paper is devoted to the problem of developing an approach to the modification and reduction of synthesized geoimages taking into account heterogeneous factors, including the variability of the goals of constructing geoimages when changing the problem being solved. The paper considers the issues of creating a generalized user model of a multi-subject geographical information system and a task model necessary for assessing the quality of a geographic image. It is proposed to evaluate the quality of geoimages from the point of view of the problem being solved on the basis of assessing the amount of information and the level of information content of the image.

Keywords:

geoimage, cognitive geovisualization, decision support, reduction of geoimage, assessment of geoimage informativeness, spatial situation, cognitive interface, GIS user model, dynamic visualization of spatial data.

1. Введение

В данной статье представлены результаты работ, выполненных в рамках разработки методического и алгоритмического обеспечения динамического синтеза когнитивных геоизображений с учетом разнородных факторов, включая вариативность целей построения геоизображения, обеспечивающего модификацию и редуцирование геоизображений при изменении решаемой задачи по теме «Развитие методологии построения интеллектуальных мультипредметных систем информационной поддержки регионального развития на примере территорий Арктической зоны Российской Федерации (АЗРФ)» (№ 0226-2019-0036).

Различные системы обработки и представления пространственной информации давно используются в качестве одного из инструментов поддержки принятия решений в различных сферах деятельности. Наибольшую эффективность системы такого рода показывают при решении широкого класса задач, связанных с территориальным планированием и управлением разного масштаба – от межгосударственного уровня, до уровня конкретных поселений и небольших участков пространства. Программно-аппаратные комплексы и информационные системы, осуществляющие сбор, хранение, обработку, анализ и визуализацию пространственно-временных данных (геоданных) и соответствующей атрибутивной информации выделяют в отдельный класс информационных систем, называемых географическими информационными системами (геоинформационными системами, ГИС).

Однако ГИС также является и человеко-машинной системой, так как основное ее предназначение это визуализация имеющейся информации и предоставление ее конечному пользователю для дальнейшего анализа и принятия решений. При работе с ГИС пользователь имеет дело с электронными картами или, говоря более точно, цифровыми геоизображениями (географическими изображениями). Геоизображение, в общем виде, представляет собой масштабную, пространственно-временную, генерализованную модель объектов или процессов окружающего мира, которая представлена в образно-знаковой форме [1, 2]. При этом, синтезируемые ГИС геоизображения обладают уникальным эвристическим и гносеологическим потенциалом, который может быть раскрыт только при непосредственном участии человека в процессе интерпретации и визуального анализа геоизображений.

Включение человека (лица, принимающего решения (ЛПР)) в контур обработки визуальной картографической информации налагает дополнительные требования на визуализацию данных и вызывает ряд дополнительных проблем, связанных с особенностями восприятия визуальной информации как отдельными группами (категориями) пользователей, так и каждым пользователем индивидуально. Эти особенности могут быть связаны с когнитивными способностями, имеющимися у конкретного пользователя, или особенностями и условиями восприятия визуальной информации отдельными категориями пользователей, профессиональными знаниями, опытом работы, как с географическими информационными системами в целом, так и с аналогичными решаемой задачами или схожими предметными областями, и многими другими характеристиками.

Таким образом, можно утверждать, что разработка различных подходов, позволяющих повысить эффективность работы человека с синтезированными геоизображениями является важной задачей, решение которой повысит ценность и эффективность применения ГИС в принятии решений.

2. Проблема синтеза, модификации и редуцирования геоизображений

Взаимодействие пользователя с ГИС представляет собой последовательность запросов к геоинформационному сервису в диалоговом режиме или с помощью языка запросов, например, SQL. Результатом обработки запросов является синтезированное системой геоизображение, которое строится подсистемой визуализации из примитивов, распределенных по тематическим

слоям. Такой подход к визуализации распространен в подавляющем большинстве современных ГИС и геосервисов (ArcGis, QGis, MapInfo, Google Earth, Google Maps, Яндекс.Карты и других). Основным недостатком такого подхода является зависимость эффективности визуализации данных как от решаемых задач, так и от субъективных оценок пользователей, задающих поисковые запросы и оценивающих полученный результат с прагматической точки зрения. В результате, ГИС формирует, как правило, избыточные визуальные образы, что негативно влияет на скорость и качество решений, принимаемых с их использованием.

Однако, несмотря на субъективность прагматической оценки геоизображений для целей поддержки принятия решений, проблема избыточности синтезируемых визуальных образов может быть в некоторой степени решена за счет разработки и применения процедур модификации и редуцирования геоизображений с учетом оценки степени их информативности и субъективной «полезности» для решения конкретных задач (класса задач).

Проблемы оценки количества информации, разработки мер информации и способов снижения информационной избыточности относятся к фундаментальным проблемам теории информации. В данной работе мы будем опираться на прагматическую меру информации, позволяющую учитывать полезность информации для конечного пользователя в контексте достижения поставленной цели, которой может являться принятие решения на основе синтезированного геоизображения.

Мы также исходим из того, что для решения проблемы избыточности синтезированных геоизображений наиболее эффективным способом является уменьшение количества примитивов до некоторого «оптимального» для решения задачи уровня. Кроме того, уменьшение визуальной загруженности изображения позволит снизить когнитивную нагрузку на лицо, принимающее решения, что позволит снизить риск принятия непродуманных и необоснованных решений и отсрочить наступление периода утомляемости. [3, 4]

3. Постановка задачи

С учетом вышесказанного, постановку задачи в общем виде можно сформулировать следующим образом. Необходимо предложить такой способ модификации геоизображения, который позволил бы минимизировать избыточность изображения и при этом максимизировать уровень его информативности:

$$\begin{cases} E(g) \rightarrow \min \\ V(g) \rightarrow \min \\ I(g) \rightarrow \max \end{cases}$$

где $E(g)$ – избыточность геоизображения для решения определенной задачи определенным пользователем; $V(g)$ — объем геоизображения – количественная мера, позволяющая дать объективную оценку «размера» изображения, например, в мегабайтах; $I(g)$ — информативность (субъективная «полезность») геоизображения для решения определенной задачи определенным пользователем; g — модифицируемое геоизображение.

Информативность можно представить в виде функции следующим образом:

$$f : G \times T \times U \rightarrow \mathbb{R}, \text{ где:}$$

$G = \{g_i\}, i = 1..N$ – множество геоизображений;

$T = \{t_j\}, j = 1..M$ – множество задач;

$U = \{u_k\}, k = 1..K$ – множество пользователей.

Однако на практике в такой строгой постановке решение задачи вряд ли достижимо, поэтому имеет смысл смягчить постановку задачи следующим образом. Пусть есть некоторое синтезированное геоизображение из множества изображений, синтезированных ГИС: $g \in G$. На основе анализа данного изображения пользователь из множества пользователей данной ГИС $u \in U$ пытается решить некоторую задачу из множества решаемых задач $t \in T$.

Тогда задача может быть выражена в такой постановке. Необходимо предложить такой способ модификации геоизображения, который позволил бы снизить избыточность изображения, при этом, не снижая уровень его информативности или не понижая уровень информативности ниже некоторого минимального уровня информативности геоизображения, необходимого для успешного решения данной задачи:

$$\begin{cases} E(g) > E(g)^* \\ V(g) \geq V(g)^* \\ (I(g) \leq I(g)^*) \vee (I(g)_{\min} \leq I(g)^*) \end{cases}$$

где $E(g)^*$ – избыточность геоизображения для решения определенной задачи определенным пользователем после процедуры модификации; $V(g)^*$ – объем геоизображения – количественная мера, позволяющая дать объективную оценку «размера» изображения после процедуры модификации; $I(g)^*$ – информативность геоизображения для решения определенной задачи определенным пользователем после процедуры модификации; $I(g)_{\min}$ – минимальный уровень информативности геоизображения, необходимый для решения задачи.

Кроме того, на данном этапе вводятся следующие допущения:

Допущение 1. Синтезированное изображение изначально обладает некоторой избыточностью: $E(g) > E(g)^*$,

Допущение 2. Уровень информативности исходного геоизображения достаточен для решения задачи пользователем.: $I(g) \geq I(g)_{\min}$,

С учетом более мягкой постановки задачи и допущений, модифицирование синтезированного геоизображения в реальной ГИС может быть связано, прежде всего, с удалением из изображения примитивов, не относящихся к решаемой задаче. Удаление таких примитивов не приведет к снижению информативности с точки зрения пользователя, но снизит общую информационную насыщенность изображения (избыточность) и объем геоизображения, измеряемый в мегабайтах. Такой способ модификации геоизображения далее будем называть редуцированием.

4. Описание подхода к редуцированию геоизображений

Из постановки задачи и определения информативности понятно, что важнейшими элементами для осуществления успешного редуцирования

геоизображений являются модель пользователя, модель задачи и модель геоизображения. Рассмотрим эти модели более подробно.

Модель пользователя. Модели пользователей применяются при реализации различных информационных систем и геоинформационные системы в этом смысле не являются исключением. Так, в работе [5] предлагается определять модель пользователя через набор его свойств. Предлагается относить пользователей к разным группам с целью адаптации интерфейсов, повышения эффективности взаимодействия с ГИС и снижения вероятности ошибок пользователя. Подчеркивается необходимость учитывать особенности психики пользователя и его ментальных особенностей при создании модели и отнесении пользователей к различным группам. В работе [6] вся совокупность пользователей ГИС воспринимается как система с естественным поведением, которая может эволюционировать и приспосабливаться к изменениям окружающего мира. Для учета характерных особенностей структуры и поведения естественных систем, модели пользователей предлагается реализовать на основе математического аппарата теории нечетких множеств. Другой подход предлагается в работе [7]. Суть его состоит в адаптации возможностей интерфейса пользователя на основе предоставления ему базового комплекта инструментов ГИС и специального механизма адаптации возможностей базового инструментария под задачи, решаемые пользователем. Эта технология также поддерживает деление пользователей на группы на основе общности решаемых пользователями задач и используемых инструментов.

Для оценки уровня информативности геоизображений, одной из наиболее подходящих моделей, на наш взгляд, является модель пользователя, описанная в работе [8]. В основу описанной в работе модели пользователя положены такие понятия, как ментальная модель и перцептивные стереотипы. Принадлежность пользователя к той или иной группе (категории) определяется схожестью представлений пользователей об объектах реального мира, которая порождается близостью их ментальных моделей и способа ранжирования атрибутов объектов реального мира по значимости. Для снижения избыточности изображения с помощью редуцирования, понимание значимости того или иного примитива-кандидата на удаление для успешного решения задачи, с точки зрения пользователя, является ключевым. Поэтому модель пользователя, построенная на основе оценки и учета значимости идентифицирующих объект атрибутов хорошо подходит для расчета информативности геоизображения.

Модель задачи. Рассматриваемые в данной работе мультипредметные ГИС, предназначены, главным образом, для оказания помощи пользователям в принятии более обоснованных решений за счет лучшей информированности. ГИС позволяет повысить эффективность и скорость принятия решений, представляя результаты запросов пользователя и анализа данных в наглядной картографической форме, удобной для восприятия. Для построения модели задачи, решаемой пользователем средствами ГИС, могут быть применены методы и подходы системного анализа и концептуального моделирования [9, 10], ситуационного концептуального управления [11], функционально-целевой подход [12] и другие. В данной работе мы предлагаем использовать модель задачи, основанную на формальном определении мультипредметного информационного ресурса, предложенном в [13]. Такая модель позволяет

учитывать мультипредметный характер ГИС и хорошо согласуется с моделью пользователя.

Пусть определено множество категорий пользователей $U = \{u_k\}$, $k = 1..K$ и множество ресурсов ГИС $R = \{r_l\}$, $l = 1..L$. Тогда можно определить функцию $z: U \times R \rightarrow \mathbb{R}$, которая характеризует прагматическую ценность некоторого ресурса ГИС с точки зрения пользователя.

Пусть также некоторый пользователь $u_k \in U$, генерирует последовательность запросов к ГИС вида: $Q^{u_k}_i (q_{\text{spatial}}, q_{\text{temporal}}, q_{\text{semantic}})$, $i = 1..N$, определяющую ограничения интересующих пользователя данных. Тогда, модель задачи может быть представлена как подбор таких ресурсов ГИС для синтеза визуального образа $g \in G$, которые удовлетворяют пространственным, временным и семантическим ограничениям, указанным в запросах пользователя и для которых функция z не ниже некоторого порогового значения.

$$\begin{cases} g = \bigcup_{l=1}^L r_l \subseteq R \mid r_{\text{spatial}} \approx q_{\text{spatial}}, r_{\text{temporal}} \approx q_{\text{temporal}}, r_{\text{semantic}} \approx q_{\text{semantic}} \\ z \geq z_{\text{min}} \end{cases}$$

Где q_{spatial} , q_{temporal} , q_{semantic} , r_{spatial} , r_{temporal} , r_{semantic} - пространственные, временные и семантические атрибуты запросов и ресурсов ГИС, соответственно. Знак примерного равенства в данном контексте означает, что пространственные, временные и семантические атрибуты отобранных ресурсов соответствуют пространственным, временным и семантическим атрибутам запроса пользователя.

Модель геоизображения. В настоящее время выделяют множество разновидностей геоизображений в зависимости от целей их использования, способов получения, основных характеристик и по другим основаниям. Созданием и разработкой общей теории геоизображений занимается отдельная научная дисциплина – геоиконика. В рамках данной работы, мы предлагаем использовать модель геоизображения, представленную в виде множества визуализированных примитивов с учетом пространственных, временных и семантических ограничений: $g = \langle P, Spa, Tem, Sem \rangle$, где

$P = \{p_x\}$, $x = 1..X$ – множество примитивов, включенных в геоизображение;

$Spa = \{pol_s\}$, $s = 1..S$ – множество полигонов, определяющих пространственные ограничения геоизображения. Каждый $pol_s = ((x_1, y_1), .. (x_n, y_n))$ представляет собой полигон, определяемый набором точек с координатами (x^*, y^*) .

$Tem = \{tem_y\}$, $y = 1..Y$ - множество интервалов, определяющих временные ограничения геоизображения. Каждый $\{tem_y\} = \{tem_y^{\text{start}}, tem_y^{\text{end}}\}$ – представляет собой пару значений, определяющих начало и конец временного интервала.

$Sem = \{sem_w\}$, $w = 1..W$ - множество семантических ограничений геоизображения, где каждому элементу множества Sem ставится в соответствие некоторый набор семантических значений, определенных в данной реализации ГИС.

Таким образом, используя формальное определение моделей пользователя, задачи и геоизображения, можно разработать конкретные процедуры редуцирования геоизображений с учетом прагматической оценки информативности получаемых визуальных образов.

6. Заключение

В данной работе описаны элементы подхода к редуцированию геоизображений с учетом их информативности. Суть подхода состоит в том, чтобы снижать избыточность и объективный объем геоизображения без ущерба для уровня информативности. Уровень информативности предлагается понимать на основе оценки прагматической меры информации, что позволяет учесть как особенности решаемой задачи, так и характерные черты пользователя географической информационной системы. В соответствии с поставленной задачей, в работе представлены формализованная постановка задачи модификации изображения, а также модели пользователя, задачи и геоизображения, необходимые для реализации процедур редуцирования. Предложенные в работе модели могут быть использованы в качестве методической основы для программной реализации мультипредметных геоинформационных систем.

Литература

1. Берлянт А. М. Геоиконика. М.: Астрея, 1996. 208 с.
2. Берлянт А. М. Картографический словарь. М.; Научный мир, 2005. 424 с.
3. Тоффлер Э. Шок будущего: Пер. с англ. М.: ООО «Издательство АСТ», 2002. 557 с.
4. Robert L. Solso, M. Kimberly MacLin and Otto H. MacLin Cognitive Psychology. Allyn & Bacon, 2008. – 592 p.
5. Иванов С. Ю. Проектирование интерфейса и модель пользователя // Новые информационные технологии в автоматизированных системах. 2013. № 16. С. 227–230.
6. Диденко Д. А. Разработка модели оценки качества информации в ГИС // Известия ЮФУ. Технические науки. 2010. № 4. С. 210–216.
7. Цветков В. Я. Особенности проектирования ГИС-пользователя на основе базового комплекта ГИС «Карта-2011» / В. Я. Цветков, С. Г. Дышленко // Землеустройство, кадастр и мониторинг земель. 2010. № 8. С. 79–84.
8. Диковицкий В. В., Ломов П. А., Шишаев М. Г. Формализация задачи построения когнитивных пользовательских интерфейсов мультипредметных информационных ресурсов // Труды Кольского научного центра РАН. Информационные технологии. 2013. Вып. 4. патиты: Изд-во КНЦ РАН, 2013. С. 90–98.
9. Олейник А. Г. Инструментальная система комплексного концептуального моделирования задач регионального управления // Информационные ресурсы России. 2005. № 2. С. 33–36.
10. Вицентий А. В., Шишаев М. Г., Гогоберидзе Г. Г. Разработка технологии геовизуализации для систем информационной поддержки задач пространственного планирования с учетом интересов и противоречий различных групп природопользователей на основе концептуальной модели конфликта // Труды Кольского научного центра РАН. 2018. Вып. 9. Информационные технологии. С. 91–98.

11. Фридман, А. Я., Фридман О. В. Модели и методы координации решений по управлению региональным промышленно-природным комплексом // Вестник КНЦ. Апатиты: КНЦ РАН, 2012. Вып. 2. С. 65–78.
12. Игнатъев М. Б., Путилов В. А., Смольков Г. Я. Модели и системы управления комплексными экспериментальными исследованиями. М.: Наука, 1986. 232 с.
13. Шишаев М. Г., Диковицкий В. В. Технология синтеза адаптивных пользовательских интерфейсов для мультипредметных информационных систем // Труды Кольского научного центра РАН. 2014. № 5 (24). С. 101–108.

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.), lead researcher of IIMM KSC RAS

DOI: 10.25702/KSC.2307-5252.2019.9.59-68

УДК 004.9

А. В. Шестаков¹, А. С. Шемякин^{1,2}, С. Ю. Яковлев^{1,2}

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² Филиал ФГБОУ ВО МАГУ в г. Апатиты

ТРЕХМЕРНОЕ МОДЕЛИРОВАНИЕ В ЗАДАЧАХ РАСЧЁТА И ВИЗУАЛИЗАЦИИ РЕГИОНАЛЬНЫХ ТЕХНОГЕННО-ЭКОЛОГИЧЕСКИХ РИСКОВ*

Аннотация

В работе обсуждаются возможности применения геоинформационных технологий для решения задач оценки промышленно-природных опасностей и рисков. Основное внимание уделено вопросам учёта рельефа прилегающей местности. На простом учебном примере расчёта зоны поражения при аварии показаны значимость этого фактора и возможности использования современных геоинформационных программных средств.

Ключевые слова:

геоинформационные технологии, ГИС, риски, чрезвычайные ситуации, 3D-моделирование.

* Работа выполнена в рамках государственного задания Министерства науки и высшего образования РФ (тема НИР № 0226-2019-0035), а также при частичной финансовой поддержке РФФИ (проект № 18-07-00167-а).

A. V. Shestakov, A. S. Shemyakin, S. Yu. Yakovlev

3D MODELING IN TASKS OF CALCULATION AND VISUALIZATION OF REGIONAL TECHNOGENICAL-ECOLOGICAL RISKS

Abstract

The paper discusses possibility of applying geographic information technologies to solve problems of assessing industrial-natural hazards and risks. The main attention is paid to consideration of surrounding area relief. On a simple educational example of calculating affected area in an accident, significance of this factor and possibility of using modern geoinformation software are shown.

Keywords:

geoinformation technologies, GIS, risks, emergency situations, 3D-modeling.

Введение

Проблема использования геоинформационных технологий для моделирования опасностей и рисков охватывает достаточно широкий круг теоретических вопросов и приложений. Чтобы ограничить рамки исследования, в настоящей работе трёхмерное моделирование будет рассматриваться как развитие спроектированной ранее информационно-аналитической системы прогнозирования, предупреждения и ликвидации последствий чрезвычайных ситуаций техногенного и природного характера (ИАС ПБЧС), связанных с хозяйственной и иной деятельностью в Арктической зоне РФ. Структура и функции ИАС ПБЧС описаны в [1, 2].

Разработка и использование пространственных моделей промышленно-природных опасностей обусловлены стремлением повысить точность и наглядность результатов. Целесообразность привлечения таких моделей определяется балансом между возможным «выигрышем» и «затратами» на их создание и/или эксплуатацию. Так, в работе [3] в качестве критерия предлагается величина возможного ущерба от аварии или чрезвычайной ситуации (ЧС). Актуальность трёхмерного моделирования промышленно-природных комплексов и опасностей обсуждается в [3, 3], при этом основное внимание уделено визуализации как средству, направленному на снижение неопределённости.

В настоящей работе, наряду с блоком визуализации (см. [3]), также рассматривается и аналитический блок, т.е. расчётные математические методы и модели анализа и оценки риска.

В первом разделе выполнен обзор методов, технологий и систем 3D-моделирования промышленно-природных рисков.

Во втором разделе продемонстрирован пример использования геоинформационных технологий для задачи расчёта и визуализации аварии.

В третьем разделе предложена структура ИАС ПБЧС, дополненной функцией пространственного моделирования.

1. Обзор существующих решений в сфере трёхмерного моделирования ЧС

Хорошее представление об уровне зарубежных исследований даёт статья [5], где показаны возможности применения трёхмерных моделей в задачах борьбы с ЧС. Интересно отметить масштабы моделируемых объектов: используются как модели отдельных зданий, так и крупных промышленных

комплексов. В некоторых случаях создаются трёхмерные модели крупных территорий. Например, в [5] упоминается, что в Южнокалифорнийском центре по прогнозированию землетрясений создана модель совокупности тектонических блоков, охватывающая значительную часть штата Калифорния в США.

Что касается отечественных исследований, отметим прежде всего недостаточность нормативно-методической базы: авторам удалось обнаружить лишь ведомственные рекомендации 2009 года [6]. Несмотря на малое общее количество российских работ, можно выделить несколько исследований.

В работе [7] описывается трёхмерная модель аварийного разлива нефти из резервуара. Расчёты в модели могут проводиться по одной из двух методик на выбор пользователя: математической модели на основе уравнения Бернулли либо методики, указанной в [8]. Входными данными модели являются геометрические параметры резервуара, геометрические параметры отверстия и время истечения нефти. Помимо этого на вход расчётной модели подаются геоинформационные данные, такие как отметки высот и изолинии. На выходе расчётной модели формируется двумерное представление разлива нефти, на основе которого формируется трёхмерная модель разлива при помощи сторонних средств. В частности, в [7] для этого используется приложение ArcScene от компании ESRI [9]. Для большей наглядности полученное трёхмерное изображение дополняется вручную созданными объёмными моделями резервуаров. Для создания моделей резервуаров использовался программный продукт Sketch Up [10].

В работе [11] предлагается теоретико-множественное описание объектов трёхмерной модели промышленного объекта, на основе которого задаются отношения объектов в трёхмерном пространстве. На основании этого описания становится возможным переход от двумерного представления пространственных данных к трёхмерному. Необходимость разработки предлагаемого теоретико-множественного описания обуславливается тем, что в задачах планирования борьбы с чрезвычайными ситуациями использование трёхмерных моделей имеет ряд преимуществ. При этом большая часть существующей пространственной информации, собранной для нужд борьбы с ЧС – двумерная. На основе предложенного теоретико-множественного описания авторы [11] разработали трёхмерную модель объектов нефтяной промышленности.

Результаты, полученные в [11], используются в [12] для задачи формирования информационного пространства на основе технологий трёхмерного геоинформационного моделирования. В рамках поставленной задачи разработана схема принятия решений по борьбе с ЧС с использованием этих технологий. На основе предложенной схемы составлен алгоритм использования трёхмерной пространственной информации при оперативном планировании борьбы с ЧС. Полученные результаты апробированы на программной системе визуализации промышленных площадок нефтеперерабатывающего предприятия.

Отметим также две практические разработки [13] и [15].

В [13] моделируется распространение аварийно химически опасного вещества (АХОВ) на территории населённого пункта в результате аварии на химически опасном производстве. Расчёты в подобных случаях должны проводиться на основе методики [14], подразумевающей выбор нужных значений из таблиц. Такой подход удобен для использования людьми, но не удобен для автоматизации расчётов на компьютере. По этой причине авторы [13]

модифицировали методику [14]: табличные данные были заменены на функции, полученные в результате кубической аппроксимации этих данных. В результате для проведения расчётов нужно выбрать необходимую функцию и подставить соответствующие параметры. Упрощённая методика реализована в виде web-приложения, использующего Яндекс-карты.

Методика [14] также использовалась в системе моделирования [15], использующей полноценные ГИС-технологии. Важной особенностью предлагаемой системы является то, что размер зоны поражения производится с учётом влияния местности. Однако в контексте работы [15] под влиянием учёта местности подразумевается не рельеф, а обобщённый тип прилегающей территории: городская застройка, сельская застройка или лесные\садовые насаждения. К сожалению, не уточняется, какая именно геоинформационная система использовалась автором.

Анализ отечественных работ показывает, что использование трёхмерной геоинформации при моделировании ЧС является актуальным. Однако, в исследованиях присутствует некоторая «разобщённость» используемых технологий: для анализа пространственной информации используются одни программные средства, для построения изображения – другие, расчётные модели реализуются в виде отдельных программ. В результате трёхмерная модель развития ЧС является результатом работы комплекса программ, что создаёт некоторые неудобства при модификациях модели и обновлениях используемых программных продуктов.

При этом современные ГИС позволяют всё реализовать в рамках единой платформы. Например, открытая ГИС QGIS [16] имеет как средства построения трёхмерных изображений [17], так и возможность создания дополнительных модулей к системе [18].

Иными словами, в настоящий момент уровень развития ГИС-технологий таков, что реализовать модель возникновения и развития ЧС можно средствами единой геоинформационной системы, имея при этом возможность использовать множество функций, заложенных в систему и обращаясь к пространственным данным через API-интерфейс. В данной работе даётся пример создания трёхмерной модели распространения условного опасного вещества (УОВ) средствами QGIS, иллюстрирующий возможности этой ГИС для решения задач борьбы с ЧС.

2. Использование QGIS для задачи моделирования распространения УОВ

Создание трёхмерной модели будем рассматривать на примере следующей ситуации. В горном районе расположен склад УОВ, представляющего собой жидкость плотностью 900 кг/м^3 , способную растекаться на ровной твёрдой поверхности слоем толщиной 5 см. Форма пятна разлившейся жидкости в этом случае будет представлять собой круг. Пусть в результате аварии произошла утечка 100000 м^3 жидкости.

Исходя из заданных условий, без учёта рельефа местности, зона поражения УОВ представляет собой круг радиусом около 800 м и площадью 2 млн м^2 (

Рис. 1).

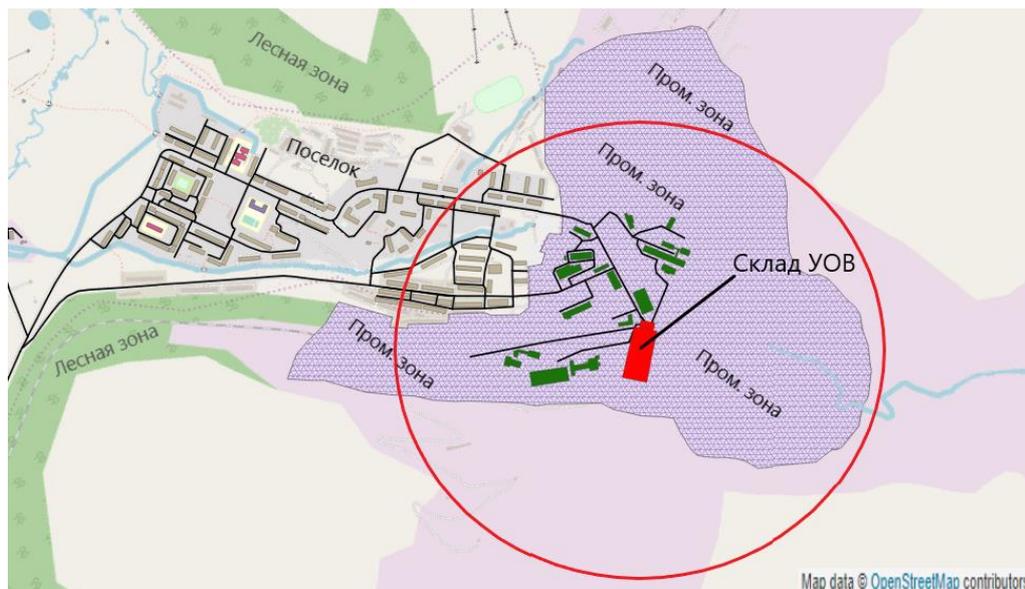


Рис. 1. Зона поражения, рассчитанная без учёта рельефа местности

Покажем, как может измениться зона поражения, если в расчётах учитывать рельеф.

Для построения трёхмерной модели будем использовать ГИС QGIS. В качестве источника данных о рельефе местности предлагается использовать ASTER GDEM [19] (цифровая модель рельефа, ЦМР). Источником двумерных картографических данных будут данные сервиса OpenStreetMap.ru [20].

Для удобства импорта в QGIS ЦМР была предварительно обработана при помощи демонстрационной версии Global Mapper [21]. Данное программное обеспечение было выбрано из-за легкости получения нужных карт высот и большого функционала для их редактирования. Предварительная обработка заключалась в том, что исходные данные ASTER GDEM были загружены в Global Mapper и затем требуемый участок земного полотна экспортировался в файл в формате GeoTIFF [22], где помимо основного изображения хранятся метаданные о географической привязке. Этот файл совместно с картографическими данными является основой для построения трёхмерной модели местности (

Рис. 2).

Средствами QGIS ЦМР совмещается с картографическими данными и в результате строится наглядное трёхмерное изображение. На основе простой расчётной модели была построена зона поражения УОВ, учитывающая рельеф местности. В этом случае площадь поражённой территории составляет примерно 17,8 тыс.м². Это обуславливается тем, что вытекающее УОВ в силу особенностей окружающего рельефа будет заполнять близлежащий карьер.

Для большей наглядности на р показаны как зона поражения

Рис. 1, (красная линия), так и зона поражения, построенная с учётом рельефа местности (оранжевая заливка).

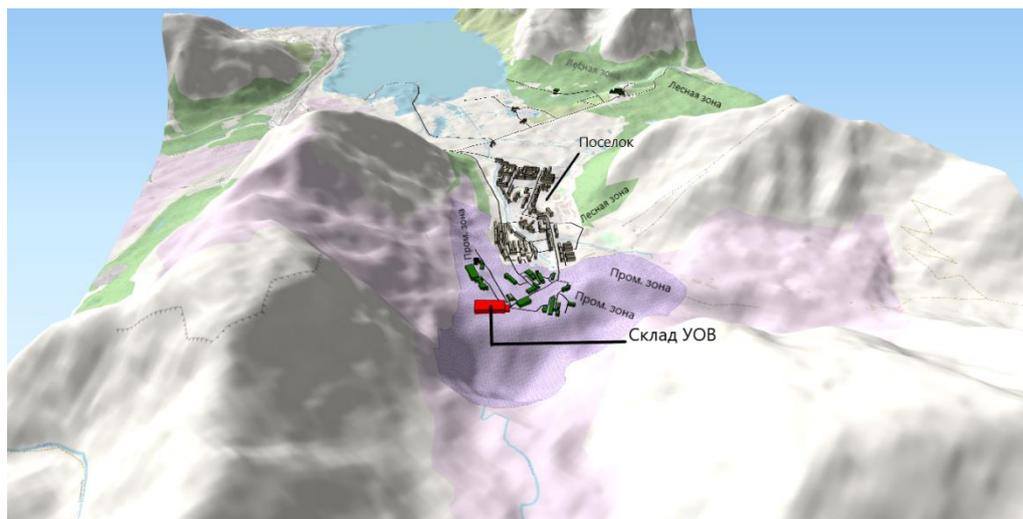


Рис. 2. Трёхмерная модель местности

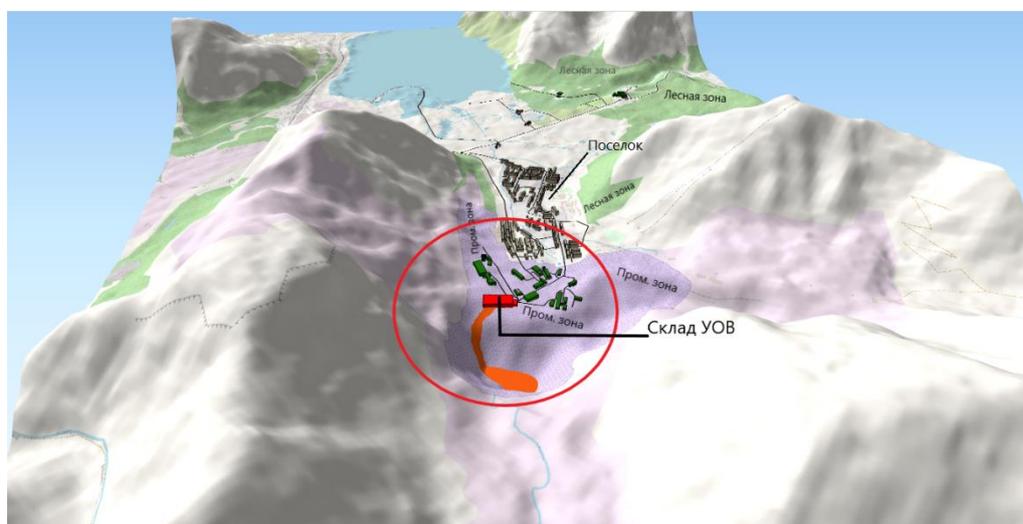


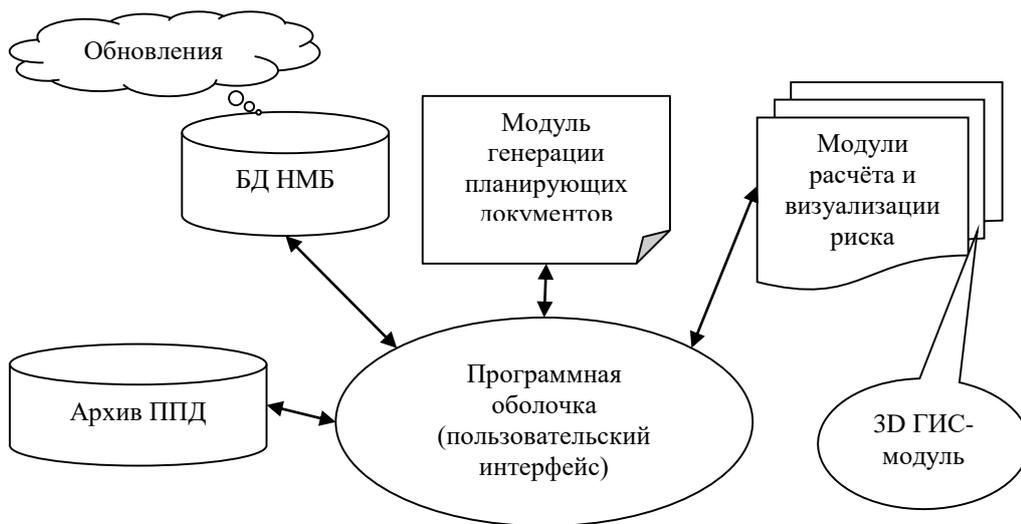
Рис. 3. Зона поражения с учётом и без учёта рельефа

Можно констатировать, что учёт рельефа местности в расчётах может существенно повлиять на планирование мероприятий по борьбе с ЧС, поскольку значительно могут изменяться размер и форма зоны поражения. В нашем примере, если мы учитываем рельеф при расчёте площади разлива, то мероприятия по локализации и ликвидации последствий аварии территориально ограничены близлежащим карьером. В противном случае – борьба с ЧС должна вестись на значительно большей площади и включать в себя мероприятия по эвакуации населения посёлка, поскольку он попадает в зону поражения. Ясно, что отклонения могут случиться и в противоположном направлении - когда учёт рельефа приведёт к существенному возрастанию возможного ущерба.

3. Структура и функции модифицированной ИАС ПБЧС

Ранее в [2] были описаны основные блоки информационно-аналитической системы. Обобщённая архитектура системы, снабжённой функцией пространственного моделирования, представлена на Рис. 4.

Предполагается, что для каждого модуля системы будет разработан программный интерфейс. В случае, когда используются готовые программные решения (например, СУБД для работы с базами данных), то по возможности будет использоваться существующий API. Организация взаимодействия между модулями через предоставляемые интерфейсы позволит реализовать работу ИАС ПБЧС в распределённой сети.



ППД – пользовательская планирующая документация
БД НМБ – база данных нормативно-методической документации
3D ГИС-модуль – модуль трёхмерного моделирования, использующий ГИС-технологии

Рис. 4. Общая архитектура системы

Работу 3D-модуля ГИС для рассматриваемого в статье примера можно схематично изобразить следующим образом – см. Рис. 5

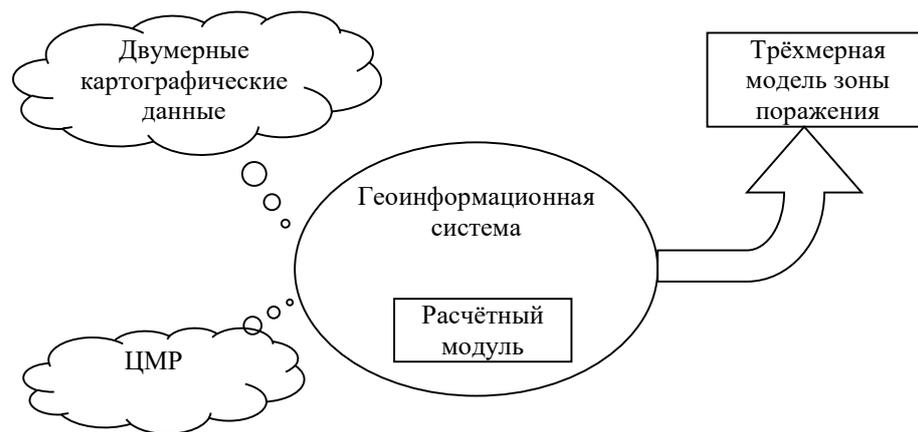


Рис. 5. Схема работы 3D ГИС-модуля

Заключение

Продемонстрированы возможности применения геоинформационных технологий в задачах расчёта и визуализации региональных техногенно-экологических рисков. В качестве модельного примера рассматривалась аварийная ситуация, в результате которой произошла утечка условного опасного вещества. С использованием штатных средств QGIS и цифровой модели рельефа была построена зона поражения, учитывающая рельеф местности. Расчётную модель можно реализовать как модуль расширения QGIS, тем самым задачу оценки опасности можно решать на платформе выбранной геоинформационной системы.

Таким образом, отличительной особенностью данной статьи является то, что задачу расчёта и визуализации риска предлагается решать в рамках единой программной платформы, в то время как в существующих работах для этих целей использовались наборы автономных программ. Это позволяет осуществить наращивание спроектированной ранее информационно-аналитической системы [1, 2] блоком пространственного моделирования.

Отметим, что в статье подробно рассмотрен лишь один вопрос – расчёт зоны поражения при аварии. Однако трёхмерное моделирование может понадобиться и при оценке других показателей, таких как вероятность (частота) аварии, возможный ущерб от аварии, достаточный состав сил и средств борьбы с ЧС и др. [1]. Эти направления авторы планируют осветить в последующих работах.

Результаты исследования смогут найти применение при реализации основных направлений государственной политики Российской Федерации в Арктике на период до 2035 года на территории Мурманской области в части разработки средств поддержки принятия решений для информационно-аналитического обеспечения региональных ситуационных центров [23].

Литература

1. Шемякин А. С., Яковлев С. Ю., Олейник Ю. А., Маслобоев А. В. Автоматизация разработки планирующей документации по снижению

промышленно-экологических рисков // Вестник Иркутского государственного технического университета. 2016. Т. 20. № 9. С. 74–85. DOI: 10.21285/1814-3520-2016-9-74-85.

2. Yakovlev S., Putilov V., Masloboev A. Information and analytical support for the industrial and ecological safety management of Arctic communications // IOP Conference Series: Earth and Environmental Science. Vol. 302 4th International Scientific conference Arctic: History and Modernity 17-18 April 2019, Saint Petersburg, Russian Federation. DOI: 10.1088/1755-1315/302/1/012032.
3. Шемякин, А. С., Яковлев, С. Ю., Маслобоев, А. В. Компьютерная визуализация в задачах информационной поддержки принятия решений // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. 46 (3). 2019. С. 540–552. DOI: 10.18413/2411-3808-2019-46-3-540-552.
4. Шемякин А. С., Яковлев С. Ю., Рыженко А. А., Тихонов Д. Е. Разработка графического компонента трёхмерной модели регионального промышленно-природного комплекса (на примере Хибинского горнорудного района) // Труды Кольского научного центра РАН. Информационные технологии. Вып. 2. Апатиты: Изд-во КНЦ РАН. 4/2011(7). 2011. С. 156–163.
5. Клименко А. С. Методы и средства повышения реалистичности моделирования и визуализации в системах виртуального окружения // В сборнике «Международная конференция RESILIENCE2014 международного центра по ядерной безопасности института физико-технической информатики». Протвино, 25–28 ноября 2014 г. 2014. С. 110–134.
6. Рекомендации по созданию трёхмерных геоизображений (моделей) территорий и объектов жизнеобеспечения, потенциально-опасных, критически важных для национальной безопасности / Нормативно-методические документы по вопросам организации выполнения научно-исследовательских и опытно-конструкторских работ. М.: ВНИИ ГОЧС, 2009. 41 с.
7. Баишева А. Р., Сайфутдинова Г. М. Геоинформационное моделирование аварийного разлива нефти при решении задач трёхмерной визуализации чрезвычайных ситуаций на территории резервуарного парка // Геоинформационные технологии в проектировании и создании корпоративных информационных систем. Уфа: ГОУ ВПО УГАТУ. 2012 С. 109–115.
8. Постановление Правительства РФ от 21.08.2000 № 613 «О неотложных мерах по предупреждению и ликвидации аварийных разливов нефти и нефтепродуктов» (ред. от 15.04.2002).
9. Работа в ArcGlobe и в ArcScene. Режим доступа: <http://desktop.arcgis.com/ru/arcmap/10.3/main/get-started/choosing-the-3d-display-environment.htm>.
10. ПО для проектирования в 3D. Режим доступа: <https://www.sketchup.com/ru>
11. Обработка двумерной пространственной информации в составе 3D модели промышленного объекта / Павлов С.В. и др. // Нефтегазовое дело. Т. 13. 2015. № 1. С. 152–158.
12. Христовуло О. И., Павлов С. В., Соколова А. В. Информационная поддержка принятия решений по предупреждению и ликвидации чрезвычайных ситуаций на промышленных объектах на основе технологий трёхмерного геоинформационного моделирования // Auditorium. Электронный научный журнал Курского государственного университета. 2019. № 1(21). С. 24–34.

13. Сакова Н. В., Скрынник А. А. Моделирование химической аварии на предприятии г. Рыбинска // Вестник ыбинской государственной авиационной технологической академии им. П.А. Соловьева. 2015. № 2(33). С. 145–149.
14. РД 52.04.253-90. Методика прогнозирования масштабов заражения сильнодействующими ядовитыми веществами при авариях (разрушениях) на химически опасных объектах и транспорте. М.: Госгидромет СССР, 1991. 24 с.
15. Грехова Ю. С. Система моделирования масштабов зон заражения АХОВ при аварии на ХОО с применением ГИС-технологий // «Научное сообщество студентов. Междисциплинарные исследования»: Электронный сборник статей по материалам XLVI студенческой международной научно-практической конференции. Новосибирск: Изд. АНС «СибАК». 2018. № 11(46) / [Электронный ресурс] Режим доступа. URL: [http://www.sibac.info/archive/science/11\(46\).pdf](http://www.sibac.info/archive/science/11(46).pdf).
16. QGIS – Свободная географическая информационная система с открытым кодом. Режим доступа: <https://qgis.org/ru/site/>.
17. Working with QGIS 3D – Part 1. Режим доступа: <https://www.lutraconsulting.co.uk/blog/2018/03/01/working-with-qgis-3d-part-1/>.
18. Поваренная книга разработчика PyQGIS. Режим доступа: https://docs.qgis.org/3.4/ru/docs/pyqgis_developer_cookbook/index.html.
19. Общее описание ASTER GDEM. – Режим доступа: <http://gis-lab.info/qa/aster-gdem.html>.
20. О проекте Openstreetmap.ru. Режим доступа: <http://openstreetmap.ru/about/org>.
21. Global Mapper – All-in-one GIS Software. Режим доступа: <https://www.bluemarblegeo.com/products/global-mapper.php>.
22. GeoTIFF Format Specification. Режим доступа: <http://geotiff.maptools.org/spec/geotiffhome.html>.
23. Маслобоев А. В. Информационное измерение региональной безопасности в Арктике / А. В. Маслобоев, В. А. Путилов. Апатиты: КНЦ РАН, 2016. 222 с.

Сведения об авторах

Шестаков Алексей Владимирович

стажёр-исследователь
e-mail: shestakov@iimm.ru

Alexey V. Shestakov
research assistant

Шемякин Алексей Сергеевич

младший научный сотрудник
e-mail: shemyakin@iimm.ru

Alexey S. Shemyakin
junior researcher

Яковлев Сергей Юрьевич

к.т.н., с.н.с., доцент
e-mail: yakovlev@iimm.ru

Sergey Yu. Yakovlev
Ph. D., senior researcher, associate professor

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

¹ *Институт информатики и математического моделирования ФИЦ КНЦ РАН*

² *Филиал ФГБОУ ВО МАГУ в г. Апатиты*

РАЗРАБОТКА ФОРМАЛЬНЫХ МОДЕЛЕЙ И СРЕДСТВ СЕТЕЦЕНТРИЧЕСКОГО УПРАВЛЕНИЯ КАДРОВОЙ БЕЗОПАСНОСТЬЮ РЕГИОНА*

Аннотация

Статья посвящена развитию методов и средств сетецентрического управления региональными социально-экономическими системами. Для организации информационно-аналитической поддержки принятия решений в сфере кадровой безопасности региона авторы предлагают использовать комбинацию трех подходов к управлению: функционально-целевого, процессного и сетецентрического. В статье дается описание особенностей взаимного применения указанных подходов к решению вопросов кадрового обеспечения. Приводится модельное представление архитектуры мультиагентной системы поддержки сетецентрического управления региональной кадровой безопасностью. Авторы детально описывают общий алгоритм функционирования разрабатываемой программной системы, основывающийся на цикле PDCA из проектного менеджмента.

Ключевые слова:

Функционально-целевой подход, процессный подход, сетецентрическое управление, мультиагентная система, кадровая безопасность, социально-экономическая система.

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

DEVELOPMENT OF FORMAL MODELS AND MEANS FOR NETWORK-CENTRIC MANAGEMENT OF PERSONNEL SECURITY IN THE REGION

Abstract

The article is devoted to the development of methods and means of network-centric management of regional socio-economic systems. For the organization of information and analytical support of decision-making in the field of personnel security of the region, the authors propose to use a combination of three approaches to management: functional-target, process and network-centric. The article describes the features of the mutual application of these approaches to solving issues of staffing. A model representation of the architecture of a multi-agent support system for network-centric management of regional personnel security is presented. The authors describe in detail the General algorithm of functioning of the developed software system, based on the PDCA cycle of project management.

Keywords:

Functional-targeted approach, process approach, network-centric management, multi-agent system, personnel security, socio-economic system.

Введение

Создание методов и средств управления региональными социально-экономическими системами является исторически сложившимся и не перестающим быть актуальным направлением научно-практических

* Работа выполнена при поддержке Министерства науки и высшего образования РФ и частично Российского фонда фундаментальных исследований - проект № 19-07-01193 А

исследований. В рамках данной тематики одним из перспективных направлений остается развитие средств информационно-аналитической поддержки управления. Разработка подобных инструментов поддержки принятия решения связана, в первую очередь, с моделированием предметной области как способом представления объекта исследования в формальном виде.

Моделирование социально-экономических систем представляет собой достаточно трудоемкий процесс, связанный с наличием большого количества разнородных элементов и отношений между ними. Оперирование такого типа системами ограничено ментальными возможностями человека. Одним из возможных способов решения этой проблемы является представление сложных систем в виде иерархических древовидных структур.

Способ построения иерархических древовидных структур, который развивается в Институте информатики и математического моделирования КНЦ РАН исследователями научной школы под руководством Путилова В.А., получил название функционально-целевой подход [1]. Основная идея данного подхода заключается в формировании иерархической системы целей, а также последующее построение дерева функций, выполнение которых приводит к достижению соответствующих целей.

В данной статье приводятся промежуточные результаты реализации научно-исследовательских проектов «Модели и методы конфигурирования адаптивных многоуровневых сетцентрических систем управления региональной безопасностью в Арктической зоне Российской Федерации» и «Методы и средства информационной поддержки управления кадровой безопасностью регионального горно-химического кластера». Общей задачей для данных проектов является разработка модели сетцентрического управления разнородными элементами региональных социально-экономических систем. Одним из способов решения данной задачи, предлагаемым авторским коллективом, является использование комбинации уже зарекомендовавших себя подходов к управлению сложными системами, а именно: функционально-целевого, процессного и сетцентрического. В рамках данной статьи более подробно рассматриваются особенности совместного применения указанных подходов, в результате которого формируется многоуровневая рекуррентная модель иерархического управления региональной кадровой безопасностью.

В рамках текущего исследования в качестве предметной области рассматривается кадровая безопасность региона. Региональная кадровая безопасность обеспечивается совокупностью взаимосвязанных социально-экономических систем региона, таких как региональный рынок труда, региональная система подготовки и переподготовки кадров, региональная система потребления трудовых ресурсов и другие. Выбор данной предметной области не случаен, так как в последние годы обостряется проблема удовлетворения кадровых потребностей экономики северных и арктических регионов РФ. Решение указанной проблемы заставляет разрабатывать новые подходы к управлению кадровой политикой, в частности, создавать новые информационные технологии поддержки управления кадровым обеспечением региональной экономики (в том числе и на основе сетцентрической парадигмы управления сложными системами).

Применение функционально-целевого подхода к кадровой безопасности

Прежде чем подходить к формализованному описанию предметной области исследования, необходимо определиться с терминологией, характерной для функционально-целевого подхода, так как могут возникать расхождения в интерпретации тех или иных терминов. В рамках текущей работы под целью системы будем понимать желаемое состояние системы или желаемый результат ее поведения. Процесс достижения цели системой неразрывно связан с понятием «задача». При этом задача семантически интерпретируется как проблемная ситуация с явно заданной целью, которую необходимо достичь. Другим термином, активно применяемым в функционально-целевом подходе, является функция, которая воспринимается как действие системы, направленное на решение соответствующей задачи. Для установления соответствия конкретной функции определенной цели используется базовая операция - «покрытие цели действием» [1,2].

В конечном итоге, используя принцип целеполагания и операцию покрытия, формируются две древовидные иерархические структуры, отражающие декомпозицию целей и покрывающих их действий (функций), при этом начальная вершина дерева целей является глобальной целью.

При построении дерева целей любой сложной системы можно ставить разные глобальные цели. В соответствии с этим утверждением определим множество глобальных целей $\{G_i^0, i = \overline{1, n}\}$. В рамках рассмотрения такой предметной области, как кадровая безопасность, можно предложить следующие формулировки глобальных целей:

- обеспечить регион всеми необходимыми трудовыми ресурсами;
- трудоустроить все экономически активное население региона;
- трудоустроить всех выпускников региональной системы профессионального образования;
- оптимизировать потоки трудовой миграции и др.

Согласно функционально-целевому подходу для построения дерева целей предметной области проводится процедура декомпозиции цели. Декомпозиция заключается в рекуррентном разбиении текущей цели на множество подцелей, совместное достижение которых обеспечивает и достижение текущей цели. Данный процесс проводится до тех пор, пока не получится множество неделимых целей (примитивов). Условие неделимости цели является возможность достижения этой цели с помощью элементарного действия (функции).

В результате декомпозиции получается дерево целей, которое можно описать в виде графового представления:

$Tr^G = \langle V^G, L^G \rangle$ – дерево целей.

Здесь $V^G = \{V^{Gk}\}$ – множество вершин дерева, представляющих собой цели, где k – номер уровня иерархии, $k = 0, \dots, N$, N – количество уровней иерархии,

$V^{Gk} = \{v_i^{Gk}\}$ – множество вершин k -го уровня иерархии, где i – номер цели, $i = 1, \dots, m_k$,

$L^G = \{l_{ij}^G\}$ – множество связей между целями, $i, j = 0, \dots, m_k$,

$l_{ij}^G = \langle v_i^{Gk}, v_j^{Gk+1} \rangle$ – связь между i -ой целью k -го уровня и j -ой целью $(k+1)$ -го уровня.

Стоит отметить, что для каждой глобальной цели строится свое дерево.

Для представления процедуры разбиения цели на подцели введем понятие рекурсии на дереве целей:

$$V_i^{G_{k+1}} = F_r(v_i^{G_k})$$

$$F_r: V^{G_k} \rightarrow V^{G_{k+1}}$$

$\forall v_i^{G_k} \in V^{G_k} \exists V_i^{G_{k+1}} \subset V^{G_{k+1}}: \forall v_j^{G_{k+1}} \in V_i^{G_{k+1}} \exists l^G = \langle v_i^{G_k}, v_j^{G_{k+1}} \rangle$, при этом возможен случай, когда: $V_i^{G_{k+1}} = \emptyset$.

В качестве примера построения дерева целей можно рассмотреть процесс декомпозиции глобальной цели «Обеспечить регион всеми необходимыми кадрами». На первом шаге итерации производится разбиение данной глобальной цели на следующие подцели, тем самым получаем первый уровень иерархии (k=1) древовидной структуры:

1. Выявить дисбаланс в кадрах.
2. Организовать систему привлечения кадров.
3. Разработать механизмы трудоустройства.

Следуя рекуррентной процедуре разбиения целей i-го уровня иерархии на подцели, формируется i+1 уровень иерархии целей. В рассматриваемом примере для цели 1-го уровня иерархии «Организовать систему привлечения кадров» получаем следующие подцели 2-го уровня:

1. Организовать привлечение кадров извне.
2. Организовать подготовку кадров внутри региона.

Проводим декомпозицию до тех пор, пока не получим неделимые цели – примитивы. Фрагмент дерева целей (3 уровня иерархии) для глобальной цели «Обеспечить регион всеми необходимыми кадрами» представлен на рисунке 1.

В соответствии с концепцией функционально-целевого подхода дереву целей ставится в соответствие дерево функций (действий). Для этой процедуры вводится операция покрытие целей действиями, которая детально описана в работе [1]. Таким образом, применяя операцию покрытия, получаем из дерева целей соответствующее дерево функций. Данное дерево действий можно представить в виде следующего формального описания:

$Tr^A = \langle V^A, L^A \rangle$ — дерево функций (действий).

Здесь $V^A = \{V^{A_k}\}$ — множество вершин дерева, представляющих собой действия, где k — номер уровня иерархии, $k = 0, \dots, N$, N — количество уровней иерархии,

$V^{A_k} = \{v_i^{A_k}\}$ — множество вершин k-го уровня иерархии, где i — номер действия, $i = 1, \dots, m_k$, при этом V^{A_k} покрывает соответствующий уровень V^{G_k} дерева целей.

$L^A = \{l_{ij}^A\}$ — множество связей между действиями, $i, j = 0, \dots, m_k$,

$l_{ij}^A = \langle v_i^{A_k}, v_j^{A_{k+1}} \rangle$ — связь между i-ым действием k-го уровня и j-м действием (k+1)-го уровня.

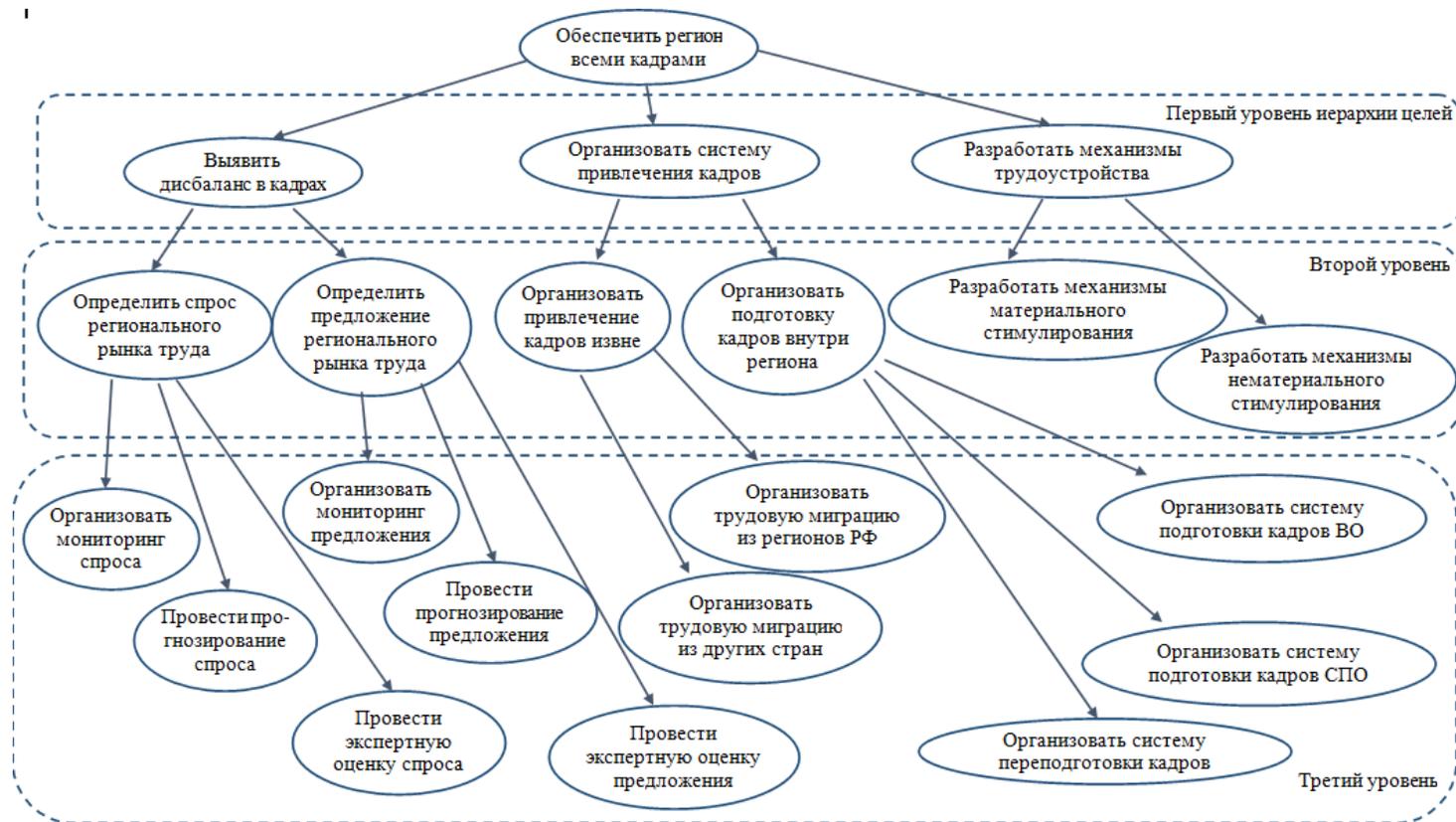


Рис. 1. Фрагмент декомпозиции глобальной цели «Обеспечить регион всеми необходимыми кадрами»

Дерево функций, в отличие от дерева целей, строится в обратном порядке – снизу-вверх. Каждому примитиву дерева целей ставится в соответствие действие, позволяющее достигнуть эту цель (рис.2). Таким образом, формируется множество примитивов дерева функций. Вершины i -го уровня формируются как совокупность действий $i+1$ -го уровня.

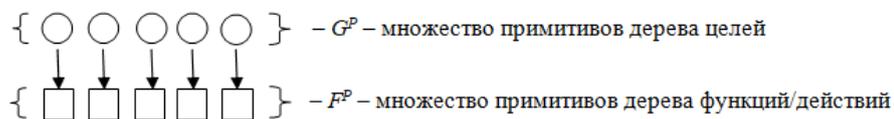


Рис.2 Отображение множества примитивов дерева целей на множество примитивов дерева функций/действий

Получаемое в результате операции покрытия дерево функций ставится в соответствие дереву целей.

Применение процессного подхода к кадровой безопасности

В предлагаемой модели управления каждый примитив дерева функций (F_i^p) можно описать соответствующим процессом или цепочкой процессов. Согласно международному стандарту ISO 9001:2000, процесс — это совокупность взаимосвязанных или взаимодействующих видов деятельности, преобразующая входы в выходы [3]. При этом на процесс действуют различные регуляторы, а также имеются исполнители $\{Ex_i\}$ рассматриваемого процесса. Упрощенная схема процесса Pr_k представлена на рисунке 3.



Рис. 3 Общая схема процесса (в соответствии с нотацией IDEF[4])

При описании примитивов дерева функций необходимо принимать во внимание, что процессы могут протекать как последовательно, так и параллельно (рис. 4). При этом возможны различные способы организации хода реализации действий, что соответствует утверждению о том, что каждый примитив дерева функций может быть представлен несколькими альтернативными цепочками процессов.

Например, рассматривая примитив «Организовать переподготовку на базе предприятий и специализированных центров переобучения» дерева функций, мы можем представить его в виде различных цепочек процессов:

1. «Определение количества лиц для переобучения» → «Формирование образовательной программы переобучения» & «Формирование преподавательского состава» & «Формирование материально-технической базы» → «Формирование группы обучающихся» →

- «Реализация программы обучения» & «Ведение сопроводительной отчетной документации» → «Проведение квалификационного испытания» → «Оформление квалификационных документов»;
2. «Определение количества лиц для переобучения» → «Формирование образовательной программы переобучения» → «Формирование преподавательского состава» → «Формирование материально-технической базы» → «Формирование группы обучающихся» → «Реализация программы обучения» → «Проведение квалификационного испытания» → «Присвоение квалификации»
 3. другие комбинации процессов.

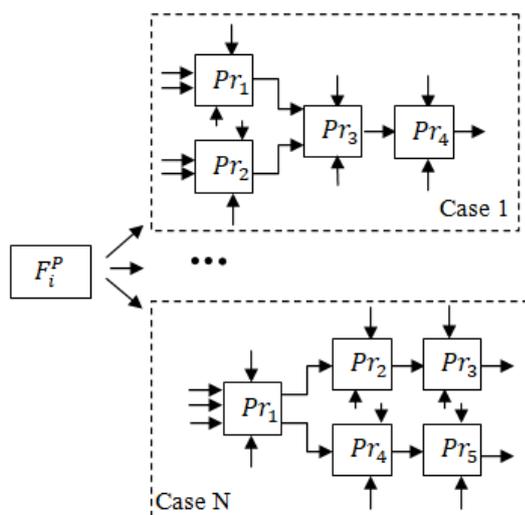


Рис. 4. Представление примитива дерева функций (F_i^P) в виде альтернативных цепочек процессов

В рассмотренном примере символ «→» стоит интерпретировать как последовательное выполнение действия, а символ «&» - одновременное/параллельное протекание процессов.

При выборе альтернативной цепочки процессов, которую нужно использовать для формирования вершин дерева функций/действий вышестоящего уровня иерархии, ставится задача индикаторной оценки альтернатив. Для получения объективной оценки цепочки процессов предлагается применить систему индикаторов, включающую в себя несколько категорий показателей процесса (финансы, время, экономический эффект и др.). Более подробно процесс индикаторной оценки рассмотрен ниже в разделе «Общая схема работы разрабатываемых средств сетецентрического управления кадровой безопасностью».

Применение сетецентрического подхода к кадровой безопасности

В последние годы активно развивается теория сетецентрического управления разнообразными социально-экономическими системами. Сотрудниками Института информатики и математического моделирования КНЦ

РАН Маслобоевым А.В. и Путиловым В.А. предложены ряд концептуальных и методических разработок в указанной научной сфере. В частности, они разработали многоуровневую рекуррентную модель иерархического управления комплексной безопасностью региональных социально-экономических систем. По мнению указанных исследователей разработанная ими модель «обеспечивает как формальную постановку и решение, так и практическую реализацию задач синтеза структуры автоматизированных систем и алгоритмов управления региональной безопасностью, оптимальных в смысле определенных критериев. Предложен подход к решению задач внутриуровневой и межуровневой координации в многоуровневых иерархических системах. Такая координация обеспечивается за счет удовлетворения требований взаимосвязи между показателями качества функционирования (целевыми функциями), оптимизируемыми различными элементами многоуровневых систем. Это позволяет достичь достаточной согласованности локальных решений, принимаемых на разных уровнях управления, в условиях децентрализованного принятия решений и высокой динамики внешней среды. Использование рекуррентной модели позволяет сформировать математические модели управления безопасностью региональных социально-экономических систем, функционирующих в условиях неопределенности.»[5]

Авторы данной статьи продолжают развитие концепции создания систем сетцентрического управления, предложенной Маслобоевым А. В. и Путиловым В. А., за счет интеграции функционально-целевого и иерархического подходов к управлению региональными социально-экономическими системами.

На текущем этапе проводимого исследования главным допущением является то, что все участники процессов обеспечения кадровой безопасности обладают достаточным уровнем заинтересованности и мотивации, чтобы решать поставленные перед ними задачи. В противном случае, не имеет смысла говорить о создании эффективной системы управления.

На рисунке 5 приведена концептуальная схема модели управления кадровой безопасностью региона. Она показывает основные уровни модели управления, на которых реализуются отдельные стадии принятия решений. Модель включает в себя три основных уровня: концептуальный, виртуальный и организационный.

Организационный уровень представляет собой совокупность организационных структур, принимающих прямое или косвенное участие в процессах, связанных с планированием, реализацией и контролем действий, направленных на достижение поставленной глобальной цели. Атомарными элементами организационного уровня являются представители организаций и предприятий, принимающих участие в эксплуатации разрабатываемых программных средств поддержки сетцентрического управления кадровой безопасностью региона. В рамках предлагаемой структуры модели управления взаимодействие осуществляется между элементами организационного и виртуального уровней и проявляется в виде взаимодействия представителя организационной структуры с пользовательским интерфейсом соответствующего программного агента.

Виртуальная среда представляет собой сетцентрическую структуру, которой свойственны следующие характеристики:

1. большое количество разнородных агентов и источников информации;

2. сетевая архитектура с выделением трех типов программных агентов;
3. организация распределенного хранения информации;
4. оперирование большими объемами информации;
5. центр принятия решения децентрализован и может меняться (мигрировать по сети) в зависимости от решаемой задачи;
6. наличие механизмов согласования целей и координации действий и/или управления между узлами мультиагентной сети.

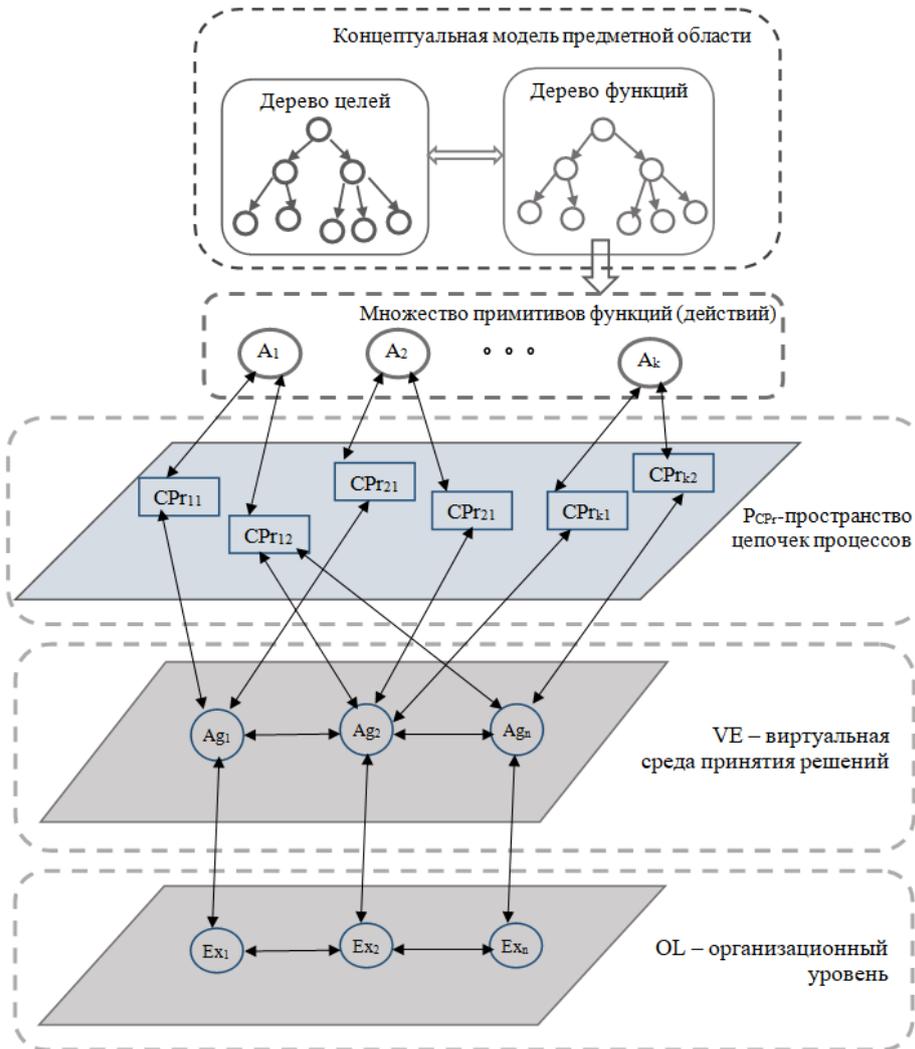


Рис. 5 Концептуальная схема модели управления кадровой безопасностью

Виртуальный уровень предлагаемой модели управления кадровой безопасностью осуществляет взаимодействие как с объектами организационного, так и концептуального уровней.

Концептуальный уровень модели формируют формализованные концептуальные модельные описания, включающие в себя графовое представление деревьев целей и функций (в соответствии с функционально-

целевым подходом), ранее разработанные концептуальные модели кадровой логистики региона [6] и управления региональной безопасностью на основе проектного менеджмента [7]. С точки зрения практической реализации предлагаемой модели в виде программного продукта формализованные концептуальные описания будут представлены в виде прикладных онтологий формата OWL.

Взаимодействие между концептуальным и виртуальным уровнями сводится к назначению программных агентов в роли исполнителей определенных процессов из сформированного пространства цепочек действий (см. рис. 5). При этом в качестве исполнителей процесса могут выступать различные сущности: ответственные за реализацию процесса; стейкхолдеры; непосредственно исполнители процесса и др. В результате каждому исполнителю Ex_i ставится в соответствие агент Ag_i виртуальной среды принятия решений.

Общая схема работы разрабатываемых средств сетецентрического управления кадровой безопасностью

В рамках проводимого исследования планируется реализовать выше описанные подходы к сетецентрическому управлению кадровой безопасностью региона в виде прототипа системы поддержки принятия решения (СППР). Общий алгоритм функционирования данного прототипа представляет собой последовательность из четырнадцати ключевых этапов. Основным назначением разрабатываемых средств сетецентрического управления кадровой безопасностью является организация информационно-аналитического сопровождения процесса решения проблемы в сфере кадровой безопасности. Предлагается рассматривать проблему с точки зрения проектного подхода, представляя решаемую задачу в виде глобального проекта [7]. Принимая данный факт во внимание, было принято решение построить общий алгоритм функционирования прототипа СППР на основе цикла Дейминга [8], применяемого в проектном менеджменте.

Первые восемь этапов предлагаемого алгоритма относятся к планированию глобального проекта, последующие два – к выполнению проекта, одиннадцатый этап – к проверке выполнения проекта, а остальные – к выработке корректирующих воздействий. Рассмотрим данные этапы более подробно:

Первый этап - инициация проблемы, требующей решения. На данном этапе один из участников организационной среды выдвигает проблему социально-экономического развития региона посредством обращения к виртуальной среде через своего агента-представителя. Возможны ситуации, когда проблема выдвигается коалицией агентов. Например, проблема о необходимости обеспечить трудовыми ресурсами экономику региона выносится на рассмотрение одним участником организационной среды, в частности, Министерством экономического развития Мурманской области, а проблема подготовки определенного количества квалифицированных работников заданного профиля может выдвигаться целой коалицией предприятий горно-химического кластера.

Второй этап – формулирование глобальной цели иницированной проблемы. Предполагается, что результатом данного этапа является сформулированная глобальная цель, при формировании которой принимают

участие, как агент-инициатор, так и другие агенты виртуальной среды, в зону ответственности которых входит решение вопросов, связанных с рассматриваемой проблемой.

Могут быть случаи, когда инициированной проблеме соответствует несколько формулировок глобальных целей, имеющих разную семантическую окраску. Например, проблема «Наличие кадрового голода в Мурманской области». Возможные формулировки глобальной цели для указанной проблемы: 1) «сбалансировать спрос и предложение регионального рынка труда»; 2) «организовать вахтовый способ привлечения кадров» и т.д.

Третий этап – формирование дерева целей, соответствующего выбранной глобальной цели. В рамках исследования процедура декомпозиции глобальной цели проводится под управлением агента виртуальной среды, который инициировал проблему. Данный агент становится координатором процесса разбиения глобальной цели на подцели. Разбиение осуществляется посредством взаимодействия коалиции агентов ($CoAg$), являющейся подмножеством всех агентов Ag виртуальной среды.

$$CoAg = \{Ag_j\}, j = \overline{1, \dots, \bar{H}}$$

$CoAg \subseteq Ag, Ag = \{Ag_i\}, i = \overline{1, \dots, \bar{N}}, H \leq N$, где H, N – количество агентов в коалиции и виртуальной среде соответственно.

В рамках коалиции выделяются две роли агентов: координатор и эксперт. Агенты-эксперты ранжируются по степени осведомленности (компетентности) о предметной области. В результате совместной работы агента-координатора с агентами-экспертами на выходе данного этапа формируется дерево целей.

Четвертый этап – формирование дерева функций. Процесс построения дерева функций в отличие от предыдущего этапа проводится в обратном направлении («снизу вверх»). Т.е. вершина верхнего уровня дерева действий получается за счет совместного выполнения функций нижестоящего уровня иерархии. Данная процедура заканчивается, когда будет достигнута вершина, соответствующая глобальной цели. В общем случае множество агентов-экспертов, привлекаемых на данном этапе, может отличаться от набора экспертов, участвующих в реализации предыдущего шага. Результатом выполнения вышеописанных этапов являются две древовидные структуры — дерево целей и соответствующее ему дерево функций.

Пятый этап – формирование цепочек процессов. Коалиция агентов для каждого примитива дерева функций формирует либо одну цепочку процессов, либо множество альтернативных цепочек (см. рис. 4).

Шестой этап – выбор множества процессов. Основной задачей данного этапа является ограничение множества процессов, которые будут рассматриваться в рамках решаемой задачи, за счет отсечения альтернативных цепочек процессов в соответствии с заданными критериями. Критерии отбора устанавливаются агентом-координатором и могут включать в себя финансовые, временные, материальные и другие виды индикаторов. В качестве интегральной оценки альтернатив цепочек процессов можно использовать взвешенную сумму индикаторов по каждой категории показателей.

$Ind_{CP} = \sum_{i=1}^N \sum_{j=1}^N (w_{i,j} \cdot Ind_{i,j})$, где i — номер категории показателей, j — номер индикатора в i -ой категории, $w_{i,j}$ — весовой коэффициент конкретного показателя, $Ind_{i,j}$ — значение конкретного показателя.

В результате на шестом этапе формируется множество актуальных цепочек процессов (CPr), каждая из которых однозначно соответствует элементу множества примитивов дерева функций.

$$CPr = \{CPr_i\} \rightarrow \{F_i^P\}, i = \overline{1, \dots, N}, \text{ где } N \text{ — количество примитивов.}$$

Седьмой этап – синтез сети процессов решаемой задачи. В рамках данного этапа решаются следующие основные задачи:

1. Определение взаимосвязей между процессами и цепочками процессов в соответствии с деревом функций.
2. Определение порядка (последовательности) выполнения цепочек действий.
3. Согласование временных характеристик выполнения процессов.
4. Оценка результатов планирования сети актуальных процессов.

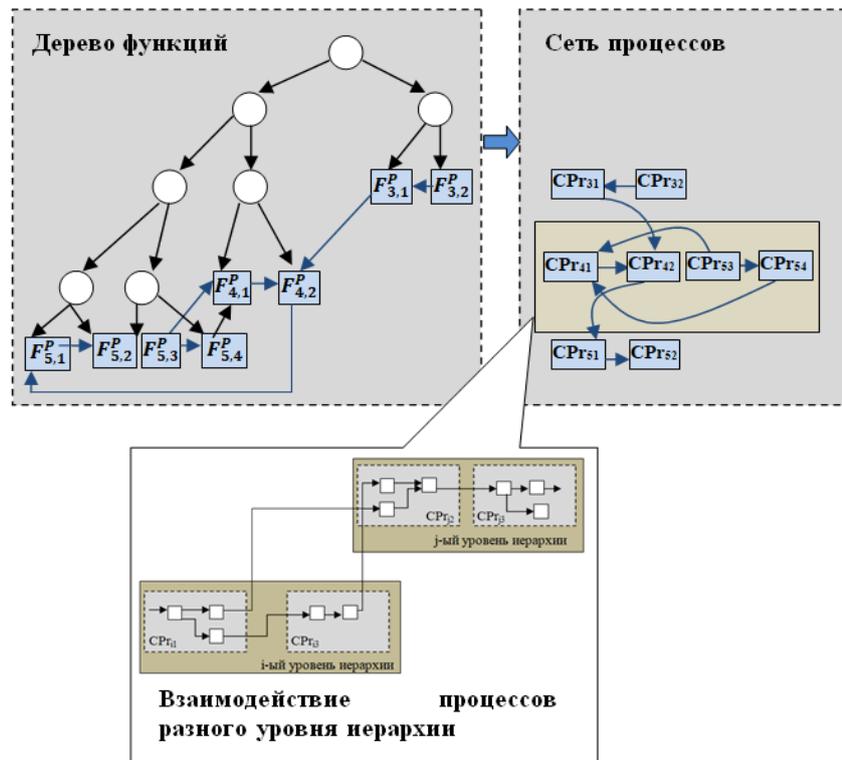


Рис. 7 Формирование сети актуальных процессов

Данный этап сводится к проверке полноты покрытия цепочками процессов всех вершин для каждого уровня иерархии дерева функций. В случае выявления конфликтных ситуаций при установлении взаимосвязей между процессами или их временными характеристиками возникает необходимость внесения изменений в ранее принятые решения. Корректировка может производиться практически на всех вышеописанных этапах, а именно: уточнение (детализация) глобальной цели, перестроение дерева целей, изменение дерева функций (в частности, множества примитивов), формирование новых или перегруппировка уже сформированных цепочек процессов и т.д. В зависимости

от конкретного случая может осуществляться возврат на тот или иной этап общей схемы работы разрабатываемых средств сетецентрического управления кадровой безопасностью. При этом на каждое корректирующее воздействие накладывается ограничение по количеству допустимых итераций, которое нужно выполнить для получения удовлетворительного результата. По истечению отведенного количества итераций на исправления, в случае отрицательного результата проведенной коррекции, осуществляется переход на более ранний этап общей схемы работы. Например, если при попытке перегруппировать уже сформированные цепочки процессов проблема не исчезает, то осуществляется переход на этап формирования альтернативных цепочек процессов (пятый этап). На выходе данного этапа формируется согласованная сеть актуальных процессов (рис. 7).

На этапах 5-7 коалиции агентов, привлекаемых к данной задаче, могут изменяться в связи с появлением новых экспертов с необходимыми компетенциями, а также исключением из экспертной группы членов с недостаточным уровнем знаний, умений и навыков.

Восьмой этап — выбор исполнителей. Для каждого примитива дерева функций, который на данном этапе уже представляет собой цепочку процессов, формируется отдельная рабочая группа из агентов виртуальной среды. Один из агентов каждой рабочей группы наделяется полномочиями координатора. Механизмы определения координатора могут быть различными: например, самовыдвижение некоторого агента, делегирование полномочий вышестоящим управляющим агентом или выдвижение в результате согласованного решения всей рабочей группы. Возможен случай, когда инициатор проблемы сам становится координатором работы отдельной рабочей группы. Алгоритм назначения исполнителя отдельного процесса представляется в виде следующего набора действий:

1. Согласованный выбор исполнителей из множества агентов, проявивших инициативу на роль исполнителя процесса. Самовыдвижение со стороны участников рабочей группы соответствует принципам сетецентрического управления.

2. Совместное решение рабочей группы по «незакрытым позициям» в сети процессов относительно назначения исполнителей.

3. В случае отсутствия полного покрытия множества процессов исполнителями после первых двух действий исполнители процессов назначаются координатором этой рабочей группы.

Ключевым параметром при назначении исполнителей конкретного процесса выступает зона его компетентности, которая является набором его знаний, навыков и умений решения задач из определенной предметной области. В каждом из рассмотренных действий алгоритма назначения исполнителя предпочтение отдается более компетентному агенту. Определение зоны компетентности является отдельной научной задачей, которая может быть сведена к интегральной оценке некоторого множества параметров агентов, например, такие как: опыт реализации проектов, предметная область, зона ответственности и т.д.

При определении исполнителей конкретного процесса может возникнуть ситуация, когда число исполнителей данного процесса больше одного. В этом случае рационально провести детализацию рассматриваемого процесса на

множество подпроцессов таким образом, чтобы у каждого подпроцесса был один исполнитель. На текущей стадии исследования вводится ограничение: каждый исполнитель процесса является ответственным лицом за результаты реализации закрепленного за ним процесса, что с точки зрения проектного менеджмента не всегда является таковым.

Девятый этап – инициализация глобального проекта, который заключается в фиксации конкретных исполнителей (*ОЕ*), сроков выполнения проекта, описания финансовых аспектов и др. Все перечисленное оформляется в регламентирующую документацию и принимается всеми участниками глобального проекта.

Десятый этап — запуск глобального проекта. Все участники глобального проекта в соответствии с их ролями, зонами ответственности и календарными планами, зафиксированными в проектной документации, приступают к выполнению своих обязанностей.

Одиннадцатый этап — контроль реализации глобального проекта. Координаторы каждой рабочей группы собирают информацию о показателях, характеризующих ход выполнения процессов, которые находятся в зоне их ответственности.

Двенадцатый этап — выработка коррекционных воздействий. В рамках предлагаемой модели управления предполагается активное использование прогнозирования для анализа возможных путей реализации запланированного глобального проекта. Прогнозирование будет реализовано с помощью имитационного моделирования, позволяющего достаточно оперативно получать прогнозы в установленных сценарных условиях, как об отдельных процессах реализации глобального проекта, так и обо всем проекте в целом. Принимая во внимание прогнозную информацию и данные мониторинга, коалиция агентов, составляющая рабочую группу по реализации цепочки процессов, принимает коллективное решение относительно внесения изменений в план реализации глобального проекта.

Тринадцатый этап — доведение корректирующих воздействий до исполнителей. Данный этап предполагает оповещение всех заинтересованных исполнителей, деятельность которых затрагивает выработанное корректирующее воздействие, посредством виртуальной среды.

Стоит отметить, что с девятого по тринадцатый этапы реализации сформированного плана выполнения глобального проекта осуществляется на организационном уровне модели управления кадровой безопасности региона. Т.е. представители организаций посредством взаимодействия с программными агентами виртуальной среды передают информацию о результатах (промежуточных и финальных) и ходе реализации процессов, закрепленных за ними, а также принимают участие в выработке корректирующих воздействий.

Четырнадцатый этап — анализ степени достижения глобальной цели по результатам выполнения отдельных стадий проекта. При анализе промежуточных результатов выполнения глобального проекта может сложиться ситуация, когда корректирующие воздействия не приводят к достижению первоначальной глобальной цели. В этом случае ставится вопрос о принципиальном продолжении реализации глобально проекта, но с худшими результатами, либо о прекращении глобального проекта и возврат на первый этап общего алгоритма.

Приведенный в данном разделе общий алгоритм функционирования СППР в области кадровой безопасности отражает концепцию комбинированного использования функционально-целевого, процессного, проектного и сетецентрического подходов к управлению социально-экономическими системами.

Заключение

На сегодняшний день создание методических и программных средств информационно-аналитической поддержки управления социально-экономическими системами является востребованным научно-техническим направлением. В рамках развития данного направления предлагается использовать комбинацию известных методов и подходов для управления кадровой безопасностью региона. По мнению авторов, интеграция разных подходов к управлению сложными системами может дать новый синергетический эффект, затрагивающий вопросы оперативности, качества и обоснованности принятия решений в области кадровой политики.

В то же время стоит отметить, что в статье изложены концептуальные моменты предлагаемых решений в области управления кадровой безопасностью. Ряд затронутых в работе вопросов остается открытым для дальнейших научных изысканий. В частности, разработка механизмов согласования целей разных участников проекта и разрешения конфликтных ситуаций в ходе его реализации.

Литература

1. Кузьмин И. А. Распределенная обработка информации в научных исследованиях / И. А. Кузьмин, В. А. Путилов, В. В. Фильчаков. Л.: Наука, 1991 304 с.
2. Путилов В. А., Горохов А. В. Системная динамика регионального развития / В. А. Путилов, А. В. Горохов. Мурманск: НИЦ «Пазори», 2002. 306 с.
3. Международный стандарт ISO 9001:2000 // Режим доступа: <http://niits.ru/public/2003/069.pdf> (22.11.2019).
4. Цуканова О. А. Методология и инструментарий моделирования бизнес-процессов: учебное пособие. СПб.: Университет ИТМО, 2015. 100 с. Режим доступа: <https://books.ifmo.ru/file/pdf/1720.pdf> (22.11.2019).
5. Маслобоев А. В. Многоуровневая рекуррентная модель иерархического управления комплексной безопасностью региона / А. В. Маслобоев, В. А. Путилов, А. В. Сютин // Научно-технический вестник информационных технологий, механики и оптики, 2014. № 6(94). С. 163–170.
6. Малыгина С. Н. Логистика кадрового обеспечения региона: формализация и структура полимодельного комплекса / Малыгина С. Н., Быстров В. В., Халиуллина Д. Н // Труды Кольского научного центра. Информационные технологии 10/2018 (9). Вып. 9. Апатиты: Изд-во КНЦ РАН. С. 36–47 DOI: 10.25702/KSC.2307-5252.2018.10.36-47.
7. Быстров В. В. Применение проектного менеджмента в задачах управления региональной безопасностью: подход и формальный аппарат / В. В. Быстров, А. В. Маслобоев, В. А. Путилов // Надежность и качество сложных систем. 2017. № 4. С. 73–84.

8. Репин В. В., Елиферов В. Г. Процессный подход к управлению. Моделирование бизнес-процессов. М.: РИА «Стандарты и качество». 2008. 408 с.

Сведения об авторах

Быстров Виталий Викторович

к.т.н., старший научный сотрудник

e-mail: bystrov@iimm.ru

Vitaliy V. Bystrov

Candidate of Tech. Sciences, senior researcher

Халиуллина Дарья Николаевна

к.т.н., научный сотрудник

e-mail: khaliullina@iimm.ru

Darya N. Khaliullina

Candidate of Tech. Sciences, researcher

Малыгина Светлана Николаевна

к.т.н., научный сотрудник

e-mail: malygina@iimm.ru

Svetlana N. Malygina

Candidate of Tech. Sciences, researcher

DOI: 10.25702/KSC.2307-5252.2019.9.84-90

УДК 004.853

М. Л. Малоземова, П. А. Ломов

Институт информатики и математического моделирования ФИЦ КНЦ РАН

**ТЕХНОЛОГИЯ ПОПОЛНЕНИЯ ТЕЗАУРУСА НА ОСНОВЕ ПРИМЕНЕНИЯ
ДИСТРИБУТИВНОГО ПОДХОДА К АНАЛИЗУ ЕСТЕСТВЕННО-ЯЗЫКОВЫХ
ТЕКСТОВ**

Аннотация

В статье рассматривается применение дистрибутивного подхода к анализу набора естественно-языковых текстов с целью выявления в нем лексических единиц для пополнения существующего тезауруса. Разработана технология пополнения тезауруса, представленного в проекте национального стандарта «Исследования в полярных регионах. Основные термины и определения» на основе латентно-семантического анализа.

Ключевые слова:

дистрибутивный подход, пополнение тезауруса, латентно-семантический анализ (ЛСА).

M. L. Malozemova, P. A. Lomov

**THESAURUS REPLENISHMENT TECHNOLOGY BASED ON THE APPLICATION
OF A DISTRIBUTIVE APPROACH TO ANALYSIS OF NATURAL LANGUAGE TEXTS**

Abstract

The article discusses the application of the distribution approach to the analysis of a set of natural language texts in order to identify its lexical units to replenish the existing thesaurus. The replenishment technology of the thesaurus has been developed, which

is presented in the draft national standard "Research in the polar regions. Basic terms and definitions". The technology is based on latent-semantic analysis.

Keywords:

distributive approach, thesaurus replenishment, latent-semantic analysis (LSA).

Введение

На сегодняшний день применение специализированных тезаурусов является распространенной практикой при работе коллективов и групп специалистов над решением определенного круга задач. Тезаурус позволяет «зафиксировать» единый набор используемых терминов, а также их взаимосвязи и смысловое значение для дальнейшего использования в различного рода коммуникациях, а также при решении широкого круга задач анализа и обработки естественно-языковых текстов. С этой точки зрения, целесообразно обеспечить его формирование и дальнейшее наполнение с наименьшими затратами труда и времени. При этом часто необходимо включать в тезаурус набор специализированных терминов, характерных для области его дальнейшего применения. В связи с этим актуальной видится задача автоматизации расширения некоторого исходного тезауруса на основе анализа набора специализированных текстов, потенциально содержащих новые понятия, относящиеся к рассматриваемой предметной области.

В данной работе предлагается технология, позволяющая выявить возможных кандидатов для пополнения тезауруса, представленного в проекте национального стандарта «Исследования в полярных регионах. Основные термины и определения», на основе применения дистрибутивного подхода.

1. Обзор существующих подходов к анализу текстов

Для решения задачи пополнения тематического тезауруса можно применять ряд подходов к анализу текстов:

1. Дистрибутивные подходы, учитывающие частоту встречаемости слов в едином контексте. Они основываются на дистрибутивной гипотезе, согласно которой лингвистические единицы, встречающиеся в схожих контекстах, имеют близкие значения. Результатом применения таких подходов является представление каждого слова в виде вектора в некотором пространстве.

2. Лингвистические подходы, предполагающие генерацию правил или лексико-синтаксических паттернов, на основе которых будут извлекаться знания из текста (лексико-синтаксические анализаторы). В данных подходах активно используются все уровни анализа естественного языка: морфология, синтаксис и семантика.

3. Краудсорсинг (crowdsourcing), предполагающий привлечение круга лиц со стороны для осуществления анализа текстового корпуса.

В рамках данной работы применялся один из дистрибутивных подходов, поэтому остановимся на этой группе подходов более подробно. Данная группа подходов основывается на технологии, называемой word embeddings (вложения слова) [1]. Вложение слова – это действительное число, векторное представление слова. Как правило, слова со схожим значением будут иметь векторные представления, которые расположены близко друг к другу в пространстве вложения.

Самыми простыми методами векторизации слов являются мешок слов, представление TF-IDF слов документа и матрица смежности (встречаемости) пары слов в документах.

Метод «горячего» кодирования (one-hot encoding) [1] – один из самых простых способов численного представления слов. Иначе этот метод еще называют «мешок слов» (bag-of-words). Создается вектор, который имеет столько измерений, сколько корпус содержит уникальные слова. Каждому уникальному слову присваивается значение, равное количеству его встречаемости в данном документе. На выходе получаются огромные и редкие векторы, которые не собирают абсолютно никакой реляционной информации. Данный способ применяется, когда нет других вариантов векторного представления слов.

TF-IDF векторы (Term Frequency – Inverse Document Frequency) [1] показывают вес слова, который пропорционален частоте употребления этого слова в документе и обратно пропорционален частоте употребления слова во всех документах коллекции. Проще говоря, это метод, который увеличивает веса слов, часто встречающихся в данном документе, и уменьшает веса слов, часто встречающихся во многих документах.

Матрица смежности (co-occurrence matrix) [1] представляет собой гигантскую матрицу, размер которой соизмерим с размером словаря. Она описывает то, как слова встречаются вместе, что, в свою очередь, фиксирует отношения между ними. Матрица смежности слов вычисляется просто путем подсчета того, как два или более слова встречаются вместе в данном корпусе. Недостатком данного метода является достаточно большое векторное представление слова, что составляет в объеме как размерность при методе горячего кодирования, только в квадрате. В итоге требуется слишком много памяти для хранения полученных результатов.

Более сложным методом считается *латентно-семантический анализ (ЛСА) [2]* – один из самых распространенных методов анализа текстовой информации. Он используется для выявления латентных (скрытых) ассоциативно-семантических связей между терминами (словами, n-граммами) путем сокращения факторного пространства термины-на-документы. «Семантическое» пространство формируется на основе алгоритма, и семантика базируется как раз на встречаемости в контексте (дистрибутивная семантика). В области информационного поиска данный подход называют также латентно-семантическим индексированием (ЛСИ).

ЛСА отображает документы и отдельные слова в пространство термины-на-документы, что позволяет представлять их в виде векторов и тем самым решать такие задачи, как, например, сравнение двух термов между собой, сравнение двух документов между собой, сравнение термина и документа и сокращение факторного пространства термины-на-документы для последующей визуализации.

Тематическое моделирование [3] является развитием латентно-семантического анализа, но оно применяется для более узкой задачи определения принадлежности документа к теме. Наиболее предпочтительным алгоритмом его реализации считается латентное размещение Дирихле (LDA), так как в результате его применения получается более корректный набор тематик в связи с уходом от нормального распределения. Основное предположение LDA состоит в том, что каждый документ с некоторой вероятностью может принадлежать множеству тем. Тема в данном контексте – это совокупность слов, где каждое слово имеет

некоторую вероятность принадлежности к ней. В результате модель LDA представляется в виде произведения матрицы «документы-темы» на матрицу «темы-слова». Данные матрицы могут быть найдены при помощи сэмпирования Гиббса, то есть путем выбора сэмпла (выборки) из совместного распределения.

Word2vec [4], в свою очередь, является развитием векторного представления слов. Был разработан группой исследователей Google в 2013 году. Суть заключается в установлении близости между словами на основе схожести их контекстов, то есть слова, встречающиеся в тексте рядом с одинаковыми словами, в векторном представлении будут иметь близкие координаты. Получаемые на выходе координатные представления векторов-слов позволяют вычислять «семантическое расстояние» между словами.

Word2vec может быть однонаправленным, учитывающим контекст слева от слова, и двунаправленным, учитывающим контекст как слева, так и справа от слова. В качестве однонаправленного *Word2vec* можно привести реализацию *InferSent* [5] от компании Facebook. Он ориентирован на так называемый *sentence embeddings*, то есть это векторное представление не просто слов, а целых предложений в векторном пространстве. Он обучается на данных о естественном языке и хорошо обобщает множество различных задач. В рамках этого метода создается словарь векторов важных слов, который кодируется в массив этих векторов. Данный метод позволяет определить и визуализировать «важность» каждого слова в предложении.

Примером реализации двунаправленного *Word2vec* является *BERT* [6]. Он представляет из себя подход к формированию языковой модели для решения задач обработки естественного языка, разработанный командой Google AI в ноябре 2018 года. *BERT* реализует двунаправленную контекстно-зависимую обработку, то есть языковая модель учитывает в определении контекста конкретного слова то, что стоит в предложении перед ним и после него. Кроме того, *BERT* учит улавливать логические связи между предложениями (например, действительно ли второе предложение должно следовать за первым или оно случайное).

2. Описание предлагаемой технологии

Разработанная в рамках исследования технология основана на применении дистрибутивного подхода к анализу естественно-языковых текстов для автоматизации пополнения тематического тезауруса. Технология позволяет извлекать из текстового корпуса определенной тематики новые лексические единицы или, иначе говоря, сущности, которые в дальнейшем могут быть использованы в качестве кандидатов для пополнения тезауруса.

На данном этапе исследования было решено применить классический латентно-семантический анализ (в виде сравнения двух термов между собой) для выявления новых терминов с целью пополнения имеющегося тезауруса. Таким образом, при помощи ЛСА необходимо найти некоторые слова в корпусе текстов, близкие к тем, которые на данный момент имеются в тезаурусе.

ЛСА предполагает формирование исходной матрицы термы-на-документы, где строки – это слова, столбцы – документы, а в ячейках содержатся веса, учитывающие частоту встречаемости слова в документе. На данной матрице, как правило, применяется SVD-разложение (разложение по

сингулярным числам), которое позволяет разложить исходную матрицу на три составляющие, линейная комбинация которых является достаточно точным приближением к исходной матрице:

$$A = U \cdot S \cdot V^t, \quad (1)$$

где U и V^t — ортогональные матрицы, S — диагональная матрица.

Используя полученное разложение (1), можно выявить основные зависимости между терминами и документами, латентно присутствующие в исходной матрице. Особенность такого разложения состоит в том, что если в матрице S оставить только k наибольших сингулярных значений, то линейная комбинация получившихся матриц будет наилучшим приближением исходной матрицы A к матрице \hat{A} ранга k .

В рамках данной работы применение ЛСА состояло из следующих этапов:

- 1) формирование текстового корпуса документами определенной тематики;
- 2) извлечение из текстового корпуса отдельных слов и коллокаций (биграмм);
- 3) препроцессинг отдельных слов и коллокаций – исключение стоп-слов и небуквенных символов, нормализация;
- 4) вычисление TF-IDF для каждой униграммы и биграммы;
- 5) формирование исходной матрицы термины-на-документы и ее SVD-разложение;
- 6) получение векторов слов тезауруса и слов текстового корпуса;
- 7) вычисление расстояний между этими векторами;
- 8) выявление наиболее близких слов текстового корпуса к терминам тезауруса.

В контексте применения латентно-семантического анализа к текущей задаче пополнения тезауруса арктической деятельности возникли две проблемы. Первая проблема состоит в том, что исходная тематическая выборка текстов может не содержать термины, которые есть в тезаурусе. Следовательно, это приводит к невозможности нахождения в текстовом корпусе каких-либо слов, близких к таким терминам тезауруса. По этой причине дополнительно был сформирован набор так называемых связывающих текстов, которые потенциально связаны с выбранной тематикой и при этом содержат термины тезауруса. Идея в том, чтобы связать термин тезауруса как-нибудь с текстовым набором через другие слова. Вторая проблема заключается в том, что большая часть терминов тезауруса представлена словосочетаниями (биграммами), поэтому из текста необходимо извлекать подобные структуры. Для ее решения была использована python-библиотека `sraCu` [7], позволяющая анализировать синтаксические деревья предложений исходных текстов и, в зависимости от отношения между токенами, выявлять необходимые конструкции.

3. Применение технологии на примере пополнения тезауруса арктической деятельности

Изначально латентно-семантический анализ был опробован на тестовом корпусе текстов (фрагментов учебных пособий) по сетевым технологиям. Он состоит из 104 документов, в которых выявлено: общее число токенов — 238 401,

уникальные слова и биграммы – 67 683. Для данного текстового корпуса предварительно был сформирован экспертный набор, представляющий эталонный результат. Он включает в себя некоторый термин из предметной области и слова, связанные с ним с точки зрения эксперта (2):

{термин: [связанные лексемы и биграммы]} . (2)

Следуя описанным выше этапам решения задачи, аналогичный по структуре набор терминов был сформирован с помощью ЛСА для сетевого корпуса текстов, и затем было оценено сходство этих двух наборов в виде полноты и точности, которые составили 0,29 и 0,015 соответственно. Низкие значения данных показателей прежде всего связаны с малым размером текстового корпуса, поскольку для более точных результатов требуется около 3–5 млн словоупотреблений. Но тем не менее, ненулевая полнота предполагает некоторую вероятность нахождения каких-то новых терминов в целевом корпусе текстов. Поэтому далее латентно-семантический анализ применялся уже к корпусу текстов арктической тематики.

Также придерживаясь описанных ранее этапов применения ЛСА, первоначально осуществлялось формирование арктического текстового корпуса, который на текущий момент состоит из 109 документов (фрагментов научных статей), в которых найдено 232 261 токен и 95 760 униграмм и биграмм.

Биграммы, учитывая упомянутые ранее особенности тезауруса, извлекались следующим образом. Первым делом языковая модель spaCy была обучена на корпусе синтаксических деревьев русского языка [8], так как в основном spaCy применяется только для английского языка. Затем биграммы извлекались по правилам, которые требуют наличие между словами биграммы определенного вида отношений – amod (adjectival modifier) и nmod (nominal modifier). Иначе говоря, это виды именных словосочетаний, где главным словом является имя существительное (например, существительное и прилагательное, существительное и существительное).

Применив поэтапно латентно-семантический анализ на арктическом текстовом корпусе, было выявлено большое количество общеупотребительной лексики, что подтверждает низкую точность данного метода. Но несмотря на это, удалось выявить новые термины, которые могут быть рассмотрены в качестве кандидатов для пополнения текущего тезауруса. Ниже представлены некоторые из обнаруженных понятий, связанные с понятиями тезауруса (выделены курсивом):

- *Айсберг*: галоклин (слой воды);
- *Антарктида*: КМАГЭ (экспедиция), гравиметрический (весовой анализ);
- *Арктический туризм*: Грумант (поселок), свердруп (единица измерения), Шхера (архипелаг);
- *Ледник*: Маркхема (шельфовый ледник), Элсмир (остров), Эйлс (ледяной массив);
- *Полярный регион*: алеут (коренное население), накипной (лишайник), кайра (морская птица), Банкс (остров).

Заключение

Автоматическое пополнение онтологий, а именно тематического тезауруса, является на сегодняшний день весьма актуальной задачей онтологического инжиниринга, а также анализа естественно-языковых текстов. Применение дистрибутивных подходов позволяет существенно облегчить данную задачу в автоматизации некоторых этапов анализа текстов в силу простоты их использования. Именно латентно-семантический анализ позволяет выявить статистические зависимости между терминами, использованными в коллекции документов, что существенно ускоряет процесс поиска кандидатов для пополнения тезауруса.

В данной статье приведена технология пополнения тезауруса проекта национального стандарта «Исследования в полярных регионах. Основные термины и определения» на основе одного из дистрибутивных подходов — латентно-семантического анализа естественно-языковых текстов. В ходе применения данной технологии удалось выявить слова, которые могут быть рассмотрены в качестве потенциальных кандидатов для пополнения заявленного тезауруса.

В дальнейшем планируется доработать текущую технологию путем использования других дистрибутивных подходов и расширенного текстового корпуса.

Литература

1. Heidenreich H. Introduction to Word Embeddings. Режим доступа: <https://towardsdatascience.com/introduction-to-word-embeddings-4cf857b12edc>.
2. Landauer T.K., Foltz P.W., Laham D. Introduction to Latent Semantic Analysis // Discourse Processes, 25. 1998. С. 259–284.
3. Li S. Topic Modeling and Latent Dirichlet Allocation (LDA) in Python. – Режим доступа: <https://towardsdatascience.com/topic-modeling-and-latent-dirichlet-allocation-in-python-9bf156893c24>.
4. Gensim models – Word2vec. Режим доступа: <https://radimrehurek.com/gensim/models/word2vec>.
5. InferSent. – Режим доступа: <https://github.com/facebookresearch/InferSent>.
6. BERT. – Режим доступа: <https://github.com/google-research/bert>.
7. SpaCy библиотека для обработки естественного языка. Режим доступа: <https://spacy.io/>.
8. Размеченный в нотации Universal Dependencies русский корпус текстов. Режим доступа: https://github.com/UniversalDependencies/UD_Russian-GSD.

Сведения об авторах

Малоземова Марина Леонидовна

инженер-исследователь

e-mail: malozemova@iimm.ru

Marina L. Malozemova

research engineer

Ломов Павел Андреевич

к.т.н., старший научный сотрудник

e-mail: lomov@iimm.ru

Pavel A. Lomov

PhD (Tech. Sci.), senior researcher

DOI: 10.25702/KSC.2307-5252.2019.9.91-97

УДК 004.047, 004.912

В. В. Диковицкий, М. Г. Шишаев, В. К. Пимешков

Институт информатики и математического моделирования ФИЦ КНЦ РАН

МЕТОД АВТОМАТИЗИРОВАННОГО ИЗВЛЕЧЕНИЯ ПОНЯТИЙ И ПАРАДИГМАТИЧЕСКИХ ОТНОШЕНИЙ ТЕЗАУРУСА ИЗ ТЕКСТОВ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ НА БАЗЕ ЛЕКСИКО-СИНТАКСИЧЕСКИХ ШАБЛОНОВ

Аннотация

Работа посвящена проблеме автоматизации извлечения знаний из неструктурированного текста с целью их прикладного использования в задачах извлечения фактов, формирования и пополнения тезауруса, анализа согласованности документов. Для извлечения и структурирования знаний используются методы статистического и лингвистического анализа.

Ключевые слова:

Семантика, анализ текста, лексико-синтаксический шаблон.

V.V. Dikovitskiy, M.G Shisaeu, V.K. Pimeshkov

METHOD OF AUTOMATED EXTRACTION OF CONCEPTS AND PARADIGMATIC RELATIONS OF THESAURUS FROM TEXTS IN NATURAL LANGUAGE ON THE BASIS OF LEXICO-SYNTACTIC TEMPLATES

Abstract

The work is devoted to the problem of automating the extraction of knowledge from unstructured text with the aim of their application in the tasks of extracting facts, the formation and replenishment of a thesaurus, analysis of document consistency. To extract and structure knowledge, methods of statistical and linguistic analysis are used.

Keywords:

Semantics, text analysis, lexical and syntactic template.

Введение

С ростом объемов текстовых данных, подлежащих обработке в рамках различных прикладных проблем, становится все более актуальной задача автоматического извлечения и обработки формализованных знаний. В настоящее время существует множество методов и подходов к автоматизированному анализу текста. Семейство технологий глубинного анализа текстов их смысла и представления его в базах знаний объединено общим названием Text Mining [1]. Разрабатываемые на основе статистического и лингвистического анализа, методов искусственного интеллекта, эти технологии предназначены для проведения смыслового анализа. [2] Методами обработки текстов на естественном языке, в основе которых лежит формирование и использование ассоциативно-онтологического представления данных [3] успешно решаются задачи выделения признаков текста для построения поисковых индексов, автоматического реферирования научных и технических документов, отнесения текста к предметной области, поиска в коллекции документов.

Интегральный подход к анализу структуры предложения с точки зрения функционально-коммуникативной грамматики предложен Г. А. Золотовой [4] и реализован в работе [5]. В работе [6] представлен семантический анализ

предложений с использованием падежных грамматик и семантических валентностей. При этом семантика предложения задается через связи главного слова (глагола) с его семантическими актантами. Моделирование синтаксиса и семантики на основе нейросетевого подхода получило широкое распространение в связи с ростом вычислительной мощности, развитием моделей и методов параллельной обработки данных (Mapreduce, Hadoop), и методов машинного обучения (Keras, Tensorflow, CatBoost). В проекте SyntaxNet представлен опыт использования рекуррентных нейронных сетей для морфологического и синтаксического анализа предложений без использования словарей и грамматик. Одной из главнейших проблем в данной области до недавнего времени являлась размерность данных и семантическая разреженность текстовой информации. Данная проблема решается алгоритмами векторного представления слов. Например, в наборе библиотек Word2Vec [7] реализованы подходы дистрибутивной семантики, позволяющие отразить слова языка в многомерное векторное пространство. Подобное отражение позволяет производить оценку близости слов и контекста их употребления.

Многоуровневый анализ текста использован для поиска высказываний деонтической логики в работе [8]. Многоуровневый анализ текста предполагает совместное использование результатов синтаксического, ассоциативного, дистрибутивного анализа текста для построения семантической сети, в форме взвешенного мультиграфа.

Многоуровневый анализ текста

Задачи анализа текста многоаспектна и включает отдельные подзадачи, в частности: Извлечение сущностей (entity recognition), извлечение признаков (feature extraction), определение отношений между объектами. Объединение решения всех перечисленных задач позволяет извлекать формализованные знания о предметной области в виде понятий и структуры их взаимосвязей.

В данной работе используется автоматическое извлечение объектов, и их свойств и отношений путем многоуровневого анализа тематических текстов предметной области. Многоуровневость анализа заключается в комбинации статистических и лингвистических методов с целью взаимного уточнения результатов анализа. На первом этапе применяются методы дистрибутивного анализа текста — определение частотных характеристик каждого слова в рассматриваемой коллекции документов, учет статистики совместного употребления, определение контекстной близости слов.

Формируется взвешенная ассоциативная семантическая сеть [8], весовой коэффициент формируется исходя из частотных характеристик совместного употребления слов. Векторизация слов используется на данном этапе для оценки контекстной близости слов с целью определения устойчивых словосочетаний и денотатов каждого слова. Также векторизация позволяет проводить анализ относительно небольших наборов текстовых документов, уточняя результаты ассоциативного отношения. Показателем контекстной близости является косинусная близость между векторами слов, полученными дистрибутивными моделями на основе больших корпусов текстов.

На втором уровне производится формирование синтаксических деревьев исходных текстов, и последующая интеграция полученных деревьев в модель

семантической сети с множественными связями. Для синтаксического и морфологического анализа использована основанная на машинном обучении библиотека SyntaxNet[9], включающая синтаксическую модель русского языка. Синтаксическую разметку составляют 44 отношения Universal Dependencies[10] (UD). UD разметка упорядоченно представляет межязыковые соответствия и основанна на существующих стандартах разметки.

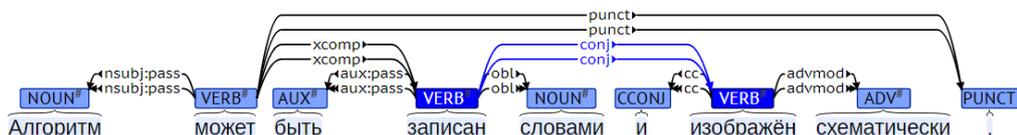


Рис. 1 Пример синтаксической разметки предложения.

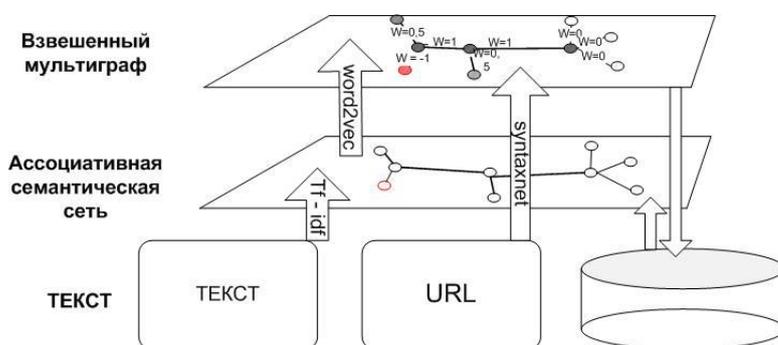


Рис. 2 Многоуровневый анализ текста

Интеграция результатов синтаксического анализа осуществляется путем вычисления составной оценки семантической близости понятий на основе сходства понятий, смежных понятий, а также их морфологических и синтаксических свойств. Результатом интеграции является семантическая сеть в виде ориентированного взвешенного мультиграфа, вершинами которого являются лексемы, а бинарные отношения обозначают синтаксическую роль лексем.

Определение парадигматических отношений основано на предположении о существовании связи лексико-семантического и синтаксического уровней языка при актуализации оценочных значений языковых единиц. Интерпретация оценочных предикатов как особых лексико-грамматических классов слов предложена в работах Г.А. Золотовой.

Для определения морфологических и синтаксических характеристик, определяющих парадигматические отношения, было произведено построение взвешенной семантической сети, полученной на наборе аннотаций русскоязычных статей Википедии. Для определения и систематизации синтаксические, морфологические и статистические характеристики понятий мультиграфа, выражающих одно парадигматическое отношение, были сгруппированы. В качестве эталона парадигматических отношений общепотребительной лексики использован тезаурус WordNet. Для приведения к

формату списка двуместных отношений синонимичные ряды тезауруса были разложены на множества двуместных отношений. Парадигматические отношения тезауруса над синонимичными рядами при этом были установлены между каждой парой слов из различных синонимичных рядов. Далее получено пересечение множеств слов присутствующих в тезаурусе и семантической сети, а также двуместные отношения между словами. В результате было получено 2534 бинарных отношений.

В таблице представлена статистика использования двуместных синтаксических отношений в различной роли парадигматических отношений тезауруса.

Таблица 1. Статистика использования двуместных синтаксических отношений в различной роли парадигматических отношений

UD \ WordNet	Элемент группы	Антоним	Мероним	Гипероним	Ассоциация	Гипоним	Атрибут
conjunct	1	263	30	694	5	360	2
adverbial clause modifier				2		1	
adjectival modifier	19	2		9		10	39
nominal modifier	36	14	30	268		464	4
nominal subject	6	7	9	53		80	
oblique nominal		2				4	
appositional modifier	5		9	10		7	
determiner						2	2
clausal subject				1			
clausal complement				1			
object				1		2	
case marking						1	
open clausal complement				1			

Из таблицы следует превалирование определенных групп синтаксических отношений текста википедии для некоторых парадигматических отношений тезауруса. Гипонимы в 39% выполняют синтаксическую роль «conjunct» и в 51% случаев роль «nominal modifier». Гиперонимам в 68% случаев соответствует синтаксическая роль «conjunct» и в 26% «nominal modifier». Для разделения гипонимов и гиперонимов учитывается направление двуместного отношения. Отличить парадигматические отношения в рамках одной синтаксической роли предложено также за счет учета определенной Word2Vec контекстной близости. Также различные синтаксические роли имеют различные связи с другими вершинами семантической сети, в частности применимы с различными предложениями и пунктуацией. Пример антонимов («автомобиль, а не самолет», «автомобиль и другие транспортные средства»). Учитываются и морфологические свойства слов, определенные SyntaxNet. Получены классы объектов, связанные одним и тем же парадигматическим отношением и определенные в множество по схожим морфологическим признакам. В таблице 3 представлены примеры гиперонимов и меронимов, полученных при анализе аннотаций статей Википедии.

Таблица 2. Морфологические свойства слов в парадигматических бинарных отношениях

conjunct - Гипероним	Кол -во	%	conjunct - Гипоним	Кол -во	%
существительное существительное	- 663	- 663	существительное существительное	- 329	- 329
Число=Мн.ч.	506	76, 3	Род=Жен.	149	45, 3
Степень=Pos	440	66, 4	NOUN__Одушевленность=Неоду ш.	137	41, 6
Падеж=Винительный	424	64, 0	Число=Ед.ч.	130	39, 5
ADJ__Одушевленность=Неодуш.	413	62, 3	CCONJ__	115	35, 0
Род=Жен.	141	21, 3	Падеж=Именительный	109	33, 1
NOUN__Одушевленность=Неоду ш.	133	20, 1	Число=Мн.ч.	45	13, 7
Падеж=Именительный	97	14, 6			
Число=Ед.ч.	75	11, 3			

глагол - глагол	27	27	глагол - глагол	26	26
VERB__Вид=Imp	12	44, 4	VERB__Вид=Imp	8	30, 8
VerbForm=Inf	12	44, 4	VerbForm=Inf	8	30, 8
Залог=Действительный	12	44, 4	CCONJ__	7	26, 9
CCONJ__	9	33, 3	Число=Ед.ч.	7	26, 9
ADV__Степень=Pos	4	14, 8	Залог=Действительный	7	26, 9

Таблица 3. Примеры меронимии и гиперонимии

Гипонимия	
река	волга
змея	горыныч
армия	махди
университет	организация
бог	гор
лингвист	чарльз
река	свислочь
государство	израиль
марка	электроника

Меронимия	
губерния	империя
митрополия	церковь
слово	словосочетание
текст	модуль
история	философия
алфавит	язык
буква	слово
масса	атом
столица	империя
вид	род
буква	алфавит
состояние	вещество
семейство	вид

Заключение

Среди общего количества пересекающихся слов, присутствующих, как в тезаурусе, так и анализируемых текстах аннотаций Википедии, было получено 2534 бинарных отношений. Были отмечены верно относительно тезауруса WordNet 1697 прагматических отношений, что составляет 67% от общего количества. Метод имеет потенциал к повышению точности результатов за счет разработки интегрированной оценки наличия прагматических отношений и учета в ней всех доступных характеристик. Систематизация которых является направлением дальнейшей работы.

Литература

1. Berry M.W. Survey of Text Mining. Clustering, Classification, and Retrieval. Berlin: Springer-Verlag, 2004. 244 p.
2. Капитонов О. А., Тютюнник В.М. Логико-лингвистическая модель семантической разметки веб-страниц // Фундаментальные исследования. – 2013. № 1–3. С. 714–717.
3. Кулешов С. В., Зайцева А. А., Марков В. С. Ассоциативно-онтологический подход к обработке текстов на естественном языке // Интеллектуальные технологии на транспорте. 2015. № 4.
4. Золотова Г. А., Онипенко Н. К., Сидорова М. Ю. Коммуникативная грамматика русского языка. // Институт русского языка им. В. В. Виноградова РАН. 2004. 44 с. ISBN: 5-88744-050-3.
5. Осипов Г. С. Методы искусственного интеллекта. 2011 г. ISBN: 978-5-9221-1323-6.
6. Барышникова Н. Ю. Обработка запросов на естественном языке на основе семантических сетей и шаблонов // Вестник АГТУ. Серия: Управление, вычислительная техника и информатика. 2016. № 4.
7. Collobert R., Weston J., Bottou L., Karlen M., Kavukcuoglu K. and Kuksa P. (2011) Natural Language Processing (Almost) from Scratch. J. // Mach. Learn. Res. 12 (November 2011), 2493-2537. WordToVec URL: <https://arxiv.org/pdf/1103.0398.pdf>.

8. V. V. Dikovitsky, M. G. Shishaev. Automated Extraction of Deontological Statements Through a Multilevel Analysis of Legal Acts // Computational and Statistical Methods in Intelligent Systems, pp. 102-110.
9. Библиотека программного обеспечения с открытым исходным кодом для Machine Intelligence TensorFlow. URL: <https://www.tensorflow.org>.
10. Фреймворк для кросс-лингвистически последовательной грамматической аннотации на 60 языках. URL: <http://universaldependencies.org>.

Сведения об авторах

Диковицкий Владимир Витальевич

к.т.н., старший научный сотрудник

e-mail: dikovitsky@gmail.com

Vladimir V. Dikovitskiy

Ph.D. (Tech. Sci.), researcher

Шишаев Максим Геннадьевич

д.т.н., главный научный сотрудник

e-mail: shishaev@iimm.ru

Maxim G. Shishaev

Dr.Sci (Tech.), lead researcher

Пимешков Вадим Константинович

стажер-исследователь

e-mail: pimeshkov@iimm.ru

Vadim K. Pimeshkov

research assistant

DOI: 10.25702/KSC.2307-5252.2019.9.97-108

УДК 004.89

О. В. Фридман

Институт информатики и математического моделирования ФИЦ КНЦ РАН

ЛОГИЧЕСКИЕ НЕЙРОННЫЕ СЕТИ: МЕТОДЫ АВТОМАТИЧЕСКОГО КОСТРУИРОВАНИЯ, РЕДУКЦИИ, ИЗВЛЕЧЕНИЯ ПРАВИЛ*

Аннотация

Рассмотрен класс логических нейронных сетей. Представлен обзор методов и алгоритмов, применяемых при построении нейронных сетей, методы извлечения правил. Рассматриваются вопросы применения логических нейронных сетей для решения различных задач.

Ключевые слова:

логические нейронные сети, моделирование, методы и алгоритмы построения нейронных сетей, извлечение правил, применение логических нейронных сетей.

* Работа выполнена в рамках темы НИР «Развитие методологии построения интеллектуальных мультипредметных систем информационной поддержки регионального развития на примере территорий Арктической зоны Российской Федерации (АЗРФ)» (№ 0226-2019-0036).

O. V. Fridman

LOGIC NEURAL NETWORKS: METHODS OF AUTOMATIC DESIGN, REDUCTION, RULE EXTRACTION

Abstract

The class of logical neural networks is considered. A review of the methods and algorithms used in the construction of neural networks, reduction, and rule extraction are presented. The problems of using logical neural networks to solve various problems are considered.

Keywords:

logical neural networks, modeling, methods and algorithms for constructing neural networks, reduction, rule extraction, the use of logical neural networks.

Введение

В настоящее время разработаны десятки различных моделей нейронных сетей, которые получили применение во многих сферах современной жизни, в частности, для задач анализа большого количества данных, для трудноформализуемых и неформализуемых задач, для задач с комплексными алгоритмами и алгоритмами с неявными параметрами и др. Нейронные сети получили широкое распространение при необходимости одновременного анализа большого числа данных, выделении тенденций и прогнозирования в зависимости от полученной информации, как, например, на рынке бумаг или валютном рынке.

Архитектуру нейронной сети можно характеризовать схемой SIMD (Single Instruction — Multiple Data), то есть выполнение одной функции для множества данных [1].

Для реализации структуры сети, необходимо ввести связи между нейронами. Основные виды связей, используемых в нейронных сетях, следующие:

- прямая связь: процесс передачи выходного сигнала происходит последовательно, с выхода одного слоя, на вход следующего;
- обратная связь: выходные сигналы передаются на вход предыдущего слоя, что позволяет производить рекуррентные операции на нейронных сетях;
- боковая связь: выходные сигналы передаются внутри слоя от нейрона к нейрону.

Перед созданием нейронной сети проводится анализ задачи, в процессе которого определяются топология и размеры сети, передаточные и активационные функции и прочие параметры. После завершения этапа формализации и переноса поставленной задачи на модель нейронной сети, начинается процесс обучения.

Обучение нейронной сети требует много времени, и чем сложнее структура, тем дольше будет проходить обучение. Для ускорения данного процесса существует лишь одно решение – улучшение самого алгоритма обучения. Наиболее эффективной, с точки зрения временных затрат модификацией для процессов обучения является распараллеливание алгоритма.

После завершения стадии обучения, структура нейронной сети фиксируется, то есть дальнейших изменений весов и связей не происходит, сеть переходит в рабочий режим.

В зависимости от поставленных целей, принято классифицировать нейронные сети. Для систем анализа изображений, задач автоматизации

управления, экономической сферы, существуют типовые нейронные сети, которые далее модифицируются под конкретно поставленную задачу. В данной работе будет рассмотрен класс логических нейронных сетей [2].

1. Логические нейронные сети

Пусть имеется некоторое множество высказываний, по которому путём неарифметических преобразований должен быть получен результат. Для работы с подобными множествами, необходимо их формализовать, представив все высказывания в виде предикатов, получив возможность обработки предикатов при помощи операций алгебры логики. Решение задачи зависит от истинности или ложности высказываний входного множества. Исходя из этого, логическую нейронную сеть можно охарактеризовать как нейронную сеть, базисом для которой является булева алгебра.

Определим функционал и вид сети. Если взять совокупность всех событий, и выделить среди них те высказывания, которые полностью покрывают смысловой диапазон, получим новое множество. В том случае, если каждая допустимая ситуация характеризуется тем, что истинное значение принимает только одно высказывание этой совокупности, полученное множество будет называться исчерпывающим множеством событий. На основе полученного множества, введя некоторую иерархию, будем строить логические цепочки, позволяющие однозначно получить результат. Систему из таких логических путей, в случае её непротиворечивости и полноты, можно рассматривать как систему принятий решений (СПР). Прежде чем перейти к построению нейронной сети из СПР, необходим ещё один шаг в формализации задачи – определение величины входных сигналов.

Поскольку в сети, нейроны представлены в некотором стандартном виде, то величины возбуждения, синаптические связи, пороги и функции должны быть либо одинаковы, либо лежать в общем диапазоне значений для всей сети.

После этого шага, можно приступить к выбору параметров сети.

В модели логической нейронной сети рассматриваются 2 типа нейронов, реализующих логические функции: нейрон-конъюнктор и нейрон-дизъюнктор [3], объединённых связями, полученными из условий задачи. Опираясь на булеву алгебру, логические нейронные сети покрывают широкий спектр задач, а также осуществлять логический вывод типа «причина-следствие», что применяется для различных систем управления, а также для систем автоматического доказательства теорем.

Главным отличием логических нейронных сетей от сетей других видов является то, что не существует одного шаблона для реализации всего класса логических нейронных сетей. Каждая такая сеть строится под задачу, путём объединения в сети конъюнкторов и дизъюнкторов. Только путём анализа конкретной задачи может быть создана логическая нейронная сеть.

Проанализировав входное множество, необходимо выделить все используемые переменные, подаваемые на вход, после чего надо построить структуру, а затем обучить её. Первая часть обучения происходит методом трассировки, когда по предъявляемому сети эталону выстраивается путь от входных значений к выходному. При предъявлении нескольких эталонов нейронная сеть строит пути последовательно, выбирая ещё не задействованные

нейроны и вводя необходимые связи. Данный алгоритм повторяется до тех пор, пока не будет достигнута полная определенность для всех входных эталонов, предъявляемых сети. Для алгоритма трассировки нейронной сети используется понятие матрицы следования – матрицы, в которой строки и столбцы — нейроны, а значения, стоящие на их пересечении отвечают за наличие или отсутствие связи. Алгоритм позволяет выстроить связи, цепочки следования из входного слоя к выходному через скрытые слои. Таким образом, получают готовую структуру логической нейронной сети. Следующим шагом обучения является приведение нейронной сети после трассировки, то есть коррекция весов и порогов для достижения необходимых результатов. В итоге, полученная логическая нейронная сеть способна на основе формализованных входных данных, выдать выходное значение на основе логических операций. Таким образом, логическая нейронная сеть, по сути, является автоматизированной системой логического вывода. Сферой применения таких сетей являются системы управления, банковская сфера, различные системы анализа, а также система автоматического доказательства теорем, основывающаяся на методе логического вывода.

2. Построение логических нейронных сетей

Рассмотрим методы и алгоритмы, применяемые при построении нейронных сетей. При построении нейронной сети используются конструктивные методы, которые предполагают последовательное построение (редукцию) нейронной сети путем поэтапного добавления (удаления) элементов архитектуры сети (связей, нейронов, скрытых слоев). Алгоритмы конструирования функционируют по определенному правилу таким образом, что любое изменение архитектуры на каждом этапе гарантированно уменьшает значение ошибки сети. Правило подчиняется следующей схеме: локализация неверно решенных примеров и коррекция архитектуры сети только для них, не затрагивая, по возможности, корректных.

В работе [3] подробно описаны алгоритмы автоматического конструирования нейронных сетей реализующих произвольные логические функции (бинарные алгоритмы).

Для конструирования нейронных сетей специального вида применяются другие методы и алгоритмы, такие как метод динамического добавления узлов, каскадная корреляция и др.

Другим направлением, которое можно выделить среди конструктивных алгоритмов, являются методы редукции нейронной сети. Все подходы к редукции объединены следующей общей схемой работы. Исходной является некоторая, уже обученная, нейронная сеть. В предположении, что размеры этой сети можно уменьшить, отсекаются определенные каким-либо образом избыточные нейроны или существующие связи между ними. Как правило, в качестве критерия выбирается степень воздействия нейрона (связи) на величину ошибки, а убираются соответственно слабо влияющие нейроны (веса). После упрощения нейронной сети повторяется процесс дообучения и повторного усечения до тех пор, пока величина ошибки не достигнет желаемой величины или топология нейронной сети не достигнет необходимой простоты. Очевидными недостатками подхода усечения являются неопределенность в принципах построения исходной нейронной сети и большой объем вычислений. Обычно метод усечения

применяется для экономии ресурсов при необходимости аппаратной реализации уже найденной нейронной сети.

Исторически, первой техникой упрощения сети была техника усечения, предложенная в 1989 г. Мозером и Смоленским [4].

Более исследованным подходом к упрощению нейронной сети является техника усечения весов, основанная на оценке значимости весов связей. В отечественной литературе техника усечения весов носит название процедуры контрастирования и впервые была предложена в работе [5].

Для направления упрощения сети методом усечения весов широко известными являются методы оптимального повреждения мозга (Optimal Brain Damage, OBD) [6] и его более позднее развитие – лучший нейрохирург (Optimal Brain Surgeon, OBS) [7].

Современным способом построения нейронных сетей является использование эволюционного подхода к поиску по пространству возможных архитектур. Исторически сложились три методологии эволюционного поиска [8]: эволюционное программирование Фогела, 1966 г., эволюционные стратегии Реченберга, 1973 г. и генетические алгоритмы (ГА) Холланда, 1975 г. [9]. Общую схему поиска архитектуры нейронных сетей с помощью ГА можно представить следующим образом. Информация об архитектуре нейронной сети специальным образом кодируется в генетический код. Затем генерируются поколения, к которым применяются стандартные генетические операции. Оценка каждого индивидуума производится следующим образом: нейронная сеть восстанавливается из генетической строки и производится ее тестирование.

Таким образом, конструктивные алгоритмы позволяют быстро создавать сети, но жестко заданные правила построения не позволяют совершать отклонения в пространстве архитектур от некоторой окрестности. Это позволяет говорить о конструктивном подходе, как о подходе, реализующем локальный поиск, и сильно подверженный попаданию в локальные минимумы.

Особое место занимают алгоритмы редукции: они позволяют получить неплохие результаты, но предполагают полноту имеющейся модели до начала работы. Иными словами, до начала работы алгоритма уже предполагается наличие некоторой нейронной сети, успешно решающей проблему, для которой делается предположение о возможном ее упрощении. Это условие сильно ограничивает круг возможного применения методов усечения.

Эволюционные алгоритмы несомненно реализуют охват всего пространства архитектур и соответственно реализуют глобальный поиск, с большей вероятностью попадания в глобальный минимум. Но подобный способ организации поиска занимает очень много времени [3].

3. Извлечение правил из обученных нейронных сетей

В работах [10, 11] рассмотрены вопросы совместного использования нейросетевых технологий с методами логического вывода и поддержки принятия решений в задачах «интеллектуального» анализа данных. Проведен анализ существующих алгоритмов и методов построения деревьев решений.

Особенностью алгоритмов и методов, применимых для решения задач интеллектуального анализа данных, является отсутствие ограничительных рамок априорных предположений о структуре выборки и виде распределений значений

анализируемых показателей, чему наилучшим образом соответствует использование нейросетевых технологий. Это обусловлено способностью нейронных сетей к моделированию нелинейных процессов, воспроизведению чрезвычайно сложных зависимостей, адаптивностью к условиям функционирования, а главное, способностью извлекать и обобщать существенные особенности из поступающей информации. Тем самым сеть осуществляет построение правил, однако эти правила содержатся в весовых коэффициентах, функциях активации и связях между нейронами, но обычно их структура слишком сложна для восприятия и определения влияния отдельного признака на выходное значение. Нейронная сеть, по сути, выступает «черным ящиком», на вход которого подаются исходные данные и на выходе получается некоторый результат, однако обоснования, почему было принято именно такое решение, не предоставляется. Правила содержатся в весовых коэффициентах, функциях активации и связях между нейронами, но обычно их структура слишком сложна для восприятия. Более того, в многослойной сети эти параметры могут представлять собой нелинейные, немонотонные отношения между входными и целевыми значениями. Таким образом, как правило, не представляется возможным отделить влияние определенного признака на целевое значение, потому что этот эффект может быть опосредован значениями других параметров. Другой сложностью является проблема выбора оптимальной топологии сети, значений параметров и структурных особенностей, которые бы наилучшим образом удовлетворяли решаемой задаче на имеющихся исходных данных. В связи с этим, особую важность приобретают вопросы совместного использования нейросетевых технологий с методами логического вывода и поддержки принятия решений, где в качестве основного подхода применяются деревья решений.

Дерево решений состоит из вершин двух типов. Вершины решений, содержащие вопросы, обозначаются окружностями. Цели или логические выводы обозначаются прямоугольниками. Вершины нумеруются и на дугах задаются условия. Каждая вершина может иметь не более одного входа. Пути движения по дереву с верхнего уровня на самые нижние определяют логические правила в виде цепочек конъюнкций [10].

Правила, выражающие закономерности, формулируются в виде продукций: «ЕСЛИ А ТО В» или в случае множества условий: «ЕСЛИ (условие 1) \wedge (условие 2) \wedge ... \wedge (условие N) ТО (Значение вершины вывода)». Их достоинством является простота и наглядность описания процесса поиска решения.

Деревья решения могут использоваться как самостоятельное средство анализа многомерных данных и поиска в них логических закономерностей. Общий принцип построения дерева заключается в рекурсивном разбиении объектов из обучающей выборки на подмножества, которые содержат объекты, относящиеся к одному классу.

Методы выделения закономерностей с помощью деревьев решений обладают свойством наглядности и позволяют находить такие связи, которые заключены не только в отдельных признаках, но и в сочетании признаков. Они предоставляют возможность прогнозировать и связывать различные параметры изучаемого явления в единое целое, что во многих случаях дает этим методам

значительное преимущество по сравнению с классическими методами многомерного анализа.

Одним из направлений использования деревьев решений для интеллектуального анализа данных является их применение для извлечения правил из нейронных сетей. Совместное использование нейросетевых технологий с методами логического вывода способно улучшить понимание структуры изучаемого явления за счет предоставления результата, полученного в ходе обучения нейронной сети, в виде иерархической, последовательной структуры правил типа «ЕСЛИ – ТО».

Можно выделить два подхода к извлечению правил из многослойных нейронных сетей. Первый подход заключается в извлечении набора глобальных правил, которые характеризуют классы на выходе непосредственно через значения входных параметров. Альтернативой является извлечение локальных правил, разделяя многослойную сеть на совокупность однослойных сетей. Каждое извлекаемое локальное правило характеризует отдельный скрытый или выходной нейрон с учетом элементов, которые имеют с ним взвешенные соединения. Затем правила объединяются в набор, который определяет поведение всей сети в целом.

Одним из алгоритмов извлечения правил из нейронных сетей, обученных решению задачи классификации, является метод NeuroRule [12]. Данный алгоритм основан на извлечении локальных правил и включает три основных этапа.

Этап 1. Обучение нейронной сети, когда двухслойный персептрон обучается вплоть до получения достаточной точности классификации. В первоначальный момент выбирается большое число промежуточных нейронов, и после обучения излишние нейроны и связи отбрасываются.

Этап 2. Прореживание нейронной сети. Обученная нейронная сеть содержит все возможные связи между входными нейронами и нейронами скрытого слоя, а также между скрытыми и выходными нейронами. Полное число этих связей обычно столь велико, что из анализа их значений невозможно извлечь доступные для понимания пользователем классифицирующие правила. Прореживание заключается в удалении излишних связей и нейронов, не приводящем к увеличению ошибки классификации сетью. Результирующая сеть обычно содержит немного нейронов и связей между ними, и функционирование такой сети поддается исследованию.

Этап 3. Извлечение правил. На этом этапе из прореженной нейронной сети извлекаются правила. Для этого проводят подготовку к извлечению правил, которая заключается в кодировании непрерывных величин, как на входе, так и внутри сети. Осуществляется кодирование признаков классифицируемых объектов, если они представляют собой непрерывные величины. Для их представления можно использовать бинарные нейроны и принцип кодирования типа «термометр». Значения, которые принимают нейроны скрытого слоя кластеризуются и заменяются значениями, определяющими центры кластеров. Число таких кластеров выбирается небольшим. После такой дискретизации активностей промежуточных нейронов производится проверка точности классификации объектов сетью. Если она остается приемлемой, то подготовка к извлечению правил заканчивается. Далее осуществляется извлечение правил, при этом движение по сети происходит от классифицирующих выходных нейронов

к входам сети. Предполагается, что эти правила достаточно очевидны при проверке и легко применяются к большим базам данных.

К недостаткам большинства алгоритмов извлечения правил можно отнести отсутствие универсальности и масштабируемости. В связи с этим, наибольший интерес представляет алгоритм TREPAN [13] и его модификации, которые лишены этих недостатков и не предъявляют никаких требований к архитектуре сети, входным и выходным значениям, алгоритму обучения и т.д. Данный подход осуществляет построение дерева решений на основе знаний, заложенных в обученную нейронную сеть, причем достаточно того, что сеть является неким «черным ящиком» или «Оракулом», которому можно задавать вопросы и получать от него ответы. Более того, алгоритм является достаточно универсальным и может применяться к широкому кругу других обученных классификаторов. Он также хорошо масштабируется и не чувствителен к размерности пространства входных признаков и размеру сети.

Основное преимущество данного подхода заключается в обобщающей способности искусственных нейронных сетей, что позволяет получать более простые деревья решений. Возможность генерации дополнительных примеров и использование «Оракула» для отнесения их к тому или иному классу позволяет компенсировать недостаток данных, часто наблюдающийся при построении деревьев на нижних уровнях. Таким образом, алгоритм позволяет извлекать структурированных знаний не только из чрезвычайно упрощенных нейронных сетей, но и из произвольных классификаторов, что делает возможным его применение в широком круге практических задач [11].

4. Применение логических нейронных сетей

Теперь рассмотрим вопросы применения логических нейронных сетей и конкретные примеры их применения при решении различных задач.

В работе [2] описаны различные задачи, решаемые при помощи логических нейронных сетей, в частности задачи анализа и прогнозирования экономических показателей (биржевых сделок), таких как котировки и объемы продаж активов на фондовом рынке, выбор наилучшего проекта строительства, определение стратегии международных торговых сделок и т.д., Решение этих задач основано на анализе многофакторных моделей. Разнообразие факторов политического, экономического и социального характера, а также такие конкретные показатели, как уровень инфляции, кредитные ставки банков, уровень безработицы, обменные валютные курсы, изменение объемов производства, денежная эмиссия и эмиссия ценных бумаг и др., представляют значительные трудности при формальном описании задачи, а главное — при определении зависимостей прогнозируемого результирующего показателя (котировка, объем продаж, цена и др.) от совокупности этих факторов. Даже предположение о линейности или мультипликативности не уменьшают неопределенности при аналитическом построении таких зависимостей. Таким образом, задача анализа и прогнозирования экономических показателей относится к классу трудно формализуемых [14].

В работе [15] объектом исследования является управление производственными запасами. Предмет исследования — инструментальные модели и методы управления производственными запасами в условиях неопределенности. Для строгого логического мышления, исключаящего

неопределенность, приходится оперировать не отдельными событиями и даже не исчерпывающими множествами таких событий, а композициями таких множеств. Между событиями, принадлежащими различным множествам, возможна зависимость, порождающая сложные высказывания. В работе использованы модели управления запасами с учетом денежной стоимости денег (модель со скидками, с дефицитом, с заёмными денежными средствами), традиционные стратегии управления («наибольшей осмотрительности», «дополнительного резерва» «процент от спроса»), для принятия решений разработана логическая нейронная сеть под задачу.

Решение задачи классификации [16] становится все более актуальным в связи с развитием технологии и разрастанием обрабатываемых объемов данных. Использование нейронных сетей обязательно при решении задач классификации, т.к. нейронные сети обладают способностью выявления значимых признаков и скрытых закономерностей. Преимуществами логической нейронной сети являются: более высокая точность классификации, большая скорость обучения и переобучения. Задача классификации встречается в самых разных областях человеческой деятельности и решается для последующего прогнозирования состояния исследуемой системы, например, в системах распознавания, в медицинской диагностике, банковской кредитной системы, контроля качества и т.д. и имеет практическую значимость. Существует огромное количество алгоритмов классификации, применяющиеся в решении определенных задач.

Под задачу классификации легко подстраивается логическая нейронная сеть с булевыми выходами. Логические нейронные сети хорошо подходят, когда классов, в которые нужно определить тот или иной объект не очень много. Это утверждение обусловлено тем, что в такой нейронной сети на выходном слое нейронов столько, сколько классов, и принадлежность объекта классификации к данному классу определяется 1 на выходе данного нейрона и 0 на всех остальных. Нейронные сети обладают способностью выявления значимых признаков и скрытых закономерностей. Преимуществами логической нейронной сети являются: более высокая точность классификации, большая скорость обучения и переобучения. Недостатками такой модели является разрастание нейронной сети при большом количестве классов, к которым надо отнести тот или иной объект и большом количестве характеристик объектов.

В работе [17] для синтеза антецедента (условия активации) решающих продукционных правил в случае представления характеризующих объект признаков в виде булевых переменных используются искусственные логические нейронные сети (ЛНС). Применяемый в настоящее время синтез ЛНС [18, 19] по сути, является итерационным процессом идентификации предикатов первого порядка, что не отвечает основным принципам самоорганизационного моделирования.

Между тем сетевые принципы синтеза моделей, лежащие в основе метода группового учета аргументов (МГУА), доказали свою перспективность в аналогичных случаях при анализе сложных открытых систем в медицинской предметной области [20].

В связи с этим аналогично МГУА предлагается использовать нейроны логической сети с ограниченным и небольшим количеством входов, постепенно усложняя структуру логической функции, «продвигаясь внутрь» по рядам ЛНС.

Ограничение на количество входных аргументов одного нейрона приближено к естественному интеллекту, который «одновременно удерживает»

в оперативной обработке ограниченное количество информации об объектах, процессах или семантических группах (включая альтернативные решения). Минимальным в этом случае является искусственный логический нейрон с двумя входами и одним выходом — бинарный нейрон

В качестве апробации возможностей ЛНС МГУА для синтеза условий активации использовалась задача синтеза адекватного решающего правила прогноза развития тромбоза легочной артерии (ТЭЛА) в послеоперационный период протезирования крупных суставов.

Полученные результаты позволяют предположить эффективность предлагаемого подхода синтеза решающих правил продукционного типа для баз знаний медицинских экспертных систем, как системообразующего модуля автоматизированных систем поддержки принятия решений в хирургии, поддерживающих оптимальную терапию и-или профилактику заболеваний, возникающих в послеоперационный период и приводящих впоследствии с высокой степенью риска к инвалидности или летальности. Это обусловливается тем, что предлагаемый подход ЛНС МГУА основан на синергетическом подходе к обработке гетерогенной структуры данных, представленных логическими значениями, позволяющих существенно уменьшить субъективизм при регистрации признаков — факторов риска.

В работе [21] рассматривается возможность применения логических нейронных сетей для решения задач авторизации для клиент-серверных приложений. В статье описана архитектура и метод построения процедуры авторизации для комплекса средств аутентификации информационных систем, построенных по технологии клиент-сервер. Ввиду необходимости анализа результатов аутентификации и принятия решения по результатам этого анализа модуль авторизации реализован в виде логической нейронной сети. Данная структура позволяет учесть и промоделировать все возможные ситуации, возникающие при аутентификации пользователей, так же предоставляет возможность разрешения спорных ситуаций, когда личность и принадлежность пользователя к данной системе установлено не точно. Метод принят для реализации ввиду оптимально подходящего для реализации решения поставленной задачи, т.е. задачи назначения прав и организационных мер пользователю в информационной системе. Этот метод позволяет реализовать достаточно гибкое и универсальное решение поставленной задачи, а именно позволяет расширять области принимаемых решений, при минимальной затрате усилий на модернизацию, позволяет принимать к анализу большее или меньшее количество результатов аутентификации, количество которых напрямую зависит от количества идентификаторов.

После завершения работы процедур аутентификации пользователей на входные нейроны подаются значения, характеризующие результаты верификаций того или иного идентификатора пользователя. Значения подаются на входной слой нейронов логической нейронной сети, модуля авторизации, тем самым активируя те или иные нейроны в зависимости от порога. Активация нейронов характеризует наступление того или иного события.

Количество событий напрямую связано с количеством идентификаторов используемых в комплексе и вероятности получения нечетких результатов аутентификации данных идентификаторов. Процесс работы модуля авторизации достаточно прост и состоит из трех этапов: получение данных от модулей

аутентификации, анализ полученных данных, принятие решения по результатам анализа с последующей передачей этого решения информационной системе.

Это далеко не полный перечень существующих задач, решаемых с использованием нейросетевого подхода, в частности, логических нейронных сетей.

Заключение

Искусственные нейронные сети в настоящее время широко применяются практически во всех областях связанных с обработкой информации. Далеко не полный список приложений нейронных сетей выглядит следующим образом:

- приближение функций: регрессионный анализ, предсказание временных рядов (финансовые приложения);
- распознавание образов: радары, распознавание лиц, почерка, речи, медицинская диагностика и т. д.;
- обработка данных: классификация, кластеризация, фильтрация, сжатие данных;
- управление: технологические процессы и системы, транспортные средства, манипуляторы, робототехника;
- системы принятия решений: игры, экспертные системы.

Проведенный обзор показал, что применение логических нейронных сетей позволяет создавать доступные широкому пользователю компьютерные системы мониторинга, управления и принятия решений практически во всех сферах деятельности. Проекты таких систем в области оценки финансов и риска, в управлении и диагностике, в искусстве и развлечениях показывают универсальность подхода, возможность переориентирования нейронных сетей на новые применения. В особенности это касается наиболее распространенных совершенных нейронных сетей – однослойных, в каждом решении использующих все факторы.

Литература

1. Хайкин С. Нейронные сети: полный курс, 2-е изд., испр. Издательский дом «Вильямс». М., 2006. 1104 с.
2. Барский А.Б. Логические нейронные сети. / А.Б. Барский; НОУ «ИНТУИТ», 2016. 492 с.
3. Аксенов С. В. Организация и использование нейронных сетей (методы и технологии). / С. В. Аксенов, В. Б. Новосельцев; Под общ. ред. В. Б. Новосельцева. Изд-во НТЛ. Томск, 2006. 128 с.
4. Mozer M. C., Smolensky P. Skeletonization: a technique for trimming the fat from a network via relevance assessment. // *Advances in Neural Information Processing Systems*. 1989. Vol. 1. Pp. 107–115.
5. Горбань А.Н. Обучение нейронных сетей. / А. Н. Горбань; М., СП ParaGraph, СССР – США, 1990. 160 с.
6. Le Cun Y., Denker J.S., and Solla S.A. Optimal Brain Damage / Y. Le Cun, J.S. Denker, S.A. Solla // *Advances in Neural Information Processing Systems II (Denver1989)*. 1990. Pp. 598–605.
7. Hassibi B., Stork D. G. Second Order Derivatives for Network Pruning: Optimal Brain Surgeon / B. Hassibi, D.G. Stork // *Neural Information Processing Systems*. 1992. Pp. 164–171.

8. Pedersen M. W., Hansen, L. K., Larsen, J. Pruning With Generalization Based Weight Saliences: gamma-OBD, gamma-OBS. / M. W. Pedersen, L. K. Hansen, J. Larsen // In Neural Information Processing Systems. NIPS. 1995. Pp. 521-527.
9. Holland J.H. Adaptation in Natural and Artificial Systems. / J.H. Holland; Univ. of Michigan Press., Second Ed., MA: The MIT Press edition. 1992. 211 p.
10. Гридин В. Н. Совместное использование нейросетевых технологий и деревьев решений для поиска логических закономерностей в данных. / В. Н. Гридин, В. И. Солодовников // Информационные технологии и нанотехнологии. 2017. С. 1756–1762.
11. Гридин В. Н. Построение деревьев решений и извлечение правил из обученных нейронных сетей. / В. Н. Гридин, В. И. Солодовников, И. А. Евдокимов, С.В. Филипков // Искусственный интеллект и принятие решений. 2013. № 4. С. 26-33.
12. Ежов А.А. Нейрокомпьютинг и его применение в экономике и бизнесе / А. А. Ежов, С. А. Шумский; М.: МИФИ. 1998. 224 с.
13. Craven M. W. Extracting tree-structured representations of trained networks / M.W. Craven, J.W. Shavlik // Advances in Neural Information Processing Systems. MIT Press, Cambridge, MA. 1996. Vol. 8. Pp. 24–30.
14. Барский А. Б. Основы построения реагирующих объектов для систем интеллектуального отображения // А .Б. Барский, Л. Б. Милютин, А. Е. Тимофеев // Компьютеры в учебном процессе, 2005. № 5. С. 81–90.
15. Сафронова Т.А., Живицкая, Е.Н. Логические нейронные сети для управления производственными запасами в условиях неопределенности. Режим доступа: <http://be5.biz/ekonomika1/r2012/2015.htm>.
16. Жилон Р. А. Применение логической нейронной сети к задаче классификации // Вестник КРАУНЦ. Физ.-мат. науки. –018. № 3(23). С. 180–183.
17. Димитриченко Д. П. Применение переменных логических функций и нейронных сетей в системах принятия решений // Вестник КРАУНЦ. Физ.-мат. Науки. 2016. № 4–1(16). С. 9–100.
18. Агеева У. О. Бионическое интеллектуальное протезирование конечностей и логические нейронные сети / У. О. Агеева, В. Г. Агеева, А. Б. Барский // Информационные технологии. 2016. Т. 22, № 5. С. 379–386.
19. Барский А. Б. Медицинские информационно-справочные системы на логических нейронных сетях / А. Б. Барский, А. А. Дмитриев, О. А. Барская // Информационные технологии. 2010. № 1. С. 1–32.
20. Дунин В. О. Разработка средств интеллектуального анализа и обработки медицинской информации / В. О. Дунин, В. А. Егоров // Современные информационные технологии. 2013. № 18. С. 173–178.
21. Борискин С.М. Логическая нейронная сеть как решение задачи авторизации для комплекса средств аутентификации. // Актуальные проблемы современной науки. 2011. № 1.

Сведения об авторах

Фридман Ольга Владимировна

к.т.н., старший научный сотрудник

e-mail: ofridman@iimm.ru

Olga V. Fridman

PhD, senior researcher

Ю.А. Олейник, А.А. Зуенко

Институт информатики и математического моделирования ФИЦ КНЦ РАН

АНАЛИЗ ВОЗМОЖНОСТЕЙ СОВМЕЩЕНИЯ ПАРАДИГМ ПРОГРАММИРОВАНИЯ В ОГРАНИЧЕНИЯХ И ОБЪЕКТНО-ОРИЕНТИРОВАННОГО ПРОГРАММИРОВАНИЯ*

Аннотация

В работе обосновывается принципиальная схожесть ограничений в задачах удовлетворения ограничений и отношений в реляционных базах данных. На основе этой схожести делается предположение о возможности переноса опыта интеграции объектного представления в реляционные базы данных на технологию программирования в ограничениях. Проводится обзор вариантов такой интеграции и анализ их применимости для решения задач удовлетворения ограничений.

Ключевые слова:

программирование в ограничениях, объектно-ориентированные базы данных, объектно-реляционные базы данных, постреляционные базы данных.

Y.A. Oleynik, A.A. Zouenko

POSSIBILITY ANALYSIS OF CONSTRAINT AND OBJECT-ORIENTED PROGRAMMING PARADIGMS COMBINING

Abstract

The paper substantiates similarity of constraints in the constraint satisfaction problems and relationships in relational databases. Based on this similarity, an assumption is made about the possibility of transferring the experience of integrating the object representation into relational databases to the constraint programming technology. A review of variants for such integration and an analysis of their applicability for solving constraint satisfaction problems is carried out.

Keywords:

constraint programming, object databases, object-relational databases, postrelational databases.

Введение

Программирование в ограничениях является довольно молодым направлением. Оно находит применение в решении комбинаторных задач, задач составления расписаний, распределения ресурсов и других направлениях, область которых продолжает расширяться. Задачи, описываемые и решаемые с помощью парадигмы программирования в ограничениях, называются задачами удовлетворения ограничений (Constraint Satisfaction Problem – CSP).

Ограничения можно также рассматривать, как отношения над множествами областей определения переменных задачи. Таким образом, решения задачи CSP — это такие значения переменных, которые делают все отношения истинными. При такой интерпретации становится очевидным родство CSP с реляционными базами данных (РБД), из которого вытекает ряд их общих недостатков.

* Работа выполнена в рамках темы НИР «Развитие методологии построения интеллектуальных мультипредметных систем информационной поддержки регионального развития на примере территорий Арктической зоны Российской Федерации (АЗРФ)» (№ 0226-2019-0036).

1. Парадигма программирования в ограничениях

В общем виде задача удовлетворения ограничений задается тремя множествами[1]: X – множество переменных задачи; D – множество областей определения переменных задачи; C – множество ограничений над переменными задачи. Под ограничением понимается любое соотношение между переменными предметной области. В качестве ограничений могут выступать арифметические выражения; логические формулы; таблицы; выражения, формулируемые на языке специализированных теорий.

Решением задачи CSP является полное присваивание, которое удовлетворяет всем ограничениям, и в некоторых случаях необходимо получить все решения. Иногда требуется найти такое решение, в котором значения переменных оптимизировали бы некоторый заданный функционал. CSP-задачи принадлежат классу NP-полных задач.

Технология программирования в ограничениях предоставляет мощные и гибкие методы, алгоритмы решения задач комбинаторного поиска. Особенности технологии:

- С точки зрения конечного пользователя задача CSP формулируется в декларативном виде, на языке близком к языку математики. Порядок задания ограничений несущественен.

- Любой алгоритм удовлетворения ограничений должен содержать две обязательных компоненты: а) компоненту, реализующую вывод (распространение); б) компоненту, реализующую поиск.

- Вывод (распространение) реализуется как целенаправленное сужение изначально заданных областей определения переменных.

- Эвристики, используемые в процедурах поиска, разрабатываются не под конкретную задачу, а являются универсальными.

- Благодаря архитектуре систем программирования в ограничениях, появляется возможность совместно обрабатывать количественные и качественные ограничения.

- Обеспечивается возможность сопровождать модели, открытые для оперативных модификаций. При добавлении/удалении из модели ограничений нет необходимости писать новые методы решения задачи.

В CSP условия задачи описываются декларативно, а алгоритмы поиска решений (распространение и ветвление) остаются скрыты от пользователя, аналогично в РБД декларативно описывается структура и наполнение базы данных, а алгоритмы поддержания согласованности и целостности данных скрыты в используемой СУБД. Очевидно, для реализации такого подхода, средства описания как РБД, так и CSP должны быть типизированы, а сами типы переменных (данных) и отношений над ними должны быть достаточно примитивны для их успешной автоматической интерпретации и обработки. Так, достаточно популярная библиотека программирования в ограничениях Choco[2] позволяет создавать переменные всего 4 различных типов: целый, логический, вещественный и множество целых, причем поддержка вещественного типа до сих пор неполная.

Поскольку примитивность типов является существенным ограничением для описания баз со сложными данными и отношениями над ними, был предпринят ряд шагов для расширения функционала РБД, часть из которых будет

рассмотрено ниже. Эти подходы видятся вполне применимыми и для расширения функционала средств решения задач CSP.

2. Предпосылки интеграции объектного представления в реляционные базы данных

Для обеспечения полноценной интеграции с языками программирования разработчики современных СУБД все более стремятся перенести парадигмы объектно-ориентированного программирования на технологию БД [3]. Таким образом, можно условно выделить объектно-ориентированный подход к представлению баз данных.

Как правило, реляционные СУБД (РСУБД) не позволяют добавлять новые типы данных, то есть набор типов данных в РСУБД заранее определен и фиксирован. Классические РСУБД не поддерживают композитных атрибутов (доменами таких атрибутов являются пользовательские типы данных, составленные из predetermined набора элементарных типов), несмотря на то, что они не противоречат реляционной модели.

Сильная сторона реляционных СУБД заключается в том, что они содержат встроенный язык запросов SQL, реализующий операции проекции и соединения отношений и предоставляющий необходимые средства для выполнения незапланированных запросов. Запросы к базе данных возвращают результаты в виде таблиц, которые тоже могут выступать как объект запросов.

К основным недостаткам РСУБД относят и то, что их семантическая составляющая развита слабо. Это, с одной стороны, затрудняет использование РСУБД в системах поддержки принятия решений и в предметных областях, имеющих сложно структурированные данные, где требуются дополнительные средства представления семантики данных. С другой стороны, программирование алгоритмов выборки данных на языке SQL порождает громоздкие конструкции.

Эти недостатки привели к появлению направления семантического моделирования и широкому использованию объектного подхода при организации хранения и обработки информации БД.

Объектно-ориентированный подход [4] был создан для решения задачи повышения уровня абстракции данных и стал фактическим стандартом разработки программного обеспечения. В объектно-ориентированных языках программирования (ООЯП) предметная область описывается в виде совокупности экземпляров различных типов, которые определяются программистом. ООЯП поддерживают три парадигмы: 1) инкапсуляция данных; 2) наследование; 3) полиморфизм.

Разработчики современных СУБД стремятся тем или иным способом реализовать объектные парадигмы, расширяя возможности базовых моделей данных. В то же время, для баз данных, использование в чистом виде объектно-ориентированного подхода зачастую может оказаться вредным, поскольку ограничивает возможность выполнения незапланированных запросов.

3. Направления развития объектного представления в базах данных

В зависимости от того, каким образом реализуются парадигмы объектно-ориентированного программирования, среди современных СУБД можно выделить три основных направления:

- объектно-ориентированные (ООСУБД);
- объектно-реляционные (ОРСУБД);
- постреляционные СУБД.

Объектно-ориентированные СУБД (ООСУБД) [3] предназначены для постоянного хранения объектов ООЯП и в них обеспечивается та или иная форма настройки по адресам. В качестве основы организации ООСУБД был принят стандарт ODMG (Object Data Management Group – консорциум разработчиков БД), описывающий способы хранения объектов в базах данных[5]. Основными элементами стандарта являются:

Объектная модель – основа стандарта, расширяет модель консорциума ODMG (Object Management Group) свойствами, необходимыми для взаимодействия с БД (связи, транзакции). Особенности объектной модели ODMG являются:

- наделение объектов такими свойствами, как атрибуты и связи;
- методы объектов (поведение);
- множественное наследование;
- идентификаторы объектов (ключи);
- определение таких совокупностей объектов как списки, наборы, массивы и т.д.;
- блокировка объектов и изоляция доступа;
- операции над базой данных.

Язык описания объектов (ODL – Object Definition Language) – средство определения схемы базы данных. ODL является расширением IDL (Interface Definition Language - язык описания интерфейсов) модели ODMG и предоставляет средства для определения объектных типов, их атрибутов, связей и методов. ODL создает слой абстрактных описаний. Схема базы данных становится независима как от языка программирования, так и от СУБД. ODL рассматривает только описание объектных типов данных, не вдаваясь в детали реализации их методов. Это позволяет переносить схему БД между различными ODMG-совместимыми СУБД и языками программирования.

Язык объектных запросов (OQL - Object Query Language) - SQL - подобный декларативный язык, который предоставляет эффективные средства для извлечения объектов из базы данных, включая высокоуровневые примитивы для наборов объектов и объектных структур. OQL-запросы могут вызываться из OO-языка, точно также из OQL-запросов могут делаться обращения к процедурам, написанным на OO-языке. OQL предоставляет средства обеспечения целостности объектов (вызов объектных методов и использование собственных операторов изменения данных).

Связывание с OO-языками. Стандарт связывания с C++, Smalltalk и Java определяет Object Manipulation Language (OML) — язык манипулирования объектами, который расширяет базовые OO-языки средствами манипулирования и хранения объектов. Также включаются OQL, средства навигации и поддержка транзакций. Каждый OO-язык имеет свой собственный OML, поэтому разработчик остается в одной языковой среде, ему нет необходимости разделять средства программирования и доступа к данным.

Продукты на основе этого стандарта не имели коммерческого успеха, поскольку требовали преобразования существующих данных в формат СУБД. При этом ООСУБД изначально интегрированы с ООЯП, поэтому в код

приложения нет необходимости встраивать дополнительные конструкции типа SQL. ООСУБД входят в состав компилятора, который обрабатывает исходный текст и автоматически создает в БД структуры для хранения объектов. Объект сохраняется в БД при вызове специального метода. Основным недостатком ООСУБД, который они унаследовали от иерархических СУБД, – невозможность осуществления незапланированных запросов к БД без изменения её структуры (добавления новых структур данных или методов), а значит, изменения и самой программы. Кроме того, если с одной и той же БД работают приложения, количество которых заранее не определено, то появление нового приложения требует перекомпиляции уже имеющихся.

Другим направлением являются объектно-реляционные СУБД, появившиеся в результате расширения функционала реляционных СУБД и языков обращения к ним [6]. ОРСУБД наряду с хранением реляционных данных, обеспечивают постоянное хранение объектов. Такие системы используют отображение “объектный класс – домен” и поддерживают создание пользовательских типов данных. Это позволяет создавать композитные атрибуты, применять парадигмы наследования и полиморфизма. Так, если поле таблицы имеет базовый тип, то в нем может содержаться также и объект любого дочернего типа. Методы типа описываются в его теле, что поддерживает инкапсуляцию данных. Методы могут описывать преобразования к другим типам, что обеспечивает поддержку приведения типов на уровне доменов. Множественное наследование в современных ОРСУБД не поддерживается, хотя не исключается в будущем.

Помимо объектных расширений, ОРСУБД поддерживают все реляционные операторы. Поэтому ОРСУБД свободны от недостатков ООСУБД и поддерживают следующие возможности:

- незапланированные запросы;
- каскадное удаление (ссылочная целостность);
- поиск указателей и идентификаторов объектов скрыт от пользователя;
- использование преимуществ инкапсуляции для внутренней структуры отношений.

Поскольку ОРСУБД, по сути, реализуют реляционную модель с дополнительной поддержкой пользовательских типов на уровне доменов, то они наследуют все основные недостатки этой модели, связанные с семантикой.

Третьим направлением к объединению возможностей РБД и объектного представления данных является создание объектно-ориентированной надстройки над уже существующими реляционными СУБД. Получившиеся в результате СУБД были названы постреляционными. Этот вариант был предложен Майклом Стоунбрейкером [7] и реализован под его началом в университете Беркли в виде СУБД Postgres[8]. Постреляционные СУБД – результат эволюции РСУБД. Они позволяют создавать пользовательские типы данных (с некоторыми ограничениями) и таблицы с композитными столбцами. В качестве домена столбца может выступать таблица, а точнее – тип данных, соответствующий структуре таблицы и создаваемый одновременно с таблицей. Композитные типы создаются на основе некоторого базового множества элементарных типов, априорно поддерживаемых СУБД. Современные постреляционные СУБД не позволяют напрямую (средствами языка SQL или его объектных расширений)

создавать массивы для хранения произвольных структур, несмотря на то, что сами пользовательские типы данных определяются без особого труда. Создание массивов пользовательских типов средствами таких СУБД требует описания на одном из объектно-ориентированных языков самого массива, а также операторов преобразования массива в строку и обратно. Постреляционные СУБД позволяют хранить функции обработки данных, использующие в качестве параметров определяемые пользователем типы. В PostgreSQL поддерживается возможность множественного наследования таблиц, которая вызывает ряд проблем (например, уникальность значений ключевых полей таблиц-наследниц). Проблемы возникают из-за неадекватности отображения «объектный класс – таблица», характерного для постреляционных СУБД [3]. Реляционную таблицу нельзя считать аналогом объектного класса, несмотря на то, что она тоже описывается набором атрибутов, по следующим причинам:

- Таблица, в отличие от объектного класса, является коллекцией однотипных объектов. В связи с этим, таблица должна иметь ключевой атрибут, который не характерен для объектного класса.

- Таблица не имеет методов. Конечно, можно описать функции, работающие с данной таблицей. Но между вызовами “table.method(3)” и “procedure(table,3)” нельзя ставить знак равенства, поскольку разрушается инкапсуляция объектного класса – ключевая парадигма объектно-ориентированного программирования.

Другими словами, более корректно отображение «динамический массив – таблица». Также постреляционные СУБД не поддерживают полиморфного поведения объектов, лежащих в одной иерархии наследования. Более того, если таблица содержит поле базового типа, то добавление объекта дочернего типа приводит к усечению последнего.

Заключение

Программирование в ограничениях является новым направлением, опирающимся на реляционное представление информации. Данное направление, по мнению авторов, не получило еще должного развития с точки зрения интеграции с объектно-ориентированным подходом. Принципиальное отличие технологии программирования ограничений от технологий управления данными состоит в том, первая предназначена для организации вывода на знаниях, а не манипулирования структурами данных. Отсутствие таких препятствий, как необходимость обеспечения одновременного доступа и целостности данных создает значительные предпосылки для упомянутой интеграции.

В статье рассмотрены различные способы интеграции объектного и реляционного подхода в СУБД. Наиболее перспективным из представленных и наименее противоречивым видится направление ОРСУБД. Учитывая успешный опыт применения ОРСУБД, по мнению авторов, именно отображение «объектный класс – домен» должно лежать и в основе разрабатываемых способов интеграции парадигмы программирования в ограничениях с объектно-ориентированным подходом.

Дальнейшие направления исследований в этой области предположительно будут связаны с разработкой контекстно-ориентированных механизмов формирования и активации систем ограничений на основе объектно-

ориентированного представления информации о предметной области, что обеспечит ускорение комбинаторного поиска в слабо формализованных предметных областях по сравнению с существующими средствами программирования ограничениях, поддерживающими только простые типы переменных.

Литература

1. Russel S., Norvig P. Artificial Intelligence: A Modern Approach. 3rd ed. / S. Russel, P. Norvig. Pearson Education, 2010. 1152 p.
2. Jussien N., Rochart G., Lorca X. Choco documentation. Режим доступа: <http://www.choco-solver.org>.
3. Дейт К. Дж. Введение в системы баз данных, 6-е издание: Пер. с англ. К.; М.; СПб.: Издательский дом «Вильямс», 2000. – 848 с.
4. Буч Г. Объектно-ориентированный анализ и проектирование: с примерами приложений на C++. / Г. Буч. «Издательство Бином», «Невский диалект», 1998. 560 с.
5. Калиниченко Л.А. Стандарт систем управления объектными базами данных ODMG-93: краткий обзор и оценка состояния. / Л.А. Калиниченко // СУБД. – 1996. № 1. С. 102–109.
6. Бич Д. К объектным базам данных. // Открытые системы. 1994. № 4. С. 50–55.
7. Стоунбрейкер М. Объектно-реляционные системы баз данных // Открытые системы. 1994. № 4. С. 43–49.
8. Булах Е.В. Средства доступа к базам данных в Internet и свободно доступная СУБД POSTGRES95 // СУБД. 1997. № 2.

Сведения об авторах

Олейник Юрий Андреевич

младший научный сотрудник

e-mail: yoleynik@iimm.ru

Yurii A. Oleynik

junior research fellow

Зуенко Александр Анатольевич

к.т.н., ведущий научный сотрудник

e-mail: zuenko@iimm.ru

Alexander A. Zouenko

Ph.D. (Tech. Sci.), leading researcher

А. А. Зуенко, О. Н. Зуенко

Институт информатики и математического моделирования ФИЦ КНЦ РАН

ПРИМЕНЕНИЕ МЕТОДОВ ДЕКЛАРАТИВНОГО ПРОГРАММИРОВАНИЯ В ЗАДАЧАХ CONSTRAINED CLUSTERING*

Аннотация

Анализ методов решения задач кластеризации с частичным привлечением учителя (Constrained Clustering или Semi Supervised Clustering) показал, что большинство из них являются приближенными. Точные методы технологии программирования в ограничениях позволяют решать задачи комбинаторного поиска, характеризующиеся большой размерностью. Однако, данные средства всё ещё редко применяются в задачах Constrained Clustering. Сделан вывод о перспективности разработки новых методов удовлетворения ограничений для различных постановок задач Constrained Clustering.

Ключевые слова:

Кластерный анализ, кластеризация с частичным применением учителя, декларативное программирование, программирование в ограничениях.

A. A.Zuenko, O. N.Zuenko

APPLICATION OF DECLARATIVE PROGRAMMING METHODS IN THE CONSTRAINED CLUSTERING PROBLEMS

Abstract

An analysis of the methods of solving Semi Supervised Clustering(Constrained Clustering) problems showed that most of them are approximate. Exact methods of constraint programming technology allow solving combinatorial search problems of large dimension. However, these tools are still rarely used in Constrained Clustering problems. The conclusion is drawn on the prospects of developing new constraint satisfaction methods for various statements of Constrained Clustering problem.

Keywords:

Cluster analysis, semi supervised clustering, declarative programming, constraint programming.

Введение

Задача кластеризации относится к задачам комбинаторного поиска. Коснемся несколько более подробно задач комбинаторного поиска.

Рассмотрим задачу подбора правильной комбинации для кодового замка. Она решается только прямым последовательным перебором всех возможных комбинаций. Это один из редких примеров задач комбинаторного поиска, для которых сложно предложить какой-либо метод ускорения. Даже уже при 1000 окошках для ввода цифр она порождает 10^{1000} комбинаций, что под силу перебрать не каждому компьютеру. Если же предположить, что нам об этих комбинациях что-либо известно, например, комбинация — это год рождения одного из членов семьи, или известны некоторые цифры комбинации или свойства цифр, то задача подбора существенно упрощается.

* Работа выполнена в рамках темы НИР «Развитие методологии построения интеллектуальных мультипредметных систем информационной поддержки регионального развития на примере территорий Арктической зоны Российской Федерации (АЗРФ)» (№ 0226-2019-0036).

Анализ знаний из предметной области способен ускорить процесс получения решения задач комбинаторного поиска. Этот вывод кажется банальным, но если рассмотреть классическую задачу кластерного анализа, то там учет фоновых знаний не предусмотрен.

Исходными данными в задаче классического кластерного анализа является матрица расстояний между кластеризуемыми объектами, которая получается из объекто-признаковой таблицы. Далее свойства объектов не анализируются. На основании введенного расстояния объекты раскладываются (разбиваются) на кластеры, в один кластер попадают наиболее близкие друг другу объекты. Однако, часто отнесение двух объектов в один кластер только лишь на основе метрики является семантически некорректной операцией.

Таким образом, основным недостатком большинства существующих методов кластеризации является невозможность учитывать пользовательские ограничения на то, какие объекты обязательно должны/не должны попадать в один кластер. В статье рассматривается подход к кластеризации, когда при отнесении объектов к одному или различным кластерам анализируются не только расстояния между объектами, но и значения их признаков. Кроме того, чтобы лучше смоделировать задачу и снизить ее сложность, могут быть добавлены пользовательские ограничения. В этом случае задача кластеризации становится задачей *Constrained Clustering*, целью которой является получение кластеров, удовлетворяющих не только критерию кластеризации, но и пользовательским ограничениям.

В статье представлен обзор методов *Constrained Clustering*, а также необходимые сведения из классического кластерного анализа.

Задача кластеризации

Рассмотрим набор данных из n объектов $O = \{o_1; \dots; o_n\}$. Каждый объект описывается значениями p атрибутов, также называемых переменными. Обозначим o_{ij} значение j -го атрибута объекта o_i . Большинство алгоритмов кластеризации опираются не на анализ свойств объектов, а используют только таблицы расстояний между объектами.

Пусть $C_1; C_2; \dots; C_k$ – подмножества множества O . Тогда $\{C_1; C_2; \dots; C_k\}$ является разбиением O на k кластеров, если для всех $c \in \{1; 2; \dots; k\}$ выполняется:

1. $C_c \neq \emptyset$;
2. $\bigcup_c C_c = O$;
3. $c \neq c', C_c \cap C_{c'} = \emptyset$.

Предположим, что имеется мера различия между любыми двумя объектами $o_i; o_j \in O$, которую обозначим d_{ij} . Мера различия d_{ij} обычно рассчитывается с использованием метрики расстояния, определенной в пространстве атрибутов.

В кластерном анализе наиболее популярной метрикой расстояния является Евклидово расстояние и квадрат Евклидова расстояния [1]. Пусть дано два объекта $o_i; o_j$, тогда евклидово расстояние между ними определяется как:

$$d_{ij} = \sqrt{(o_{i1} - o_{j1})^2 + (o_{i2} - o_{j2})^2 + \dots + (o_{ip} - o_{jp})^2}.$$

Квадрат Евклидова расстояния определяется как:

$$d_{ij}^2 = (o_{i1} - o_{j1})^2 + (o_{i2} - o_{j2})^2 + \dots + (o_{ip} - o_{jp})^2.$$

Другой популярной мерой расстояния является Манхэттенское расстояние [1], определяемое как:

$$d_{ij} = |o_{i1} - o_{j1}| + |o_{i2} - o_{j2}| + \dots + |o_{ip} - o_{jp}|.$$

Аналогично, обозначим меру сходства между двумя объектами o_i и o_j как s_{ij} .

1. Мера сходства широко используется в спектральной кластеризации. Как правило, рассчитывается Гауссова функция различия [2]:

$$s_{ij} = \exp\left(-\frac{d_{ij}^2}{2\sigma_i^2}\right),$$

где σ_i — параметр.

Существуют и другие меры сходства, например нормализованная корреляция Пирсона (*normalized Pearson correlation*), мера Жакара (*Jaccard measure*), мера коэффициента игральной кости (*dice coefficient measure*) и др. [1].

Задача кластеризации состоит в разбиении совокупности объектов на классы (группы, кластеры) таким образом, чтобы выполнялось некоторое условие/критерий оптимизации.

При кластеризации формируемые классы должны удовлетворять двум свойствам:

1. Однородность (*homogeneity*).
2. Хорошая разделимость/различимость (*well separated*).

Эти требования обычно выражаются посредством критерия оптимизации. Тогда задача кластеризации может быть определена как проблема оптимизации, состоящая в поиске такого разбиения объектов, которое оптимизирует заданный критерий. Говоря об однородности, предъявляются некоторые требования к объектам внутри одного кластера. Например, может быть задан критерий, состоящий в том, что оптимальным является разбиение с минимальным диаметром (относительно других возможных разбиений). Напомним, что диаметр разбиения — это максимальный диаметр среди всех диаметров кластеров, принадлежащих данному разбиению. Диаметр кластера — это расстояние между наиболее удаленными точками кластера.

Вводя критерии разделимости, предъявляются определенные требования к объектам, которые попадают в различные кластеры. В качестве примера подобного критерия может служить критерий, состоящий в том, что оптимальным является разбиение (относительно других возможных разбиений) с максимальным значением функции *split*. Для конкретного разбиения функция *split* показывает минимально возможное расстояние между двумя различными кластерами данного разбиения.

Различные методы кластеризации используют различные метрики и различные критерии. Не существует метода кластеризации, одинаково хорошо работающего для всех типов данных

Классификация методов кластеризации

Общепринятой классификации методов кластеризации не существует. Приведем следующие группы методов:

1. Вероятностные методы. К ним относятся методы *k*-средних (*k-means*, *k-median*, *k-medoid*).
2. Методы на основе искусственного интеллекта. В частности, методы в рамках нейросетевого подхода.
3. Теоретико-графовые методы. Например, метод *k*-средних с критерием диаметра.
4. Логические методы. В частности, деревья решений на основе представления задачи в виде КНФ (*BooleanSATisfability*).
5. Иерархические методы.
6. Все другие методы, не вошедшие в классификацию.

Чаще методы кластеризации делят на иерархические и неиерархические. Неиерархические (упорядоченные – *partitioning*) отличаются тем, что заранее известно количество кластеров. В иерархических (*hierarchical*) количество кластеров заранее неизвестно и является частью решения. Часто иерархическая кластеризация предшествует неиерархической. Результатом иерархической кластеризации является дендрограмма – древовидная диаграмма.

Для остановки кластеризации существует определенный критерий. Например, когда внутри кластерное расстояние становится больше, чем межкластерное.

Иерархические методы делятся на дивизимные (сверху вниз по дереву) и агломеративные (снизу вверх по дереву). Примеры: метод ближайшего соседа, метод наиболее удаленного соседа.

Также методы кластеризации делят на точные и приближенные.

Точные и приближенные методы кластеризации

В искусственном интеллекте задачи комбинаторного поиска решаются двумя группами методов.

1. Методы систематического (конструктивного) поиска — позволяет найти глобальный экстремум.
2. Методы локального поиска.

Локальный поиск не обладает свойством полноты. Систематический поиск обладает свойством полноты, но позволяет исследовать существенно меньшие пространства поиска.

Точные методы обеспечивают поиск глобального оптимума. Глобальный оптимум – это точное решение задачи оптимизации, которое удовлетворяет критерию оптимизации. Приближенные (эвристические) методы, также называются «жадные алгоритмы», в качестве решения могут вернуть локальный оптимум. Алгоритмы локального поиска, которые опираются на понятие соседнего состояния, обладают малым количеством памяти для фиксации

пройденных состояний. Их существенным недостатком является то, что они сильно зависят от выбора начального состояния.

К неиерархическим эвристическим алгоритмам можно отнести методы *k-means* [3], *k-median*, *k-medoid*, метод спектральной кластеризации, метод *FTP* (*Furthest Point First*) [4].

Далее перечислим наиболее популярные точные методы решения задач неиерархической кластеризации. Это, прежде всего, теоретико-графовые методы и метод ветвей и границ. Теоретико-графовые методы сводят задачу кластеризации к задаче раскраски графа. Наиболее известный алгоритм, относящийся к данному классу, разработан для критерия минимального диаметра [5]

Bench&bound – метод ветвей и границ. Это целый класс методов, которые отличаются критерием кластеризации. В зависимости от этого критерия, выстраивается процедура обхода дерева поиска, а именно, строится своя функция оценки (стоимостная функция), позволяющая отсекалть заранее неперспективные ветви дерева поиска [6, 7].

Задача *Constrained Clustering*

Задача классического кластерного анализа — это задача разбиения множества объектов на классы, когда какая-либо априорная информация о принадлежности объектов этим классам отсутствует. Классификация — это задача отнесения предъявляемых объектов к заранее указанным классам, сопровождаемая этапом предварительного обучения. На этапе обучения для некоторых объектов (объектов обучающей выборки) указывается, каким классам они принадлежат. Выводится правило классификации. Задача *Constrained Clustering* использует некоторые фоновые знания из предметной области. Количество классов и сами классы неизвестны, но для некоторых пар объектов известно, например, что они попадают или не попадают в один кластер. Поэтому задача *Constrained Clustering* также называется задачей кластеризации с частичным привлечением учителя (*semi-supervised clustering*).

Основная идея состоит в том, что при кластеризации хорошо бы использовать фоновые или базовые знания из предметной области. Эти базовые знания предложено представлять в виде пользовательских ограничений. В таких задачах выделяют два вида пользовательских ограничений:

1. Ограничения на кластеры (*cluster-level constraints*), указывающие требования к кластерам.
2. Ограничения на объекты кластеров (*instance-level constraints*), уточняющие требования к парам конкретных объектов.

Примерами ограничений на кластеры является минимальная/максимальная населенность кластера, средняя населенность кластера (отношение минимума к максимуму) – это ограничение говорит о том, что кластеры должны быть примерно одного размера.

Наиболее интересны ограничения на объекты кластеров. Выделяют два вида таких ограничений [8]: *must-link* и *cannot-link*. Данные ограничения также называются попарными. Ограничение *must-link* означает, что два объекта должны быть в одном кластере. А *cannot-link* означает, что два объекта должны быть в разных кластерах.

Ограничение *must-link* между двумя объектами o_i и o_j , обозначаемое как $ML(o_i; o_j)$, предписывает, что оба объекта o_i и o_j должны быть в одном кластере. Напротив, ограничение *cannot-link* между объектами o_i и o_j , обозначаемое как $CL(o_i; o_j)$, предписывает, что эти два объекта не должны находиться в одном кластере.

Согласно [9], ограничения на объекты кластеров имеют следующие свойства:

1. Ограничения *must-link* транзитивны: Пусть CC_a и CC_b являются связанными компонентами (подграфами, полностью связанными посредством ограничений *must-link*), пусть o_i и o_j будут объектами в CC_a и CC_b соответственно, тогда:

$$ML(o_i; o_j); o_i \in CC_a; o_j \in CC_b \Rightarrow ML(o_x; o_y); \forall o_x; o_y : o_x \in CC_a; o_y \in CC_b.$$

2. Для ограничений *cannot-link* справедливо следующее: пусть CC_a и CC_b являются связанными компонентами (подграфами, полностью связанными посредством ограничений *must-link*), пусть o_i и o_j будут объектами в CC_a и CC_b соответственно, тогда:

$$CL(o_i; o_j); o_i \in CC_a; o_j \in CC_b \Rightarrow CL(o_x; o_y); \forall o_x; o_y : o_x \in CC_a; o_y \in CC_b.$$

Идея о введении ограничений *must-link* и *cannot-link* может показаться простой, но на самом деле эти ограничения являются мощным инструментом для многих приложений. Увеличение количества ограничений на пары объектов может существенно улучшить точность результата кластеризации.

Ограничения *must-link* и *cannot-link* также могут использоваться для выражения других пользовательских ограничений.

В работе [10] введены ограничения δ -*constraint* и ε -*constraint*, которые могут быть обработаны совместно с ограничениями на пары объектов:

- Ограничение δ -*constraint* (также называемое *minimum split constraint*) выражает, что расстояние между любой парой точек, которые находятся в двух разных кластерах, должно быть не меньше, чем значение δ . Это ограничение может быть представлено в виде конъюнкции ограничений *must-link* для всех пар объектов с расстоянием меньше, чем δ .

- Ограничение ε -*constraint* для любого кластера C_c , содержащего более одного объекта, для каждого объекта $o_i \in C_c$ должен существовать другой объект $o_j \in C_c$, такой, что расстояние между o_i и o_j не превышает $d_{ij} : d_{ij} \leq \varepsilon$. В работе показано, что это ограничение эквивалентно дизъюнкции ограничений *must-link*. Для каждой точки o_i вычисляется множество χ объектов o_j таких, что: $d_{ij} \leq \varepsilon$. Ограничение ε -*constraint* может быть представлено как дизъюнкция ограничений *must-link* между объектом o_i и точками множества χ .

Другим ограничением, которое относится к ограничениям на объекты кластеров, является ограничение на максимальный диаметр (максимально возможное расстояние между объектами одного кластера). Это ограничение

задает верхнюю границу γ на диаметр кластеров, таким образом, расстояние между любой парой точек, принадлежащих одному и тому же кластеру, должно быть не более значения γ . Это ограничение может рассматриваться как конъюнкция ограничений *cannot-link* между всеми парами объектов с расстоянием, превышающим γ .

Методы решения задач *Constrained Clustering*

За последнее десятилетие было выполнено много работ, направленных на расширение классических алгоритмов для обработки ограничений *must-link* и *cannot-link*, например, расширение для *COBWEB* [11], *k-means* [12, 13], иерархической кластеризации [14] и т.д.

Методы решения задач *Constrained Clustering* получаются при помощи модификации базовых методов кластеризации. Можно выделить три вида таких модификаций.

1. Модификация формул расстояния.
2. Модификация целевой функции, которая должна учитывать штрафы за нарушение ограничений *must-link* и *cannot-link*.
3. Модификация стратегии поиска.

Однако, общего решения, чтобы расширить традиционные алгоритмы для различных типов ограничений, не существует.

Далее рассматривается еще один подход, который может быть применен для решения задач *Constrained Clustering*. Данный подход может являться основой для разработки высокоэффективных методов систематического поиска.

Методы декларативного программирования в задачах *Constrained Clustering*

Существующие к данному моменту методы *Constrained Clustering* являются развитием методов классической кластеризации и наследуют их недостатки. Большинство данных методов это методы локального поиска, т.е. они позволяют часто найти лишь локальный экстремум. Дело в том, что пространство поиска в задачах кластеризации достаточно большое, и до недавнего времени не существовало технологий систематического обхода пространства поиска такой размерности. В 90-х годах появились мощные *SAT solvers*, а затем и технология программирования в ограничениях. Эта технология позволяет систематически исследовать куда большие пространства поиска, чем ранее.

В связи с этим, в последнее время возрос интерес к разработке декларативного подхода для *Data Mining* [15], в том числе задачи *Constrained Clustering*. Декларативный подход, возможно, менее эффективен, нежели классические алгоритмы, для специфических задач, зато он более гибок и носит более общий характер, а также существенно упрощает интеграцию новых знаний в задачу. Декларативный подход включает методы программирования в ограничениях, методы сведения задачи к задаче пропозициональной выполнимости (*Boolean satisfiability*), методы целочисленного линейного программирования, методы логического программирования. Кроме того, именно в рамках данного подхода стало возможно задачу *Constrained Clustering* решать при помощи конструктивного (систематического) перебора пространства поиска.

В [16] предлагается подход в рамках программирования в ограничениях для извлечения набора *k*-паттернов. В [17] представлен язык, основанный на

ограничениях, для выражения запросов, чтобы обнаруживать паттерны в *Data Mining*. Подход на основе сведения задачи *Constrained Clustering* к задаче выполнимости булевой функции (*SAT*) предложен в [18] для кластерного анализа с 2 кластерами. Данный подход объединяет различные виды пользовательских ограничений: *must-link*, *cannot-link*, а также ограничения диаметра и разбиения. В статье [19] рассматривается подход к *Constrained Clustering*, основанный на целочисленном линейном программировании. В работе [20] разрабатывается точный подход для задачи *Constrained Clustering* с критерием минимизации внутрикластерной суммы квадратов, основанный на целочисленном линейном программировании.

Среди средств оптимизации общего назначения, программирование в ограничениях является мощной парадигмой для решения задач комбинаторного поиска, которая основывается на широком спектре методов из искусственного интеллекта, информационных технологий и исследования операций. Основная идея программирования в ограничениях состоит в том, чтобы решать задачи путем наложения ограничений, которые должны быть удовлетворены при помощи решения. Главные принципы программирования в ограничениях это: 1) пользователи формулируют задачу декларативно, как задачу удовлетворения ограничений; 2) *solver* ищет решение путем удовлетворения ограничений и поиска.

Программирование в ограничениях также реализует идею декларативного программирования.

С точки зрения теории методы программирования в ограничениях можно разделить на три основные группы:

1. Методы распространения ограничений – суть этих методов состоит в уменьшении пространства поиска за полиномиальное время.

Когда алгоритмы распространения ограничений останавливаются, достигнув некоторой неподвижной точки, то в дело вступают алгоритмы второй группы:

2. Алгоритмы поиска – прежде всего, это алгоритмы поиска «в глубину с возвратами», которые используют эвристики при разборе случаев.

3. Алгоритмы структурной декомпозиции задач удовлетворения ограничений. Методы этой группы предназначены для разбиения задачи либо на независимые, либо на слабо связанные задачи. Цель разбиения состоит в том, чтобы решать каждую из подзадач в отдельности, а затем комбинировать полученные решения. Чаще всего граф задачи преобразуют к дереву. Для деревьев есть доказанный быстродействующий алгоритм распространения ограничений.

Программирование в ограничениях могло бы стать полезным инструментом для задач кластерного анализа с ограничениями, поскольку данный подход обладает такими преимуществами как: декларативность, которая дает возможность добавлять новые ограничения, и возможность находить оптимальное решение, удовлетворяющее всем ограничениям (если оно существует). Однако, по мнению авторов, данный подход не получил пока должного развития в области разработки высокоэффективных конструктивных методов решения задач *Constrained Clustering*.

Заключение

Анализ методов решения задач *Constrained Clustering* показал, что, как и в случае задачи классического кластерного анализа, большинство из них являются приближенными, а не точными. Данные методы не предназначены для систематического исследования пространства поиска и, в общем случае, не позволяют найти глобальный оптимум при рассмотрении задачи *Constrained Clustering* как задачи комбинаторной оптимизации. Появление мощных *SAT-solvers* и технологии программирования в ограничениях позволило решать многие интересные практически значимые задачи комбинаторного поиска, характеризующиеся большой размерностью, с помощью точных методов (методов систематического поиска). Однако, проведенный анализ показал, что данные средства декларативного программирования всё ещё редко применяются в задачах *Data Mining*, в целом, и в задачах *Constrained Clustering*, в частности. В связи с изложенным, представляется перспективной разработка новых методов удовлетворения ограничений для различных постановок задач *Constrained Clustering*.

Литература

1. Han J. *Data Mining: Concepts and Techniques*. 3rd Edition / J. Han, M. Kamber, J. Pei; Morgan Kaufmann Publishers Inc. - San Francisco, CA, USA, 2011. 703 p.
2. Luxburg U. A Tutorial on Spectral Clustering // *Statistics and Computing*. 2007. Vol. 17, № 4. pp. 395–416.
3. MacQueen J.B. Some Methods for Classification and Analysis of MultiVariate Observations // In: L. M. Le Cam and J. Neyman, editeurs. *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press. 1967. V. 1. pp. 281–297.
4. Gonzalez T. Clustering to Minimize the Maximum Intercluster Distance // *Theoretical Computer Science*. 1985. V. 38. pp. 293–306.
5. Hansen P. Complete-Link Cluster Analysis by Graph Coloring. / P. Hansen, M. Delattre // *Journal of the American Statistical Association*. 1978. V. 73, № 362. pp. 397–403.
6. Brusco M. J. An Enhanced Branch-and-Bound Algorithm for a Partitioning Problem // *British Journal of Mathematical and Statistical Psychology*. 2010. Vol. 56, № 1. pp. 83–92.
7. Brusco M. J. A Repetitive Branch-and-Bound Procedure for Minimum Withincluster Sum of Squares Partitioning // *Psychometrika*. 2006. Vol. 71, № 2. pp. 347–363.
8. Wagstaff K. Clustering with Instance-Level Constraints / K. Wagstaff, C. Cardie // In *Proceedings of the 17th International Conference on Machine Learning*, pages. 2000. pp. 1103–1110.
9. Davidson I. A Survey of Clustering with Instance Level Constraints / I. Davidson, S. Basu // *ACM Transactions on Knowledge Discovery from Data*. 2007. V. 1. pp. 1–41.
10. Davidson I. Clustering with Constraints: Feasibility Issues and the k-Means Algorithm / I. Davidson, S.S. Ravi // In *Proceedings of the 5th SIAM International Conference on Data Mining*. 2005. pp. 138–149.

11. Wagstaff K. Clustering with Instance-Level Constraints / K. Wagstaff, C. Cardie // In Proceedings of the 17th International Conference on Machine Learning. 2000. pp. 1103–1110.
12. Wagstaff K. Constrained K-means Clustering with Background Knowledge / K. Wagstaff, et al. // In Proceedings of the 18th International Conference on Machine Learning. 2001. pp. 577–584.
13. Bilenko M. Integrating Constraints and Metric Learning in Semi-Supervised Clustering. / M. Bilenko, S. Basu, R.J. Mooney // In Proceedings of the 21st International Conference on Machine Learning. 2004. pp. 11–18.
14. Davidson I. Agglomerative Hierarchical Clustering with Constraints: Theoretical and Empirical Results. / I. Davidson, S. S. Ravi // In Proceedings of the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases. 2005. pp. 59–70.
15. Raedt L. D. Constraint Programming Meets Machine Learning and Data Mining / L. D. Raedt, et al. // Dagstuhl Seminar 11201, Dagstuhl Reports. 2001. V. 1, № 5. pp. 61–83.
16. Guns T. k-Pattern set mining under constraints / T. Guns, S. Nijssen, L.D. Raedt // IEEE Transactions on Knowledge and Data Engineering. 2013. V. 25, № 2. P. 402–418
17. Métivier J.-P. Constrained Clustering Using SAT. / J.-P. Métivier, et al. // In Proceedings of the 11th International Symposium on Advances in Intelligent Data Analysis. 2012. pp. 207–218.
18. Davidson I. A SAT-based Framework for Efficient Constrained Clustering. / I. Davidson, S. S. Ravi, L. Shamis // In: Proceedings of the 10th SIAM International Conference on Data Mining. 2010. pp. 94–105.
19. Mueller M., Kramer S. Integer Linear Programming Models for Constrained Clustering // In Proceedings of the 13th International Conference on Discovery Science. 2010. pp. 159–173.
20. Babaki B. Constrained Clustering using Column Generation. / B. Babaki, T. Guns, S. Nijssen // In: Proceedings of the 11th International Conference on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems. 2014. pp. 438–454.

Сведения об авторах

Зуенко Александр Анатольевич

к.т.н., в.н.с.

e-mail: zuenko@iimm.ru

Alexander A. Zouenko

Ph.D. (Tech. Sci.), leading researcher

Зуенко Ольга Николаевна

аспирант, стажер-исследователь

e-mail: ozuenko@iimm.ru

Olga N. Zouenko

postgraduate, trainee researcher

Р. А. Македонов, А. А. Зуенко

Институт информатики и математического моделирования ФИЦ КНЦ РАН

ДЕКЛАРАТИВНЫЙ ПОДХОД К ПЛАНИРОВАНИЮ ЗАКУПОК МАЛЫМ ПРЕДПРИЯТИЕМ*

Аннотация

В работе предлагается применять технологию программирования в ограничениях для решения задачи планирования закупок малым предприятием. Выбор в пользу данной технологии обусловлен тем, что часто классические числовые математические модели не в состоянии адекватно описать задачи предметной области, предполагающие наличие гибких ограничений, то есть разнообразных логических и аналитических условий на поставку, сроки и стоимость доставки, хранение и продажу товара.

Ключевые слова:

планирование закупок, ABC-анализ, XYZ-анализ, программирование в ограничениях, задача удовлетворения ограничений, комбинаторный поиск, комбинаторная оптимизация.

R. A. Makedonov, A. A. Zuenko

DECLARATIVE APPROACH TO PROCUREMENT PLANNING BY A SMALL ENTERPRISE

Abstract

The paper proposes to apply constraint programming technology to solve the problem of procurement planning by a small enterprise. The choice in favor of this technology is due to the fact that often classical numerical mathematical models are not able to adequately describe the subject domain tasks, which require flexible constraints, that is, a variety of logical and analytical conditions for the delivery, delivery time and cost, storage and sale of goods.

Keywords:

procurement planning, ABC-analysis, XYZ-analysis, constraint programming, constraint satisfaction problem, combinatorial search, combinatorial optimization.

Введение

Одной из важных задач, стоящих перед малым и средним бизнесом, является прогнозирование спроса и формирование предложения, способного удовлетворить потребности рынка в условиях ограничений по финансам и времени. Соответственно, возникает потребность в разработке программного обеспечения, нацеленного на эффективное решение подобных задач. Сложность заключается в том, что часто классические числовые математические модели не в состоянии адекватно описать задачи предметной области, что обусловлено неопределенностью ситуации, в которой приходится принимать решение. При моделировании практических задач требуется совместно обрабатывать качественные и количественные зависимости. Качественные зависимости

* Работа выполнена в рамках темы НИР «Развитие методологии построения интеллектуальных мультипредметных систем информационной поддержки регионального развития на примере территорий Арктической зоны Российской Федерации (АЗРФ)» (№ 0226-2019-0036).

(логические формулы, продукционные правила и т.д.) используется для формализации различных аспектов неопределенности.

Цель исследований, представленных в с, состоит в том, чтобы оценить применимость технологии программирования в ограничениях, которая является мощным инструментом решения сложных задач комбинаторного поиска и комбинаторной оптимизации, для задачи планирования закупок малым предприятием в условиях ограничений на финансы.

Исходные данные для задачи берутся из результатов совмещенного *ABC-XYZ*-анализа. *ABC*-анализ – это инструмент, который позволяет определить рейтинг товаров по получаемой прибыли (как один из возможных критериев) [1, 2]. Он основан на правиле Парето: согласно которому 20% ассортиментных позиций обеспечивает 80% прибыли. Таким образом, весь ассортимент торгового предприятия в результате *ABC*-анализа можно разделить на группы по степени важности: группа *A* – очень важные товары, которые приносят наибольшую прибыль; группа *B* – товары средней степени важности; группа *C* – наименее важные товары, это претенденты на исключение из ассортимента.

XYZ-анализ – это инструмент, позволяющий разделить продукцию по степени стабильности продаж и уровню колебаний потребления [2]. Анализ заключается в расчете каждой товарной позиции колебания расхода (коэффициента вариации). Этот коэффициент показывает отклонение расхода от среднего значения и выражается в процентах. Результатом *XYZ*-анализа является группировка товаров по трем категориям, исходя из стабильности их поведения:

- категория *X*, в которую попадают товары с колебанием продаж от 5 % до 15 % (условно), характеризующиеся стабильной величиной потребления и высокой степенью прогнозирования;
- категория *Y*, товары с колебанием продаж от 15 % до 50 %. Это товары с сезонными колебаниями и средними возможностями их прогнозирования;
- категория *Z*, товары с колебанием продаж от 50 % и выше. Это товары с нерегулярным потреблением и непредсказуемыми колебаниями.

При совмещенном *ABC-XYZ*-анализе определяется девять групп товаров (*AX, BX, CX, ..., CZ*), из которых наибольший интерес вызывают товары групп *AX* и *BX*, как наиболее прибыльные и прогнозируемые.

Исходными данными для задачи является прогнозируемое количество продаж продукции, получаемое из совмещенного *ABC-XYZ*-анализа и заданные границы колебаний продаж. Другой особенностью задачи является наличие гибких ограничений, позволяющих отразить разнообразные логические и аналитические условия на поставку, сроки и стоимость доставки, хранение и продажу товара.

Программирование в ограничениях как основа эффективной совместной обработки количественных и качественных соотношений

В качестве методологической основы исследований выбрана технология программирования в ограничениях. Для применения предлагаемой технологии необходимо, чтобы прикладная задача была сформулирована в виде задачи удовлетворения ограничений.

Согласно [3–5] задача удовлетворения ограничений (CSP — Constraint Satisfaction Problem) состоит из трех компонент: $\langle X, D, C \rangle$. X — множество

переменных $\{X_1, X_2, \dots, X_n\}$. D — множество доменов $\{D_1, D_2, \dots, D_n\}$, где D_i является доменом (областью определения) переменной X_i . C — множество ограничений $\{C_1, C_2, \dots, C_m\}$, которые предписывают допустимые комбинации значений переменных. Каждый домен D_i описывает множество допустимых значений $\{v_1, \dots, v_k\}$ для переменной X_i . Под ограничением понимается любое соотношение между переменными предметной области. В качестве ограничений могут выступать арифметические выражения; логические формулы; таблицы; выражения, формулируемые на языке специализированных теорий.

Решением задачи CSP является полное присваивание, которое удовлетворяет всем ограничениям. В некоторых случаях необходимо получить все решения. Иногда требуется найти такое решение, в котором значения переменных оптимизировали бы некоторый заданный функционал. CSP-задачи принадлежат классу NP-полных задач.

Технология программирования в ограничениях предоставляет мощные и гибкие методы, алгоритмы решения задач комбинаторного поиска. Особенности технологии:

1. С точки зрения конечного пользователя задача CSP формулируется в декларативном виде, на языке близком к языку математики. Порядок задания ограничений несущественен.

2. Любой алгоритм удовлетворения ограничений должен содержать две обязательных компоненты: а) компоненту, реализующую вывод (распространение); б) компоненту, реализующую поиск.

3. Вывод (распространение) реализуется как целенаправленное сужение изначально заданных областей определения переменных.

4. Эвристики, используемые в процедурах поиска, разрабатываются не под конкретную задачу, а являются универсальными.

5. Благодаря архитектуре систем программирования в ограничениях, появляется возможность совместно обрабатывать количественные и качественные ограничения.

6. Обеспечивается возможность сопровождать модели, открытые для оперативных модификаций. При добавлении/удалении из модели ограничений нет необходимости писать новые методы решения задачи.

Программирование в ограничениях особенно полезно там, где нет возможности получить удовлетворительное/адекватное решение средствами «обычной» математики. Оно используется при решении задач планирования, проектирования, прогнозирования, в инженерных и экономических расчетах, при создании графических интерфейсов, в системах понимания естественного языка и др. [6, 7].

Весь 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. Начнем с описания части, отвечающей за вывод. Как правило, под выводом на ограничениях понимается процесс сокращения размерности пространства поиска, 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*. Данные алгоритмы исключают из областей определения переменных заведомо лишние значения, то есть значения, которые не входят ни в одно из допустимых присваиваний. Завершение алгоритмов-распространителей может произойти с тремя возможными исходами. Во-первых, в результате работы алгоритмов-распространителей может быть получено решение исходной задачи CSP: все домены переменных сужаются до одноэлементных множеств. Во-вторых, может быть установлено, что область определения некоторой переменной пуста, и тогда решения соответствующей задачи CSP не существует. Наконец, алгоритмы-распространители могут остановиться, достигнув некоторой неподвижной точки, а решение задачи CSP еще не получено. В этом случае в дело вступают методы, реализующие стратегии интеллектуального поиска. Эти методы, 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.

Архитектура систем программирования в ограничениях напоминает архитектуру вычислительных машин (рис. 1). В качестве своеобразной шины данных в системах программирования в ограничениях выступает так называемое хранилище ограничений, представляющее собой, список пар «переменная — перечень значений переменной, которые она может принимать на текущем этапе вывода». Аналогом центрального процессора выступает алгоритм — дирижер (управляющая подпрограмма), реализующий базовую стратегию поиска.

В качестве сопрограмм выступают подпрограммы, реализующие вывод на ограничениях, — распространители ограничений. Для различных типов ограничений и типов областей определения переменных разрабатываются различные алгоритмы-распространители. В частности, для арифметических и логических ограничений целесообразно иметь различные алгоритмы-распространители, учитывающие специфику данных типов ограничений. Продолжая аналогию с архитектурой вычислительных машин, процедуры-распространители соотносятся с дополнительными устройствами (оперативная память, жесткий диск и т.п.), достаточно гибко подключаемыми через шину данных. Кроме стандартных типов ограничений и алгоритмов-распространителей, в рамках библиотек программирования в ограничениях существует возможность создания пользовательских типов ограничений и соответствующих процедур распространения. В частности, автором были введены новые типы ограничений — *C*- и *D*-системы, и созданы высокоэффективные методы рассуждений на данных структурах, которые подробно описаны ниже.

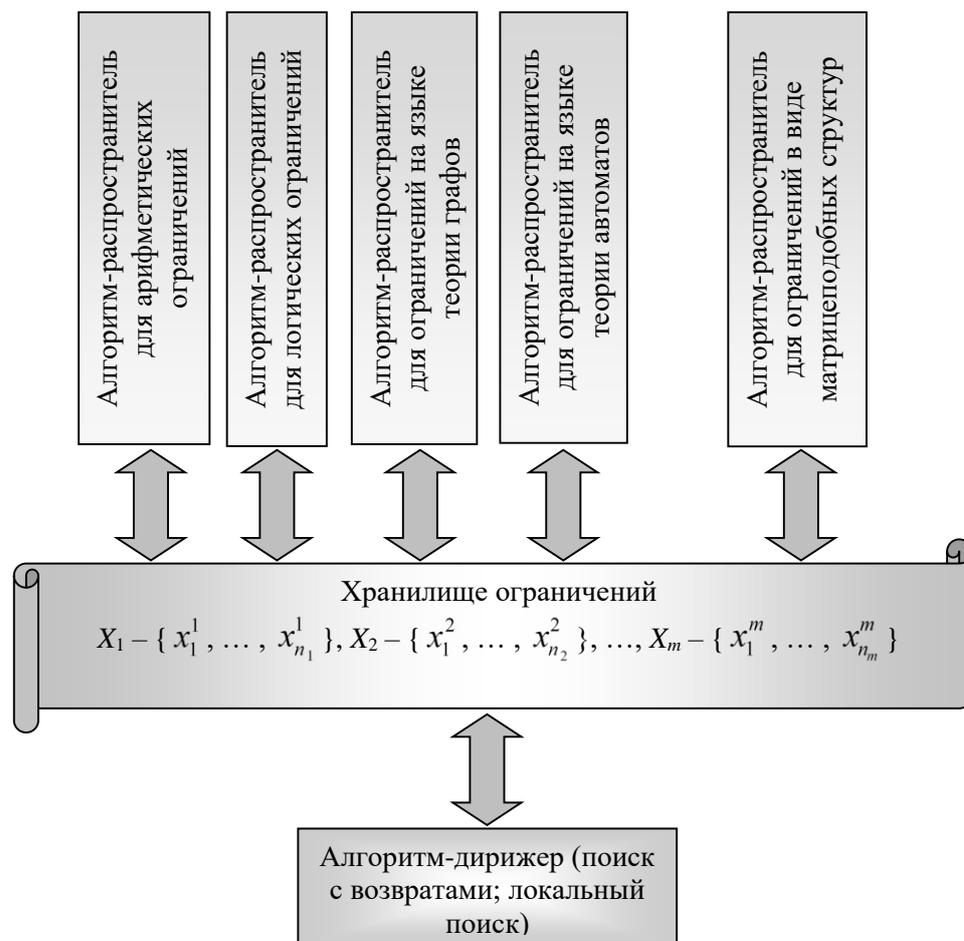


Рис. 1. Архитектура системы программирования в ограничениях

Пользователь-программист может изменять логику управляющей программы, в частности в случае поиска с возвратами пользователь может задавать эвристики для выбора переменной на текущем шаге, значения переменной. Могут задаваться действия, выполняемые в случае обнаружения тупиковой вершины и т.п.

В процессе решения конкретной задачи CSP используется следующий механизм активации ограничений (соответствующих процедур-распространителей) по событиям: ограничение активируется, если домен какой-либо переменной, которая входит в данное ограничение, был усечен. Также при активации ограничений имеются определенные приоритеты: сначала активируются наиболее простые ограничения, затем – более сложные, то есть содержащие большее количество переменных. Так, обычно, сначала активируются унарные ограничения, затем бинарные и т.д. Дело в том, что для простых ограничений (унарных, бинарных, иногда используются тернарные) алгоритмы их распространения имеют низкую вычислительную сложность.

Именно особенности архитектуры систем программирования в ограничениях создают предпосылки для унификации совместной обработки разнородной (количественной и качественной) информации (ограничений).

Однако на практике эффективность процедур подобной совместной обработки страдает из-за отсутствия высокопроизводительных процедур-распространителей для качественных ограничений.

Теперь перейдем к описанию авторских методов вывода на нечисловых ограничениях.

Пример прикладной задачи

Далее для иллюстрации применения предлагаемого подхода рассмотрим упрощенный пример планирования закупок для магазина автозапчастей.

Постановка задачи

На начало месяца в магазине автозапчастей отдел закупок получил денежные средства в размере 3000 руб. на закупку и доставку запчастей. Чтобы выполнить годовой план по прибыли, необходимо к концу месяца получить не менее 20% прибыли от продаж.

Ассортимент товаров известен и представлен тремя видами: Товар 1 (вес 4 кг), относится к категории АХ; Товар 2 (вес 14 кг), относится к категории СХ; Товар 3 (вес 1 кг), относится к категории ВХ.

Условия на закупку товара

Запчасти приобретаются у трех поставщиков с различными условиями закупки. Подразумевается, что товар всегда имеется в необходимом количестве, если иное специально не оговорено. Условия закупки товара указаны в таблицах 1–3.

Таблица 1. Условия закупки Товара 1

Поставщик 1	Не более 10 шт., цена 30 руб.	Более 10 шт., цена 27 руб.	Условие на объем закупки
Поставщик 2	Любое количество стоит 29 руб.		
Поставщик 3	При покупке 4 шт. цена 28 руб. за шт.		Закупка упаковками (по 4 шт.)

Поставщик 3 работает на оптовом складе с упаковками, поэтому Товар 1 продается только в количестве кратном 4.

Таблица 2. Условия закупки Товара 2

Поставщик 1	Особые условия закупки		
Поставщик 2	Любое количество стоит 48 руб.		
Поставщик 3	Не более 10 шт., цена 50 руб.	Более 10 шт., цена 45 руб.	Условие на объем закупки

Поставщик 1 продает товар Товар 2 по цене 35 руб., при условии его покупки от 15 шт. и совместно с Товаром 3 (один к одному). В противном случае цена 50 руб.

Таблица 3. Условия закупки Товара 3

Поставщик 1	Не более 20 шт, цена 15 руб.	Более 20 шт, цена 12 руб.	Условие на объем закупки
Поставщик 2	Товар отсутствует		
Поставщик 3	Любое количество стоит 17 руб.		

Согласно договору с Поставщиком 1 объем закупок у него должен быть не менее 2000 руб. в месяц. Также для поддержания ассортимента необходимо, чтобы был закуплен товар всех трех видов.

Условия на доставку товара

Условия на доставку товара от поставщиков зависят от объема закупки и задаются следующей таблицей.

Таблица 4. Условия доставки товара

Поставщик 1	Сумма заказа менее 2500 руб., стоимость 400 руб.	Сумма заказа не менее 2500 руб., доставка 0 руб.
Поставщик 2	Сумма заказа менее 2000 руб., стоимость 450 руб.	Сумма заказа не менее 2000 руб., доставка 0 руб.
Поставщик 3	Нет своей доставки	
Транспортная компания	Если вес менее 100 кг, то доставка 120 руб., от 100 до 200 кг включительно – доставка 250 руб., свыше 200 кг – доставка 500 руб.	

Условия на продажу товара

Розничная цена на каждый вид товара фиксирована и равна 39 руб., 65 руб. и 22 руб. соответственно. Предполагается, что запчасти будут продаваться в соответствии со статистическими данными, полученными на основе прошлых периодов. Данные, полученные на основе ABC- и XYZ-анализа, приведены в таблице.

Таблица 5. Результаты ABC- и XYZ-анализа

Товар	Среднее кол-во продаж	Средн. кв. откл.	Мин. кол-во продаж	Макс. кол-во продаж
Товар 1	50	3	47	53
Товар 2	30	15	15	45
Товар 3	100	15	85	115

Для выполнения условий дилерского соглашения необходимо реализовать Товар 2 в количестве не менее 20 шт. за месяц.

Условия на хранение товара

Непроданные в запланированный период запчастей остаются на складах и требуют дополнительных затрат на обслуживание: электроэнергию, отопление, заработную плату работникам склада и т.п.

Таблица 6. Условия хранения товара на складе

Товар 1	Если не более количества Товара 2, тогда 1 руб. за ед.	Если более количества Товара 2, тогда 6 руб. за ед.	Особое условие хранения
Товар 2	Стоимость 5 руб. за шт.		
Товар 3	Если не более 50 шт., тогда 1 руб. за ед.	Если более 50 шт., то 3 руб. за ед.	Условие на объем хранения

Товар 1 можно хранить вместе с Товаром 2, поэтому стоимость его хранения с Товаром 2 невелика. Однако если Товар 1 занимает целую ячейку, то цена за его хранение существенно возрастает.

Требуется найти объем закупки каждого товара у каждого поставщика, доставляющий требуемую прибыль (согласно статистике прошлых периодов) и удовлетворяющий заданным ограничениям.

Модель задачи

Будем считать, что каждый товар закупается порциями: одна – у одного поставщика, вторая – у другого, третья – у третьего. При этом если некоторый товар у поставщика мы не покупаем, то объем такой порции равен 0.

Введем обозначение x_i^j – j -ая порция i -того товара и представим переменные с помощью таблицы.

Таблица 7. Количество закупаемого товара

	Поставщик 1	Поставщик 2	Поставщик 3
Товар 1	x_1^1	x_1^2	x_1^3
Товар 2	x_2^1	x_2^2	x_2^3
Товар 3	x_3^1	x_3^2	x_3^3

Сумма столбцов по строке таблицы дает количество купленного товара $x_i = \sum_j x_i^j$.

Аналогично введем переменные c_i^j – цена i -того товара у j -того поставщика.

Таблица 8. Стоимость закупаемого товара

	Поставщик 1	Поставщик 2	Поставщик 3
Товар 1	c_1^1	c_1^2	c_1^3
Товар 2	c_2^1	c_2^2	c_2^3
Товар 3	c_3^1	c_3^2	c_3^3

Сумма, уплачиваемая j -тому поставщику за весь покупаемый у него товар, будет равна $\sum_i x_i^j c_i^j$, а общая сумма закупки у всех поставщиков равна

$$S = \sum_j \sum_i x_i^j c_i^j.$$

Домены переменных, получаемые из постановки задачи: $x_1^1, x_1^2, x_1^3 \in [0;111]$, $x_2^1, x_2^2, x_2^3 \in [0;66]$, $x_3^1, x_3^3 \in [0;250]$, $x_3^2 = 0$, $c_1^1 \in \{27;30\}$, $c_1^2 \in \{29\}$, $c_1^3 \in \{28\}$, $c_2^1 \in \{35;50\}$, $c_2^2 \in \{48\}$, $c_2^3 \in \{45;50\}$, $c_3^1 \in \{12;15\}$, $c_3^2 \in \{50\}$ (цена ставится произвольная, т.к. $x_3^2 = 0$), $c_3^3 \in \{17\}$.

Сформулируем ограничения на закупку Товара 1:

$$C1: [(x_1^1 \leq 20) \wedge (c_1^1 = 30)] \vee [(x_1^1 > 20) \wedge (c_1^1 = 27)],$$

$$C2: c_1^2 = 29,$$

$$C3: (x_1^3 \geq 4) \wedge (c_1^3 = 28).$$

Аналогичным образом записываются ограничения **C4-C9** на закупку Товара 2 и Товара 3. Дополнительные условия на закупку запишем в виде следующих ограничений:

ограничение на объем закупки у Поставщика 1

$$C10: x_1^1 c_1^1 + x_2^1 c_2^1 + x_3^1 c_3^1 \geq 2000,$$

ограничение на представление ассортимента всех товаров

$$C11: x_1^1 + x_1^2 + x_1^3 > 0,$$

$$C12: x_2^1 + x_2^2 + x_2^3 > 0,$$

$$C13: x_3^1 + x_3^2 + x_3^3 > 0.$$

Введем переменные: $s_j = \sum_i x_i^j c_i^j$ – стоимость закупки у j -того

поставщика, $v_j = 4x_1^j + 14x_2^j + x_3^j$ – вес товара от j -того поставщика, d_j – стоимость доставки товара от j -того поставщика. Далее сформулируем ограничение на доставку товара от Поставщика 1:

$$C14: [(s_1 < 2500) \wedge (d_1 = 120)] \vee [(s_1 \geq 2500) \wedge (d_1 = 0)] \vee$$

$$\vee [(v_1 < 100) \wedge (d_1 = 120)] \vee [(v_1 \geq 100) \wedge (v_1 \leq 200) \wedge (d_1 = 250)] \vee$$

$$\vee [(v_1 > 200) \wedge (d_1 = 500)].$$

Здесь учитывается возможность доставки средствами Поставщика 1 и доставки транспортной компанией.

Ограничения **C15** и **C16** на доставку товара от Поставщика 2 и Поставщика 3 записываются аналогичным образом.

Введем обозначение $Q = \sum_i (s_i + d_i)$ – общая стоимость покупки

и доставки товара. Тогда ограничение на сумму имеющихся средств имеет вид:

$$C17: Q \leq 3000.$$

Введем переменные y_i – количество продаж i -того товара за месяц. Домены для переменных y_i берутся из таблицы 5: $y_1 \in [47;53]$, $y_2 \in [15;45]$, $y_3 \in [85;115]$. Сформулируем дополнительное ограничение на продажу Товара 2:

$$C18: y_2 \geq 20 .$$

Обозначим через P общую выручку от реализации товара:

$$P = 39y_1 + 65y_2 + 22y_3 .$$

Введем переменные h_i – стоимость хранения i -того товара на складе и сформулируем ограничения на хранение товара:

$$C19: [(x_1 - y_1 \leq x_2 - y_2) \wedge (h_1 = x_1 - y_1)] \vee [(x_1 - y_1 > x_2 - y_2) \wedge (h_1 = 6(x_1 - y_1))] \text{ (ограничение на хранение Товара 1),}$$

$$C20: h_2 = 5(x_2 - y_2) \text{ (ограничение на хранение Товара 2),}$$

$$C21: [(x_3 - y_3 \leq 50) \wedge (h_3 = x_3 - y_3)] \vee [(x_3 - y_3 > 50) \wedge (h_3 = 3(x_3 - y_3))] \text{ (ограничение на хранение Товара 3).}$$

$$\text{Обозначим } H \text{ – общую стоимость хранения излишков товара: } H = \sum_i h_i$$

и запишем ограничение на прибыль:

$$C22: P - Q - H \geq 600 .$$

Заключение

Проведенные исследования показали перспективность применения технологии программирования в ограничениях для задачи планирования закупок малым предприятием. Применение данной технологии дает возможность совместно обрабатывать не только количественные, но и качественные зависимости предметной области, служащие для формализации гибких ограничений на поставку, сроки и стоимость доставки, хранение и продажу товара. Использование технологии программирования позволяет обеспечить возможность сопровождения модели, открытой для оперативных модификаций. Потребность в модификации модели возникает довольно часто: когда задача не имеет решения и требуется “ослабить” условия задачи, когда требуется ввести в рассмотрение новые критерии и т.п. При добавлении/удалении из модели ограничений нет необходимости писать новые методы решения задачи, что выгодно отличает предлагаемую технологию от методов теории исследования операций [8].

Среди направлений дальнейших исследований следует отметить целесообразность разработки специализированных глобальных ограничений [9] для повышения эффективности обработки логических условий (качественных ограничений), поскольку при росте размерности задачи (пространства поиска) недостаточно эффективные процедуры удовлетворения нечисловых ограничений вызвать значительные трудности.

В качестве одного из дальнейших направлений развития работы в задачу можно ввести критерии оптимизации. Тогда задача удовлетворения ограничений превратится в задачу ограниченной оптимизации (COP — Constraint Optimization Problem). Систематизация подобных критериев представляется

актуальной задачей, ввиду возможности создания дополнительных средств ускорения вычислительных процедур на основе анализа подобных критериев.

Литература

1. Требинский В. В. Опыт применения ABC- и XYZ-анализов в управлении ассортиментом розничного предприятия на примере книжного магазина // Вопросы современной науки и практики. Университет им. В. И. Вернадского. 2008. № 1(11). Том 2. С. 132–136.
2. Хамлова О. ABC-анализ: методика проведения // Управление компанией. 2006. № 10. С. 54–57.
3. Russel S., Norvig P. Artificial Intelligence: A Modern Approach. 3rd edition. Prentice Hall, 2010. 132 p.
4. Ruttkay Zs. Constraint satisfaction a survey // CWI Quarterly. – 1998. – Vol. 11. pp. 163–214.
5. Bartak R. Constraint Programming: In Pursuit of the Holy Grail // Proceedings of the Week of Doctoral Students (WDS99). 1999. Part IV. pp. 555–564.
6. Margaux N. Cumulative scheduling with variable task profiles and concave piecewise linear processing rate functions / N. Margaux, C. Artigues, P. Lopez // Constraints. 2017. Vol. 22(4). pp. 530–547
7. Kreter S. Models and solution procedures for the resource-constrained project scheduling problem with general temporal constraints and calendars / S. Kreter, J. Rieck, J. Zimmermann // European Journal of Operational Research. 2016. № 251(2). pp. 387–403.
8. Гермейер Ю. Б. Введение в теорию исследования операций. М.: Наука, 1971. 384 с.
9. R’egin J. Generalized arc consistency for global cardinality constraint // Proceedings of the Thirteenth National Conference on Artificial Intelligence. Portland. 1996. pp. 209–215.

Сведения об авторах

Зуенко Александр Анатольевич

к.т.н., ведущий научный сотрудник

e-mail: zuenko@iimm.ru

Alexander A. Zouenko

Ph.D. (Tech. Sci.), leading researcher

Македонов Роман Александрович

младший научный сотрудник

e-mail: makedonov@iimm.ru

Roman A. Makedonov

junior researcher

П. А. Ломов

Институт информатики и математического моделирования ФИЦ КНЦ РАН

ПРИМЕНЕНИЕ ГРАФОВЫХ СУБД В ЗАДАЧАХ АНАЛИЗА ДАННЫХ*

Аннотация

Данная статья посвящена вопросу применения графовых систем управления базами данных (СУБД) при решении задач анализа данных. Рассматриваются структуры распространенных графовых моделей данных – модель "субъект-предикат-объект" и модель графа свойств, а также графовые алгоритмы и их применение для анализа данных. Приводится описание качественных отличий между графовыми и реляционными СУБД, которые целесообразно учитывать при их выборе в качестве средства для хранения и анализа данных с различной структурой.

Ключевые слова:

графовая СУБД, RDF, граф свойств, графовые алгоритмы

P. A. Lomov

APPLICATION OF GRAPH DATA BASES FOR DATA ANALYSIS

Abstract

This article is devoted to the use of graph databases for solving data analytical problems. The structures of common graph data models – the "subject-predicate-object" model and the property graph model, as well as their main differences are considered. Along with this, common graph algorithms and their application for solving some types of analytical problems are presented. The article describes the qualitative differences between graph and relational DBMSs, which should be taken into account when choosing a tool for storing and analyzing data with a different structure.

Keywords:

graph DBMS, RDF, property graph, graph algorithms.

Введение

Современная информационная система как правило характеризуется большим объемом данных, обладающих большим количеством и сложной структурой связей. В зависимости от выбранной модели данных представление таких связей может осуществляться по-разному. Например, в реляционной модели они могут быть представлены в виде «соединений» реляционных таблиц по внешнему ключу. Это позволяет в процессе выполнения запроса объединять отдельные элементы данных (кортежи таблиц) и получать в результате более сложный фрагмент данных, содержащий ответ на заданный запрос.

Наличие сложной структуры связей в данных позволяет в ходе их анализа находить несколько альтернатив для решения некоторой прикладной задачи, учитывать некоторые дополнительные факторы, а также выявлять решения похожих задачи, имевшие место в прошлом.

Связи элементов данных могут быть представлены явно путем применения для этого средств, предоставляемых моделью данных, или могут формироваться «на лету» при использовании хранилища данных. В последнем

* Работа выполнена при финансовой поддержке РФФИ (проект № 17-47-510298 p_a).

случае пользователь должен сформулировать условия их установления в запросе, а программный компонент, отвечающий за его выполнение, вычислить такие связи между элементами данных.

Эффективность работы хранилища с данными, обладающими большим количеством внутренних связей является на сегодняшний день очень важным фактором его выбора для непосредственного использования в решении задач анализа данных, требующих получения сложных фрагментов данных с различной структурой. В этой связи большое распространение на сегодняшний день получили графовые СУБД. Они относятся к типу так называемых нереляционных (NoSQL, Not only SQL) хранилищ и ориентированы на явное отражение всего множества связей между элементами данных. Среди известных реализаций графовых СУБД можно выделить Neo4J, AnzoGraph, Titan, JanusGraph, OpenLink Virtuoso, Oracle Spatial and Graph (дополнение к Oracle Database), OrientDB, а также облачные сервисы Amazon Neptune, Azure Cosmos DB. В качестве языков запросов к ним выступают Gremlin, SPARQL, Cypher, GraphQL.

В данной статье рассматривается вопрос применения графовых СУБД для решения задач анализа данных, а также некоторые их особенности.

1. Модели данных графовых баз данных

В основе графовых моделей данных лежит понятие графа, как совокупности множеств вершин и ребер:

$$G = (V, E),$$

где V – множество вершин, а $E \subset V \times V$ – множество ребер.

Заметим, что в моделях данных конкретных графовых СУБД могут присутствовать те или иные дополнения данного определения, например, возможность определения взвешенных и/или направленных ребер. Наибольшее распространение на сегодняшний день получила модель графа свойств или атрибутивного графа (Property graph model) (рис.1).

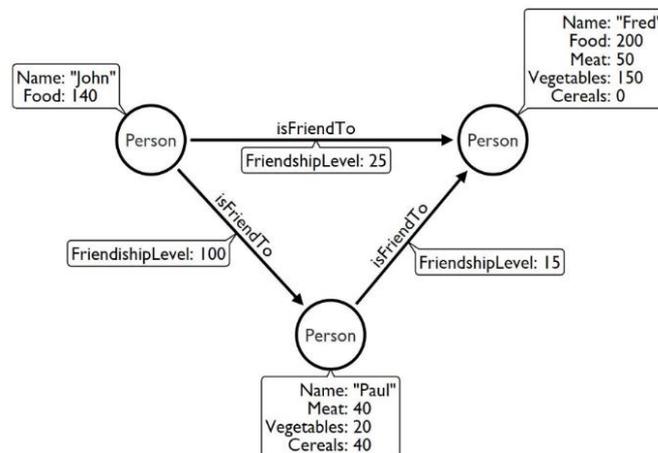


Рис. 1. Пример графа свойств

Она подразумевает наличие наименований у вершин (node labels) и ребер (relationship types). Ребра при этом могут иметь направление. С вершинами и ребрами также могут быть ассоциированы наборы свойств в виде пар— (ключ, значение). Возможные типы значений могут варьироваться в зависимости от конкретной реализации модели от простых (например, строкового, целочисленного и т.д.) до более сложных (например, списки, словари).

Другой распространенной графовой моделью является моделью триплетов или RDF-моделью (Resource Description Framework, RDF). Она состоит из триплетов или троек «субъект-предикат-объект», которые также можно интерпретировать в виде пары вершин (субъект и объект) и инцидентного им ребра (предикат). Однако если сравнивать модель графа свойств с RDF-моделью, то основным отличием будет являться наличие у элементов модели графа свойств внутренней структуры (т.е. некоторого набора свойств). В свою очередь в RDF-модели элементы триплета определяются просто в виде строк символов, являющихся URI-идентификаторами (Universal Resource Identifier, URI) или литеральными значениями. Иными атрибутами RDF-модель их не наделяет.

Другим отличием будет возможность определения экземпляров ребер, то есть задание несколько ребер одинакового типа между вершинами. Модель графа свойств обеспечивает данную возможность. Каждое ребро в этом случае будет иметь уникальный идентификатор, тогда как в RDF-модели триплет «субъект-предикат-объект», определяющий якобы еще один экземпляр ребра между, будет считаться дубликатом, виду того что URI предиката будет тем же самым, что и в исходном триплете. Следствием этого является также то, что RDF-модель не обладает имманентной возможностью характеризовать ребра, то есть определять их свойства, например, вес ребра. Хотя это ограничение модели обходиться с помощью приема реификации [1], то есть трансформации ребра в промежуточную вершину, смежную с исходными. Однако смысл, стоящий за такой модификацией, необходимо будет поддерживать уже на уровне приложения, осуществляющего манипуляции с графом.

Еще одним отличием RDF-модели является возможность задавать «четверки» (quads), то есть определять принадлежность триплета некоторому графу и тем самым разделять граф на компоненты. Модель графа свойств явно такой данной возможности не предоставляет. Однако такого сегментирования можно добиться заданием дополнительных меток вершин и типов ребер, указывающих на принадлежность к некоторому подграфу, или заданием нескольких графов, несвязанных между собой и имеющих в качестве идентификаторов некоторые «центральные» вершины. Однако, как и случае реификации, учет таких приемов метамоделирования должен быть реализован в бизнес-логике приложения или приниматься во внимание пользователем при формулировке и интерпретации результатов запросов.

2. Основные графовые алгоритмы для анализа данных с помощью графовых СУБД

Рассмотрим применение графовых баз данных для решения задач анализа данных. Заметим, что в данном случае подразумеваются те возможности, которые обычно предоставляются самой графовой СУБД и ориентируются именно на оперирование графовой моделью. В основе этих возможностей лежат алгоритмы,

применяемые для решения известных задач теории графов. Их можно разделить на следующие на 3 группы:

- алгоритмы решения задачи поиска кратчайшего пути на графе и ее вариации;
- алгоритмы определения центральности вершины (Degree centrality);
- алгоритмы кластеризации вершин графа.

В первую группу входят алгоритмы поиска в длину (Bread-first search, BFS) и ширину (Depth-first search, DFS), нахождение кратчайшего пути между каждой парой вершин из двух наборов (All pairs shortest path, APSP) нахождение пути от заданной вершины до всех вершин некоторого набора (Single-Source Shortest Paths, SSSP), поиск минимального связующего дерева (Minimum Spanning Tree, MST), т.е. набора дуг с минимальной суммарной стоимостью, соединяющих все вершины графа без учета циклов.

Вторую группу составляют алгоритмы определения центральности вершины (Degree centrality). К ним относится вычисление центральности на основе близости вершин – центральность по близости (Closeness centrality), то есть определение вершин с кратчайшим расстоянием до всех остальных узлов. При решении практических задач это позволяет найти вершину, из которой наиболее быстро можно распространить некоторый сигнал по всей сети. Однако в реальных ситуациях графы часто являются несвязными. В этом случае длина пути между вершинами двух компонентов считается равной бесконечности, что в свою очередь приводит к неверным результатам оценки центральности (бесконечной центральности) для несвязанных узлов. В таких случаях применяют оценку центральности в рамках отдельных компонент графа с помощью алгоритма Вассермана-Фауста (Wasserman Faust) [2].

Иной способ подсчета центральность по близости для несвязных графов состоит в игнорирование бесконечных путей. Он реализован в алгоритме подсчета гармонической центральности (Harmonic/Valued Centrality) [3] и заключается в том, что бесконечный путь между несвязанными вершинами дает нулевой вклад в результирующую оценку центральности.

Другой подход к определению центральности базируется на идее, что важные вершины располагаются на наиболее востребованных путях в графе. Это так называемая промежуточная центральность. Ее расчет можно применять для решения задач поиска «узких» мест или критических участков различных сетей. Для подсчета в этом случае применяется алгоритм, определяющий центральность вершины в зависимости от того, какая доля всех кратчайших путей между каждой парой вершин проходит через нее. Однако в больших графах перебор всех пар вершин может потребовать неприемлемое количество времени, поэтому используют более быстрый Randomized-Approximate (RA-Brandes) алгоритм, дающий приближительные результаты. Его отличие состоит в том, что рассматриваются не все пары узлов, а лишь их подмножество, которое определяется случайным образом или с учетом степени вершины, то есть вершины со сравнительно низкой степенью имеют низкую вероятность быть включенными в рассмотрение.

Еще один вид центральности и одновременный способ ее подсчета это центральность по степени (degree centrality). Она определяется как отношение количества инцидентных вершине дуг к общему количеству вершин. Выделяют

входящую (indegree) и исходящую (outdegree) степенную центральность, в зависимости от того какие инцидентные дуги рассматривают при подсчете. На практике это позволяет выявить важных инициаторов сигналов или основных коллекторов, получающих сигнал от многих узлов сети.

Важным является так называемая центральность по собственному вектору (Eigenvector centrality), которую можно рассматривать как разновидность степенной. Она отражает своего рода «полезную» степенную центральность вершины, то есть меру связи вершины с теми, которые в свою очередь имеют большое число инцидентных ребер. Один из распространенных алгоритмов для ее определения является – PageRank. Он работает итеративно, осуществляя последовательные переходы с некоторой вероятностью от произвольной начальной вершины к другим и уточняя начальную оценку центральности вершин, до того как сойдется или закончится заданное число итераций. Его вариацией является персонализированный PageRank (Personalised Page Rank), который ограничивает множество вершин для возможных переходов и тем самым позволяет оценить центральность выбранной вершины, исходя из дуг, инцидентных вершинам этого ограниченного набора.

Еще одним видом графовых алгоритмов являются алгоритмы кластеризации вершин. Их также называют алгоритмами определения сообществ (Community Detection Algorithms) по распространенной прикладной задаче, решаемой ими. Эти алгоритмы направлены на выявления множеств вершин, связанных инцидентными им ребрами преимущественно между собой, чем с другими вершинами графа. При практическом применении эти алгоритмы используются в задачах определения групповых предпочтений, установления устойчивости социальных групп, визуализации графов. К алгоритмам этого группы в первую очередь относятся алгоритмы определения числа треугольников (Triangle count) и локального коэффициента кластеризации (The local clustering coefficient). Треугольником называется клика из 3 вершин, то есть множество из трех вершин графа, в котором каждая соединена с двумя другими. Определение локального коэффициента кластеризации для вершины и ее прямых соседей — смежных с ней вершин, вычисляется как отношение удвоенного числа треугольников, включающих данную вершину, к количеству всех возможных связей между рассматриваемыми вершинами [4]. Тем самым, данные алгоритмы позволяют установить меру связности вершины и ее соседей между собой. Глобальный коэффициент кластеризации можно определить как средневзвешенную сумму локальных коэффициентов для всех вершин.

Алгоритм определения компонент сильной связности (Strongly Connected Components, SCC) позволяет выявить такие наборы связанных вершин в ориентированном графе, где каждая вершина доступна в обоих направлениях от любой другой в этом же наборе. Для неориентированного графа применяется также похожий алгоритм для поиска связанных компонент (Connected Components), который отличается требованием только наличия пути между двумя вершинами. [5]. Оба алгоритма ввиду их сравнительно невысокой вычислительной сложности применяются для грубого выявления групп, предварительного анализа графа, поиске циклов, а также для общего понимания структуры графа.

Следующим алгоритмом поиска сообществ является алгоритм распределения меток (Label Propagation algorithm, LPA). Идея, лежащая в основе

алгоритма, состоит в итеративном присвоении метки некоторой вершины расширяющемуся кругу ее соседей по некоторым правилам. По завершении процесса распространения, исходя из присвоенных вершинам меток можно сделать вывод относительно структуры сообществ, представленных рассматриваемым графом. Вопрос присвоения или не присвоения метки некоторой вершине на каждой итерации определяется, исходя из количества соседних ей вершин, имеющих данную метку, а также веса инцидентных ребер. [6] Критериями для остановки распространения может являться требование того, чтобы каждая вершина имела ту же метку, что и большинство ее соседей. В начале работы алгоритма уникальные метки присваиваются всем вершинам. Однако в этом случае, результаты его работы могут быть различными для разных попыток. Поэтому можно также присвоить так называемые “посевные” метки (Seed label) некоторым вершинам. В этом случае весьма вероятно, что их соседи будут наследовать именно их, что позволит стабилизировать вариационность полученных результатов.

Следующий алгоритм для кластеризации является Лувенский алгоритм (Louvain Method) [7]. Считается одним из самых быстрых алгоритмов для графов большого размера (более 10 миллионов вершин). Наряду с обнаружением сообществ он также отражает их иерархию. Это позволяет выявить структуру сети на разных уровнях детализации – от больших групп до малых. Ключевым понятием данного алгоритма является модулярность. Она является мерой качества текущего распределения вершин графа по кластерам. Таким образом, в процессе работы для разбиения подсчитывается показатель модулярности, которая оптимизируется на последующих итерациях путем перераспределения вершин. Перераспределение, увеличивающее модулярность, заключается в формировании групп таким образом, чтобы внутри них вершины имели больше связей между собой, чем с вершинами других групп.

Работа алгоритма состоит из двух повторяющихся шагов. На первом шаге для выбранных вершин с целью максимизации модулярности определяются группы из их прямых и не прямых соседей. На втором шаге на основе групп с первого шага происходит формирование более крупных групп. Это происходит на основе общего числа связей между вершинами отдельных групп, то есть две группы объединяются в одну большую, если общее число связей между их вершинами больше, чем с вершинами других групп.

Среди графовых алгоритмов можно также выделить алгоритмы генерации графов Girvan-Newman (GN) [8] и Lancichini-netti-Fortunato-Radicchi (LFR) [9]. В основном они применяются для проверки алгоритмов определения сообществ. Для этого они позволяют сгенерировать граф с известным числом сообществ и характеристиками, похожими на характеристики реального графа, и провести тестирование выбранного алгоритма.

3. Особенности применения графовых баз данных

Рассмотрим некоторые особенности графовых баз данных, которые необходимо учитывать при их использовании для решения задач анализа данных. В первую очередь, следует отметить, что следует избегать применения графовой модели для анализа транзакционных данных. Особенно если их объем (т.е. количество элементов) предполагается очень большим. Например, к этому можно

отнести временные ряды. В этом случае гораздо эффективнее применять реляционную модель, в которой можно хранить элементы данных транзакций некоторого вида в рамках одной таблицы (реляционного отношения). В таком случае выполнение сложных аналитических запросов, благодаря различным механизмам индексации, будет выполняться наиболее эффективно. В этом случае анализ данных более ориентирован (как и сама реляционная модель) на учет связей между элементами данных одного вида (далее будем называть эти связи «вертикальными»). Например, в некоторой таблице «Закупка компонентов» каждый кортеж будет отражать данные одной транзакции - дата, количество, поставщик, цена, название компонента. Соответственно любой аналитический запрос будет подразумевать изначальное получение идентификаторов некоторых кортежей как ссылки на некоторые целостные сложные элементы данных (например, закупочные транзакции), а уже потом обработку значений их полей.

Таким образом, реляционная модель подходит в большей степени для анализа элементов данных с неизменной структурой, которые можно представить в рамках одной таблицы. При этом она эффективно формирует «вертикальные» связи в процессе выполнения запроса. Например, сортирует ежемесячные закупки по цене. Связи между элементами данных разных видов (далее будем называть их «горизонтальными») установлены на этапе загрузки данных в таблицу в соответствии типами ее полей. Например, названия поставщиков будут распределены по таблице «Закупка компонентов» с учетом того в какое время, какое количество, какого компонента, по какой цене было у них закуплено.

Поэтому применение реляционной модели в отношении анализа транзакционных данных, каждый вид которых имеет неизменную структуру, дающую возможность заранее сформировать «вертикальные» связи, более предпочтителен. Однако, если необходимо анализировать данные, представленные в разных таблицах, то в этом случае «вертикальные» связи нужно формировать на этапе выполнения запроса с помощью операции соединения таблиц. Увеличение числа объединяемых таблиц с большим количеством числом кортежей могут сделать время выполнения запроса неприемлемым.

Поэтому в случае, если число учитываемых при аналитических операциях связей велико и их перечень в разных запросах трудно предсказать, то в этом случае можно рассматривать возможность применения графовой модели, в которой и «вертикальные» и «горизонтальные» связи задаются явно в виде дуг графа, что сводит операции соединения таблиц к поиску путей между вершинами. К таким данным можно отнести нормативно-справочные данные (Reference Data Management, RDM), и мастер-данные (Master Data Management, MDM).

Отдельно отметим, что графовую модель и соответствующую ей СУБД рекомендуется использовать при решении задач анализа данных, в которых объект анализа в реальности представляет собой сетевую структуру. Например, социальная сеть, коммутационная сеть, логистическая сеть и т.д. В этом случае моделирование упрощается ввиду сходства элементом модели данных и частей объекта моделирования и представляется возможность «в лоб» применить графовые алгоритмы. Противоположной ситуацией является та, в которой объект моделирования необходимо сводить к графовой модели (например, набор текстовых документов). В этом случае при таком сведении для успешного проведения аналитических процедур следует изначально ориентироваться на применение и последующую интерпретацию результатов применения выбранных

графовых алгоритмов и, исходя из этого, выбирать какими элементами графовой модели будут представлены части объекта моделирования.

Заметим также, что использование графовых баз данных дополнительно открывает возможности для применения визуального анализа данных в виду того, что результатом запросов в большинстве случаев будет являться графовая структура, которая подразумевает явную и привычную визуальную метафору — множество вершин, соединенных ребрами, чего нельзя сказать в отношении реляционных баз данных, представляющих в виде ответа некоторую таблицу. Поэтому для облегчения интерпретации результатов аналитических операций имеет смысл дополнительно использовать средства визуального анализа (например, Gephi [10]), позволяющие визуально отразить свойства элементов полученного графа.

Заключение

На сегодняшний день графовые СУБД вполне можно рассматривать как альтернативу традиционным реляционным СУБД в отношении хранения и анализа данных, обладающих сложной иерархической структурой, различные элементы которой по разному могут обрабатываться в рамках аналитических операций. В первую очередь это относится к нормативно-справочным данным (Reference Data Management, RDM), и мастер-данным (Master Data Management, MDM). Наряду с этим графовые СУБД хорошо подходят для моделирования объектов предметной области, обладающих явной сетевой структурой.

Распространенными графовыми моделями данных на данный момент являются RDF-модель, представляющая данные в виде троек (субъект, свойство, значение), и модель графа свойств или атрибутированного графа (Property graph), позволяющая задавать внутреннюю структуру у вершин и дуг, в виде наборов пар (ключ, значение).

Представление данных в виде графовой модели дает возможность применять в рамках аналитических операций множество алгоритмов с хорошо исследованными свойствами. Однако эффективность их применения напрямую зависит от сформированного подхода к моделированию, так как именно это определяет какого рода информация и в какой степени будет учтена при анализе.

Использование в качестве математической основы графовых СУБД теории графов делает весьма востребованным выполнение визуализации результатов запросов для ускорения и облегчения их интерпретации аналитиком. Формирование такой визуализации как правило заключается в основном к выбору взаиморасположения вершин и узлов графа, их размера, окраски. Но уже это потенциально обладает большей вариативностью, чем отображение результата в виде таблицы при использовании реляционных СУБД. Таким образом, можно сказать, что хоть графовая модель и предполагает использование конкретной визуальной метафоры, но обеспечение ее эффективности с точки зрения скорости и правильности дальнейшей интерпретации представляет собой сложную проблему, которую следует решать с учетом условий исходной задачи и стереотипов пользовательского восприятия.

Таким образом, применение графовых СУБД становится на сегодняшний день довольно привлекательным для хранения и анализа данных, представляющих собой сложные иерархические структуры. Однако принятие

решения об их использовании должно быть обосновано необходимостью учета большого количества связей между фрагментами данных и высокой вариативностью аналитических запросов.

Литература

1. Nguyen V., Bodenreider O., Sheth A. (2014) Don't Like RDF Reification?: Making Statements about Statements Using Singleton Property. // Proceedings of the 23rd international conference on World wide web. April 07–11, 2014, Seoul, Korea DOI:10.1145/2566486.2567973.
2. Wasserman S., & Faust K. (1994). Social Network Analysis: Methods and Applications. Vol. 8. Cambridge: Cambridge University Press.
3. Marchiori M.; Latora V. (2000) Harmony in the small-world // Physica A: Statistical Mechanics and its Applications. 285 (3–4). pp. 539–546.
4. Stanley W., Faust K. (1994) Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press p. 243.
5. Galler B. A., Fischer M. J. (1964) An Improved Equivalence Algorithm.
6. Raghavan U. N., Albert R., Kumara S. (2007) Near Linear Time Algorithm to Detect Community Structures in Large-Scale Networks // Physical Review E – Statistical, Nonlinear and Soft Matter Physics. APS. Vol. 76, Issue 3, pt. 2, pp. 036106.
7. Blondel V. D.; Guillaume J.-L.; Lambiotte R.; Lefebvre E. (2008) Fast Unfolding of Communities in Large Networks // Journal of Statistical Mechanics: Theory and Experiment. 2008 (10). 10008 p.
8. Girvan M., Newman M.E.J. (2002) Community Structure in Social and Biological Networks. P. Natl. Acad. Sci. USA 99(12), pp. 7821–7826.
9. Lancichinetti S., Radicchi F. (2008) Benchmark Graphs for Testing Community Detection Algorithms. Physical Review E, 78.
10. Пакет программного обеспечения для сетевого анализа и визуализации. Режим доступа: <https://gephi.org/>.

Сведения об авторах

Ломов Павел Андреевич

к.т.н., старший научный сотрудник

e-mail: lomov@iimm.ru

Pavel A. Lomov

PhD (Tech. Sci.), senior researcher

DOI: 10.25702/KSC.2307-5252.2019.9.146-157

УДК 004.6, 004.9

М. Г. Шишаев¹, А. В. Вицентий¹ Н. М. Куприков^{1,2}

¹ *Институт информатики и математического моделирования ФИЦ КНЦ РАН*

² *Московский авиационный институт (национальный исследовательский университет)*

КОНЦЕПЦИЯ НАЦИОНАЛЬНОЙ СИСТЕМЫ УПРАВЛЕНИЯ ДАННЫМИ: СОВРЕМЕННЫЙ КОНТЕКСТ РЕАЛИЗАЦИИ

Аннотация

В статье рассмотрены основные положения концепции Национальной системы управления данными (НСУД), ее цели и организационно-правовое обеспечение. Приведены особенности Арктической зоны РФ как тестового региона реализации практических решений НСУД. Рассмотрены отечественные и зарубежные аналоги системы. Отдельное внимание уделено принципам открытости, интероперабельности и повторного использования данных FAIR

Ключевые слова:

национальная система управления данными, государственная информационная система, АЗРФ, научное облако, центр больших данных, FAIR-данные.

M. G. Shishaev, A. V. Vicentiy, N. M. Kuprikov

CONCEPT OF A NATIONAL DATA MANAGEMENT SYSTEM: MODERN IMPLEMENTATION CONTEXT

Abstract

The article discusses the main points of the concept of the National Data Management System (NSMS), its goals, organizational and legal basis. Features of the Russian Arctic zone as a test region for the implementation of practical solutions of the NSMS are given. Domestic and foreign analogues of the system are considered. Special attention is paid to the principles of findability, accessibility, interoperability and reusability (FAIR) of data.

Keywords:

National data management system, state information system, Russian Arctic Zone, Open Science Cloud, BD Hub, FAIR data.

1. Введение

В настоящее время в нашей стране реализуется национальная программа «Цифровая экономика Российской Федерации», паспорт которой утверждён решением президиума Совета при Президенте Российской Федерации по стратегическому развитию и национальным проектам 24 декабря 2018 года [1]. В целях реализации мероприятий федерального проекта «Цифровое государственное управление», входящего в эту национальную программу, разработана концепция создания и функционирования национальной системы управления данными (НСУД). Концепция, а также «дорожная карта» по созданию НСУД на 2019-2021 годы, утверждены распоряжением правительства Российской Федерации от 3 июня 2019 года [2].

Концепция определяет цели, задачи и принципы создания и функционирования НСУД, а также порядок ее создания, основные элементы и общую оценку ожидаемого социально-экономического эффекта от ее создания. Следуя определениям, данным в Концепции, национальная система управления

данными это система, состоящая из взаимосвязанных элементов информационно-технологического, организационного, методологического, кадрового и нормативно-правового характера и обеспечивающая достижение целей и выполнение задач, обозначенных в Концепции.

«Основной целью создания и обеспечения функционирования Системы является повышение эффективности создания, сбора и использования государственных данных как для предоставления государственных и муниципальных услуг и осуществления государственных и муниципальных функций, так и для обеспечения потребности физических и юридических лиц в доступе к информации. Достижение указанной цели будет осуществляться за счет нормативных правовых, методологических, информационно-технологических, организационных и кадровых механизмов посредством:

- повышения доступности государственных данных;
- обеспечения полноты, актуальности, непротиворечивости и связанности государственных данных;
- обеспечения информационной безопасности.»

Ответственность за реализацию Концепции и «дорожной карты» возложена на Минкомсвязи России (Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации), Минэкономразвития России, Аналитический центр при Правительстве Российской Федерации, совместно с заинтересованными федеральными органами исполнительной власти.

В соответствии с Концепцией, для обеспечения потребности физических и юридических лиц в доступе к информации, а также осуществления государственных и муниципальных функций и предоставления услуг, необходимо обеспечить высокую эффективность создания, сбора и использования государственных данных. Определение поставщиков и пользователей государственных данных, а также ведение реестра и обеспечение обработки запросов на получение государственных данных, возложено на федеральную государственную информационную систему «Единая информационная платформа национальной системы управления данными», которая является неотъемлемой частью НСУД. С учетом того, как в настоящее время организовано межведомственное электронное взаимодействие, основными поставщиками государственных данных для НСУД могут стать органы государственной (муниципальной) власти, осуществляющие предоставление различных государственных и муниципальных услуг (министерства, комитеты, инспекции и т.д.) [3].

При этом, непосредственно передача государственных данных в НСУД, обеспечение актуальности и достоверности данных, за редким исключением, осуществляется операторами данных. Операторами данных являются органы и организации государственного сектора, федеральные органы исполнительной власти или иные организации, уполномоченные в соответствии с законодательством Российской Федерации, законами и иными нормативными правовыми актами субъектов Российской Федерации, муниципальными правовыми актами на формирование и ведение информационных ресурсов, содержащих государственные данные (рис.1).

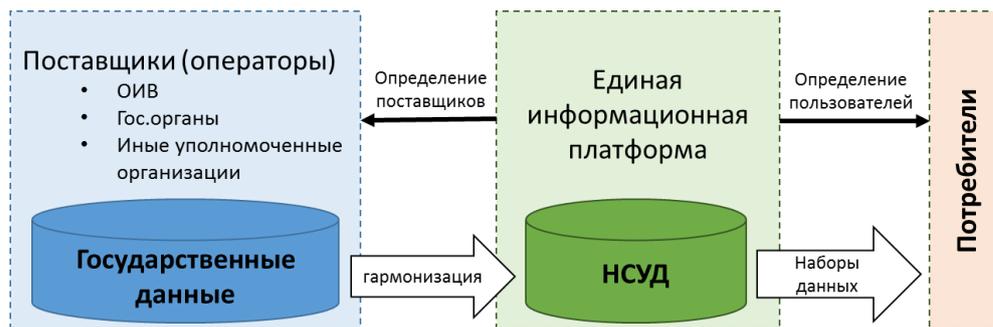


Рис. 1. Архитектура НСУД

Таким образом, с технологической точки зрения создание НСУД связано с решением проблемы логической интеграции, гармонизации и унификации управления разнородными данными, поступающими из различных источников, в режиме федеративного доступа. Построение подобной системы на национальном уровне является сложной масштабной задачей, решение которой предполагается опробовать на примере арктического макрорегиона. В данной статье предлагается краткая характеристика технологического и регионального (организационного) контекста реализации концепции НСУД, ее сравнение с существующими примерами разработки и реализации подобных проектов. Приведенный материал не претендует на полноту обзора затронутых вопросов, но лишь задает некоторые базовые ориентиры для потенциального позиционирования проблематики создания НСУД в современном контексте.

2. АЗРФ как тестовый полигон для отработки решений НСУД

Согласно докладу «Фост И.Д. Национальная система управления данными и реализация стратегии развития Арктической зоны Российской Федерации» [4], АЗРФ станет тестовым полигоном для отработки технических и организационных решений, необходимых для реализации НСУД. Выбор региона представляется не случайным. Арктика является примером макрорегиона, де-факто обладающего признаками самостоятельного объекта управления, но не имеющего статуса субъекта Федерации или иной административно-территориальной единицы (федерального округа). В то же время, территория Арктической зоны Российской Федерации явным образом определена как специфический макрорегион в рамках Указа Президента «О сухопутных территориях Арктической зоны Российской Федерации». После ликвидации Госкомсевера в 2001 году, Арктика рассматривается как самостоятельный объект государственного управления. Признаки «субъектности» макрорегиона проявляются в активной нормотворческой деятельности в отношении российской Арктики. В настоящее время, в большей или меньшей степени плодотворно, ведется разработка множества нормативных документов, регулирующих деятельность в АЗРФ, в том числе:

- федеральный закон «О развитии Арктической зоны Российской Федерации и о внесении изменений в отдельные законодательные акты Российской Федерации» (разработка закона, фактически, приостановлена [5]);

- федеральный закон «О государственной поддержке предпринимательской деятельности в Арктической зоне Российской Федерации» [6];

- стратегия «Арктика 2035» [7];

- серия государственных и международных стандартов, создаваемых в рамках технических комитетов Росстандарта «Исследования в полярных регионах» (ТК-187 и МТК 555) [8].

В то же время фактическое управление развитием АЗРФ, несмотря на делегирование координирующих функций Минвостокразвития, осуществляется опосредованно через различные ОИВ. Это естественным образом отражается и на организации информационного обеспечения развития Арктики. В настоящее время каждый орган власти, так или иначе задействованный в процессах управления макрорегионом, строит процессы своего информационного взаимодействия самостоятельно в разрезе отдельных функций. Права на доступ к данным регулируются множеством отдельных несинхронизированных нормативно-правовых актов (рис.2) [4]. Целевой ситуацией является формирование единых правил доступа к данным, общей экосистемы управления данными.

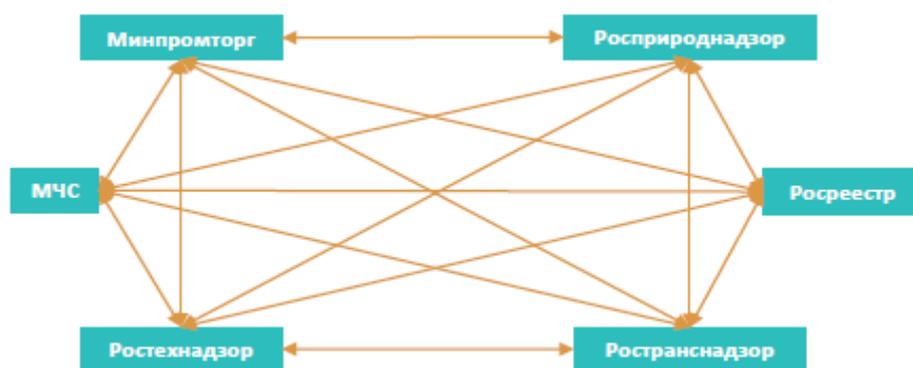


Рис.2. Взаимодействие госорганов в контексте управления данными в АЗРФ*

Таким образом, в АЗРФ отражены характерные проблемы информационного обеспечения управления территориями в российских организационно-правовых условиях.

3. Примеры крупномасштабных проектов создания интегрированных систем накопления и управления разнородными данными

Проблематика унификации процессов генерации, интеграции и управления данными, разнородными как в техническом и семантическом, так и в организационном смыслах (формируемыми различными, не обязательно взаимосвязанными, организациями) не является новой с одной стороны, и является предметом активной разработки в настоящее время – с другой. Проекты, связанные с интеграцией и унификацией управления разнородными данными из

* Фрагмент презентации «Фост И.Д. Национальная система управления данными и реализация стратегии развития Арктической зоны Российской Федерации» [4]

множественных источников, создаются и реализуются на корпоративном, отраслевом и государственном уровнях.

Российские проекты создания единых государственных информационных систем

В последние годы в Российской Федерации создано и находится в стадии разработки большое число государственных информационных систем, в различной мере реализующих архитектурные принципы и предполагаемую функциональность НСУД и имеющих отраслевую или предметную специализацию. Некоторые примеры подобных систем:

- Единая государственная информационная система социального обеспечения (ЕГИССО), позволяющая получать гражданам и органам власти актуальную информацию о мерах социальной поддержки, оказываемых из бюджетов всех уровней. Архитектура ЕГИССО в значительной мере соответствует концепции НСУД и включает несколько функциональных модулей: кабинет поставщика информации; кабинет органа, назначающего меры социальной поддержки; кабинет аналитика; личный кабинет гражданина. 19 ноября 2019 г. система выведена на штатный режим работы [9].

- Федеральная государственная информационная система в сфере здравоохранения (ЕГИСЗ). До вступления в силу Федерального закона от 29.07.2017 г. № 242-ФЗ «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам применения информационных технологий в сфере охраны здоровья» система имела иерархическую структуру, включающую централизованные и прикладные компоненты органов и организаций здравоохранения федерального и регионального уровней. В настоящее время ЕГИСЗ, региональные информационные системы здравоохранения и системы медицинских организаций рассматриваются как самостоятельные ИС, при этом федеральная система должна обеспечивать согласованное функционирование всех других информационных систем в здравоохранении за счет единых протоколов и принципов информационного обмена [10]. Архитектура системы включает несколько подсистем, в том числе различные реестры (организаций, документов, и т.п.), а также информационно-аналитические модули.

- Система межведомственного электронного взаимодействия (СМЭВ), позволяющая «федеральным, региональным и местным органам власти, кредитным организациям (банкам), внебюджетным фондам, и прочим участникам СМЭВ обмениваться данными, необходимыми для оказания государственных услуг гражданам и организациям, в электронном виде» [11]. Система была создана в рамках Федеральной целевой программы «Электронная Россия» для решения проблемы разрозненности межведомственных информационных систем (электронных сервисов), а также с целью обеспечения их взаимодействия [12]. Система решает относительно узкую, но очень важную технологическую задачу обеспечения взаимодействия между различными операторами государственных данных в условиях их организационной разнородности и с необходимым уровнем защищенности.

Кроме приведенных примеров интегрированных информационных систем, инициатива создания которых исходит с федерального уровня,

существует большое число частных проектов и концепций в разной степени воплощенных на практике, выдвигаемых и инициируемых отдельными экспертами, бизнес-структурами или локальными администрациями (см., например, [13, 14, 15]).

Международные проекты создания интегрированных систем формирования и управления данными

За рубежом наибольшее число концепций и проектов создания систем интеграции данных реализуется в научной сфере. Изначально, подобные проекты развивались, в основном, в рамках отдельных научных дисциплин. В настоящее время, с возрастанием объемов уже накопленных и темпов появления новых научных данных, актуальным стал вопрос их интеграции и совместного использования. Накопление больших объемов «сырых» данных в различных областях деятельности, в том числе — научной, и успехи в развитии методов их интеллектуальной обработки привели к появлению «data-driven» (основанного на данных) подхода к проведению исследований.

В 2015 г. в целях создания единой инновационной экосистемы больших данных Национальный научный фонд США инициировал программу создания Национальной сети региональных инновационных центров больших данных (BD Hubs), в рамках которой были созданы четыре территориальных центра больших данных, охватывающих в совокупности всю территорию Соединенных штатов. Программа нацелена на формирование сообществ исследователей и практиков, проводящих основанные на больших данных исследования, включающих органы власти городов, округов и штатов, местные промышленные, некоммерческие и исследовательские организаций [16]. В рамках программы поддерживается как формирование коллабораций на базе принципов государственно-частного партнерства, так и проекты, нацеленные на формирование технологической инфраструктуры больших данных. В качестве примера можно привести «Open Storage Network» (OSN) - проект, нацеленный на создание кибер-инфраструктуры для решения задач хранения, передачи, совместного использования и доступа к данным в режиме федеративного доступа [17]. Архитектура сети представлена на рис. 3.

Еще одним примером национального проекта интеграции научных данных является Австралийская национальная служба данных (Australian National Data Service, ANDS). ANDS обеспечивает поддержку публикации, обмена, поиска, совместного и повторного использования данных [18].

Наиболее масштабным европейским проектом на сегодняшний день является «Европейская облачная инициатива» (The European Cloud Initiative), финансируемая в рамках программы Horizon 2020 [19]. Инициатива* нацелена на укрепление позиций Европы в области инноваций, основанных на данных, повышение конкурентоспособности и согласованности соответствующих решений и создание единого цифрового рынка в Европе. В рамках инициативы предполагается создать инфраструктуру хранения и управления данными мирового уровня, обеспечить исследовательское сообщество, бизнес и государственные службы высокоскоростными коммуникациями и средствами

* В данном случае прямой перевод этого англоязычного термина означает «новый подход, стратегия, комплекс мероприятий».

обработки данных высокой производительности. Это поможет в полной мере использовать преимущества больших данных, предоставляя возможность беспрепятственно перемещать, совместно и повторно использовать междисциплинарные разнородные данные в трансграничном режиме в рамках глобальных рынков. Предполагается, что это даст импульс к развитию инновационного предпринимательства, малого и среднего бизнеса, исследований и разработок и даже появлению новых индустрий.

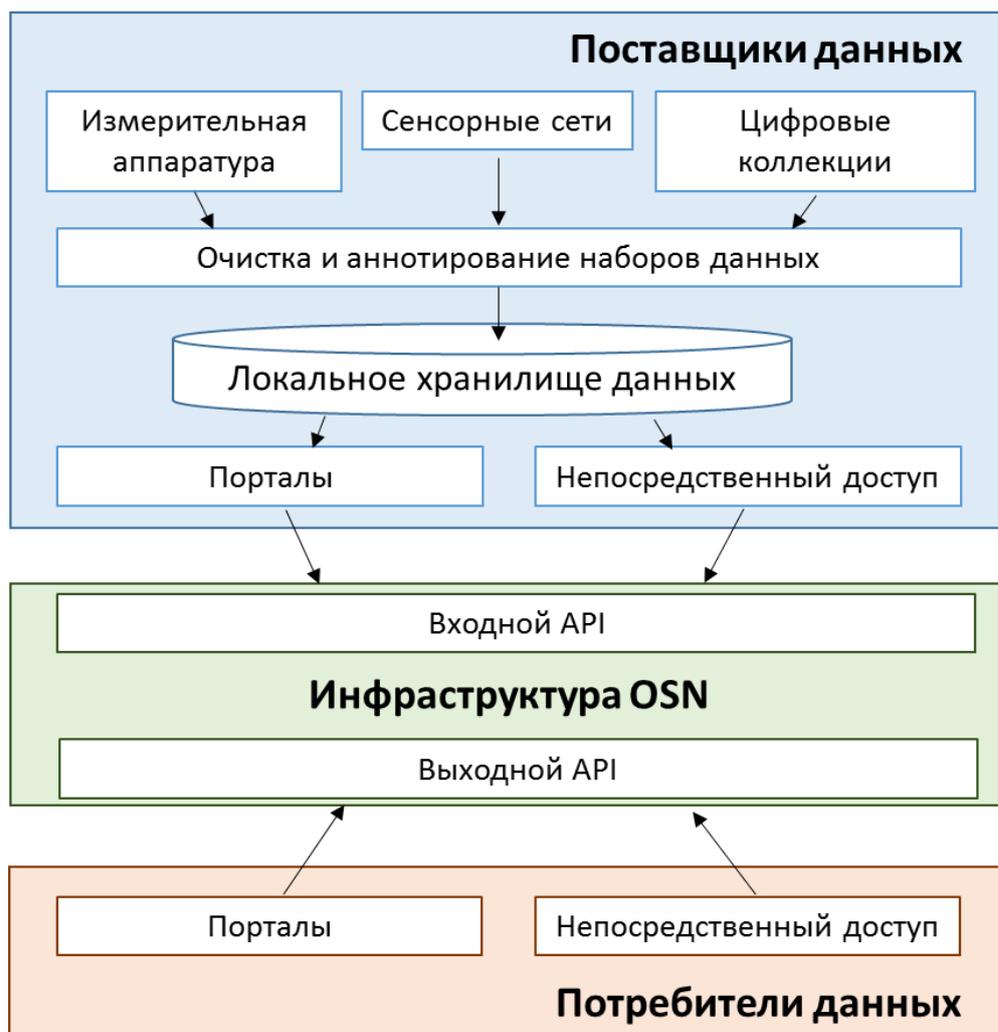


Рис. 3. Архитектура Open Storage Network

Ключевым компонентом инициативы является создание Европейского открытого научного облака, European Open Science Cloud (EOSC) — цифровой платформы для научного сообщества, обеспечивающей беспрепятственный доступ к надежным данным и функционально совместимым сервисам, охватывающим весь жизненный цикл научных данных, от обнаружения и добычи до хранения, управления, анализа и повторного использования [20]. Облако, при этом, должно охватывать разнородные научные данные вне зависимости от того,

в рамках каких государств и научных дисциплин они получены. В рамках программы Horizon 2020 на создание европейского научного облака в течение 2014-2020 гг выделено 600 млн. Евро. Облако является «зонтичным» проектом, объединяющим множество ранее созданных хранилищ научных данных и аналитических сервисов. На момент написания статьи в рамках веб-портала EOSC обеспечивался доступ к 234 службам хранения и обработки научных данных, реализованных в различных организациях странах Европы [21].

EOSC предполагает использовать федеративную модель интеграции ресурсов, сохраняющую независимость исходных ресурсов (данных и сервисов), но обеспечивающую к ним унифицированный доступ посредством федерализирующего ядра, реализуемого в рамках проекта «EOSC-hub» [22]. Модель включает в себя шесть направлений деятельности: (1) архитектура, (2) данные, (3) услуги, (4) доступ и интерфейсы, (5) правила и (6) управление (рис. 4). Модель предполагает, что элементы федеративной инфраструктуры данных будут функционировать в соответствии с принципами FAIR (см.ниже) и аккредитовываться/сертифицироваться на соответствие FAIR. Это означает, что службы данных, входящие в EOSC, будут со временем соответствовать определенным инфраструктурным стандартам и стандартам качества в рамках схемы обеспечения качества. Для выработки практических решений по реализации принципов FAIR на протяжении всего жизненного цикла научных данных 1 марта 2019 года был запущен проект «FAIRsFAIR» с финансированием около 10 млн. Евро на период до февраля 2022 года. В проекте предполагается выработать подходы к развитию культуры данных FAIR, изучению и внедрению передовых практик создания источников данных, соответствующих принципам FAIR [23].

Архитектура	Архитектура федеративной инфраструктуры, решающая проблему имеющейся фрагментированности инфраструктур научных данных и их недостаточной интероперабельности
Данные	Общий язык описания данных для обеспечения трансграничного/интердисциплинарного управления данными на основе принципов FAIR.
Сервисы	Доступные с точки зрения пользователя сервисы. Развитая среда, предлагающая широкий спектр сервисов для удовлетворения потребностей пользователей.
Доступ и интерфейсы	Механизмы / интерфейсы для доступа к EOSC. Простые способы удовлетворения обязательств по открытым данным или получения доступа к научным данным в различных дисциплинах.
Правила	Правила участия для разных акторов EOSC. Обеспечение соблюдения существующих правовых и технических ограничений и повышения правовой определенности и доверия.
Управление	Управление EOSC с целью обеспечения лидерства ЕС в “data-driven” науке, требующет новых структур управления.

Рис. 4. Модель Европейского открытого научного облака

Принципы FAIR-данных

В 2016 году сообщество заинтересованных исследователей и организаций были сформулированы так называемые принципы FAIR-данных [24, 25]. Перечисленные принципы распространяются как на данные и метаданные,

так и на обеспечивающую инфраструктуру. Специальный акцент делается на пригодность данных к автоматизированной машинной обработке (то есть на способность вычислительных систем находить, получать доступ, взаимодействовать и повторно использовать данные без какого-либо или минимального вмешательства человека):

- Findability («обнаруживаемость»): метаданные и данные должны быть легко доступны как для людей, так и для компьютеров. Поэтому для автоматического обнаружения наборов данных и сервисов необходимы машиночитаемые метаданные.

- Accessibility (доступность): после того, как требуемые данные обнаружены, необходимо обеспечить доступ к ним с помощью открытых универсальных протоколов, обеспечивающих, в том числе авторизацию и аутентификацию. Предполагается также, что метаданные остаются доступными даже после того, как доступ к собственно данным прекращен.

- Interoperability (совместимость): должна обеспечиваться возможность интеграции данных из различных источников, а также их использование приложениями для анализа, хранения и обработки.

- Reusability (возможность повторного использования): это ключевой принцип в концепции FAIR; данные должны быть описаны достаточно полно и точно, чтобы обеспечить возможность формирования на их основе различных наборов, применимых для решения прикладных задач. Кроме того, должны быть определены ограничения (лицензии) на использование данных, данные должны быть ассоциированы с их источником и представлены в принятом (стандартном) для соответствующей предметной области виде.

Центральный компонент технической экосистемы FAIR - цифровые объекты FAIR, которые могут представлять данные, программное обеспечение, протоколы или другие ресурсы. Каждый объект должен иметь постоянный идентификатор (PID) и снабжен достаточно богатыми метаданными, необходимыми для того, чтобы объект можно было надежно найти, использовать и процитировать. Кроме того, данные должны быть представлены в общедоступных, в идеале – открытых, форматах и должны быть тщательно документированы с использованием стандартов метаданных и словарей, принятых соответствующим исследовательским сообществом для обеспечения возможности взаимодействия и повторного использования (рис. 5). В соответствии с этим, выделяются две приоритетные проблемные области разработки экосистемы FAIR:

- 1) разработка, уточнение и принятие общих словарей, онтологий, спецификаций метаданных и стандартов, которые являются центральными для обеспечения взаимодействия и повторного использования;
- 2) широкое внедрение качественных профессиональных методов управления данными, хранилищ данных и соответствующих сервисов.



Рис. 5. Модель цифрового объекта FAIR*

В концепции FAIR особо отмечается, что создание технической экосистемы FAIR не может быть осуществлено с помощью чисто нисходящего (архитектурного) или восходящего (органического, основанного на спецификациях) подходов; они должны быть объединены.

Таким образом, следует заключить, что НСУД имеет множество аналогов, как в нашей стране, так и за рубежом, со схожей целевой функциональностью и в которых реализуются схожие концептуальные и архитектурные принципы.

4. Заключение

Создание Национальной системы управления данными — актуальная задача, соответствующая современным потребностям и возможностям практического воплощения ‘data-driven’-подхода к государственному управлению. Эффективная организация сбора, хранения, доступа и обработки данных позволит значительно повысить качество управленческих и сопутствующих процессов, создать основу для будущих качественных изменений в социально-экономическом развитии территорий, основанных на современных методах обработки больших данных.

Арктическая зона Российской Федерации, обозначенная в качестве тестового региона для опробования подходов и технологий реализации НСУД, обладает, как объект управления, специфическими свойствами, характерными для российских регионов и страны в целом. Критическая важность Арктики в контексте стратегических планов развития РФ также повышает актуальность задач развертывания в макрорегионе перспективных технологий и систем информационного обеспечения управления и развития территорий.

Организационная и архитектурная концепция, заложенная в НСУД, имеет большое количество аналогов как в России, так и за рубежом. Общим свойством реализуемых в РФ и других странах программ и проектов создания крупных

*Сокращенная русскоязычная версия иллюстрации из [26]

интегрированных инфраструктур данных является использование федеративного подхода к управлению данными и ориентация на использование стандартных организационно-технологических решений. Концептуальные требования к реализуемым в рамках подобных систем сервисам наиболее емко сформулированы в рамках концепции FAIR, принятой за основу в проектах создания инфраструктур больших данных национального масштаба в США, Евросоюзе и других странах. Таким образом, при реализации НСУД целесообразно максимально полно учитывать отечественный и зарубежный опыт создания подобных систем, изначально ориентироваться на использование российских и международных стандартов.

Литература

1. Паспорт национальной программы «Цифровая экономика Российской Федерации». Режим доступа: <http://static.government.ru/media/files/urKNm0gTPPnzJlaKw3M5cNLobgczMkPF.pdf>.
2. Постановление Правительства РФ от 3.06.2019 № 1189-р. «Концепция создания и функционирования национальной системы управления данными». Режим доступа: <http://static.government.ru/media/files/jYh27VIwiZs44qa0IXJZCa3uu7qqLzl.pdf>.
3. Кайль Я. Я. Взаимодействие участников процесса предоставления государственных (муниципальных) услуг на уровне субъекта Российской Федерации // Национальные интересы: приоритеты и безопасность. 2011. № 43.
4. Фост И. Д. Национальная система управления данными и реализация стратегии развития Арктической зоны Российской Федерации // Презентация на Круглом столе, 28.06.2019. Аналитический центр при Правительстве РФ.
5. Жуков М. А. Закон об Арктике. Дорога без конца // Информационно-аналитический портал «Арктика сегодня». Опубликовано: 22 января 2019. Режим доступа: <https://arctic-today.ru/index.php/arkticheskoe-zakonodatelstvo/243-akon-ob-arktike-doroga-bez-kontsa>.
6. Проект Федерального закона «О государственной поддержке предпринимательской деятельности в Арктической зоне Российской Федерации». – Режим доступа: <http://docs2.cntd.ru/document/554453249>.
7. Создаем стратегию развития Арктики. URL: <https://www.arctic2035.ru/>.
8. Технический комитет по стандартизации №187 «Проведение исследований в полярных регионах». Официальный сайт. – Режим доступа: <http://tk187.ru/>.
9. Портал ЕГИССО. <http://egisso.ru/site/>.
10. Гусев А. Обсуждаем постановление Правительства №555 о ЕГИСЗ. – Режим доступа: <https://www.kmis.ru/blog/obsuzhdaem-proekt-polozheniia-o-egisz/>.
11. Система межведомственного электронного взаимодействия. Режим доступа: https://ru.wikipedia.org/wiki/Система_межведомственного_электронного_взаимодействия.
12. Принципы взаимодействия органов государственной власти в Электронном Правительстве. Режим доступа: <http://www.ardecs.com/blog/25-03-2016/?lang=ru>.
13. Единая система информационных ресурсов официальных сайтов исполнительных органов государственной власти Санкт-Петербурга. URL: <https://netrika.ru/project/edinaya-sistema-informacionnyh-resursov-oficialnyh-saytov-ispolnitelnyh-organov-gosudarstvennoy-vlasti-sankt-peterburga>.

14. Национальная информационная система Российской Федерации (НИС РФ). Режим доступа: <https://instella.ru/nis-rf/>.
15. Селезнев А.Г. Развитие российской экономики в рамках цифровой экосистемы на Евразийском пространстве // Креативная экономика. 2019. Том 13, № 1. С. 21–36.
16. Big Data Regional Innovation Hubs (BD Hubs). NSF - National Science Foundation. Режим доступа: https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505185.
17. Núñez-Corrales S. et al. (2018) Open Storage Network: National Data Storage Cyberinfrastructure for the 21st. DOI: 10.13140/RG.2.2.31543.78249.
18. Australian National Data Service. Режим доступа: <https://www.andis.org.au/>.
19. The European Cloud Initiative. Режим доступа: <https://ec.europa.eu/digital-single-market/en/%20european-cloud-initiative>.
20. European Open Science Cloud. Режим доступа: <https://ec.europa.eu/digital-single-market/en/european-open-science-cloud>.
21. EOSC Marketplace. Режим доступа: https://marketplace.eoscportal.eu/services/c/storage?q=&service_id=&sort=_score
22. Implementation Roadmap for the European Open Science Cloud // European Commission. Brussels, 2018. Режим доступа: https://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf#view=fit&pagemode=none.
23. Fostering FAIR Data Practices in Europe. HORIZON 2020 project website. – Режим доступа: <https://cordis.europa.eu/project/rcn/222563/factsheet/en>.
24. Wilkinson M., Dumontier M., Aalbersberg I. et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 DOI:10.1038/sdata.2016.18.
25. FAIR Principles – GO FAIR. – URL: <https://www.go-fair.org/fair-principles/>.
26. Turning FAIR Into Reality. Final Report and Action Plan from the European Commission Expert Group on FAIR Data // Luxembourg: Publications Office of the European Union, 2018. DOI: 10.2777/1524.

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.), lead researcher of IIMM KSC RAS

Вицентий Александр Владимирович

к.т.н., старший научный сотрудник
e-mail: alx_2003@mail.ru

Vicentiy Alexander

PhD (Tech. Sci.), senior researcher

Куприков Никита Михайлович

к.т.н., младший научный сотрудник
e-mail: kuprikov@asms.ru

Kuprikov Nikita

PhD (Tech. Sci.), junior researcher

DOI: 10.25702/KSC.2307-5252.2019.9.158-171

УДК 004.6, 004.9

А. Л. Щур¹, А. М. Федоров^{1,2}, И. О. Датьев¹

¹ Институт информатики и математического моделирования ФИЦ КНЦ РАН

² Филиал ФГБОУ ВО МАГУ в г. Апатиты

ЦИФРОВИЗАЦИЯ В ГОСУДАРСТВЕННОМ И МУНИЦИПАЛЬНОМ УПРАВЛЕНИИ: МИРОВОЙ ОПЫТ, ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ОНЛАЙНОВЫХ СОЦИАЛЬНЫХ СЕТЕЙ

Аннотация

В статье рассмотрен процесс развития идей электронной демократии в мире, выраженных в появлении форм электронного правительства и становлении электронного взаимодействия между органами власти и населением. Представлена хронология изменения понятийной и компонентной баз цифровизации процесса управления, выделены специфические особенности применения онлайн-социальных сетей в этом процессе на муниципальном, региональном и государственном уровнях. Приведены примеры использования социальных сетей при управлении на различных уровнях. Описаны особенности, проблемы и перспективы использования онлайн-социальных сетей, в том числе при принятии решений в Арктической зоне Российской Федерации.

Ключевые слова:

Цифровизация, э-правительство, э-участие, муниципальное управление, социальные сети, АЗРФ.

A. L. Shchur, A. M. Fedorov, I. O. Datyev

DIGITALIZATION IN STATE AND MUNICIPAL GOVERNANCE: WORLD EXPERIENCE, PROBLEMS AND PROSPECTS FOR USING ONLINE SOCIAL NETWORKS

Annotation

The article discusses the process of developing ideas of e-democracy in the world, expressed in the emergence of forms of e-government and the establishment of electronic interaction between government and the public. A chronology of changes in the conceptual and component bases of digitalization of the management process is presented, specific features of the use of online social networks in this process at the municipal, regional and state levels are highlighted. Examples of the use of social networks for management at various levels are given. Features, problems and prospects of using online social networks are described, including the decision-making process in the Arctic zone of the Russian Federation.

Keywords:

Digitalization, e-government, e-participation, municipal government, social networks, AZRF.

Введение

Одним из самых заметных изменений в жизни современного общества, произошедших в начале 21 века, является повсеместное распространение цифровых технологий, случившееся не в последнюю очередь благодаря резкому росту скорости и доступности каналов спутниковой связи. Так называемый феномен *цифровизации*, изначально воспринимавшийся лишь в виде технических каналов обмена данными, с появлением технологий Web 2.0 быстро перерос свое изначальное предназначение и стал носить всеобъемлющий общественный характер. Благодаря мессенджерам, социальным сетям и другим сервисам,

позволяющим так или иначе осуществлять процесс не только хранения информации, но и свободного обмена ею между людьми, стало возможным говорить о появлении цифрового общества как отпечатка общества реального. Сегодня уже трудно вообразить человека, никак не представленного в онлайн-пространстве, и наоборот, с каждым днем становится все меньше людей, не использующих Интернет-ресурсы в качестве источника информации и социального взаимодействия.

Кроме социального сектора, *цифровизация общества* активно повлияла и продолжает влиять на другие сферы человеческой жизнедеятельности: экономическую и политическую.

Интернет-эра породила не только новые формы ведения бизнеса (онлайн-реклама, оплата товара через онлайн-сервисы), но и целые новые его направления, как, например, онлайн-магазины. В 1995 году Джефф Безос запустил Amazon.com, программист Пьер Омидьяр - AuctionWeb (eBay). Спустя 4 года в Китае была основана Alibaba Group. На мировых рынках появились новые лидеры с миллиардными капитализациями (Aliexpress, Amazon.com, eBay и другие), деятельность которых целиком основана на использовании информационно-коммуникационных технологий (ИКТ). На региональном и муниципальном уровнях широкое распространение получила интернет-коммерция на базе онлайн-социальных сетей, посредством создания специализированных групп-сообществ.

На сегодняшний день электронная коммерция является неотъемлемой составляющей мировой экономики и представлена такими направлениями как электронная торговля (e-trade), электронные деньги (e-cash), электронный маркетинг (e-marketing), электронный банкинг (e-banking), электронные страховые услуги (e-insurance).

Последней пришла в Интернет большая политика. Изначально государственный аппарат большинства стран мира и политические деятели не воспринимали онлайн-сервисы в качестве серьезного социологического фактора. Но по мере вовлечения в цифровое общество все большего числа людей, социальные сети также начали использоваться как политические арены – для диспутов и выдвижения политических кандидатур. Ярким примером, задавшим тон широкому политическому применению интернет-ресурсов, в том числе социальных медиа, стала президентская кампания Барака Обамы в 2008 году. Многие ее наблюдатели отметили, что именно онлайн-деятельность предвыборного штаба демократов (куда входил Крис Хьюз, один из сооснователей социальной сети Facebook), активно использовавшая Twitter, Facebook и YouTube, внесла большой вклад в победу Обамы на выборах [1].

Следующий американский президент Дональд Трамп с еще большей эффективностью использовал возможности социальных сетей в своей избирательной кампании 2016 года. Как заявляют эксперты, именно это позволило ему одержать победу, в которой многие сомневались [2].

Таким образом, за последние полтора десятилетия цифровое общество, представленное в онлайн-масс-медиа, постепенно превратилось из отпечатка реального общества в его полноценную компоненту, влияющую на реальность в той же мере, в какой реальные события влияют на онлайн.

В данной статье предпринята попытка сопоставить мировой и российский опыт развития электронной демократии, а также оценить текущую и

потенциальную роль онлайн-социальных сетей в государственном и муниципальном управлении. Отдельное внимание уделено особенностям и перспективам использования электронных технологий в управлении имеющих региональную специфику территорий, таких как Арктическая зона Российской Федерации.

Становление принципов электронной демократии в мире

Первые попытки использования информационных технологий в качестве инструмента принятия решений начались в обществе еще до появления сети Интернет. Еще в 70-х и 80-х годах прошлого века в США была проведена целая серия экспериментов «теледемократии», целью которых было объединение воедино ранних домашних компьютеров, телекоммуникаций и интерактивного кабельного ТВ.

В 1972 году был проведен эксперимент MINERVA, предназначенный для того, чтобы «массы граждан могли вести дискуссии друг с другом, и который позволил бы им принимать групповые решения, не покидая своих домов и не собираясь в огромном зале». Система включала в себя телефонные конференции, радио, двухстороннее кабельное телевидение и спутники.

В 1980-х годах прошел ряд экспериментов в Гонолулу, на Гавайях и в Южной Калифорнии, в ходе которых исследователи звонили случайным группам граждан по телефону, предлагали изучить брошюру, содержащую некую политическую информацию и несколько точек зрения на нее, а затем попросили проголосовать за наиболее понравившуюся.

Лишь малая часть этих проектов получила поддержку и была интегрирована в конституционные механизмы государственного управления. Но они стимулировали творческое мышление среди политических теоретиков на тему предполагаемой взаимосвязи между интерактивными коммуникационными технологиями и новыми формами участия в демократическом управлении [3].

Первое официально признанное использование современных цифровых технологий для принятия решений на государственном уровне относится к декабрю 1999 года. Тогда парламент Шотландии впервые в мировой практике согласился рассмотреть электронную петицию от Всемирного фонда дикой природы, созданную путем электронного сбора имен и адресов в процессе обкатки системы e-Petitioner, разрабатывавшейся для нужд парламента.

Система была создана в рамках договоренности между Интернациональным центром теледемократии на базе Единбургского университета и Комитетом по публичным петициям парламента. Соглашение позволило обеим сторонам начать оценку использования и гражданского воздействия электронных петиций в Шотландии. Эксперимент продлился с декабря 1999 года до 29 февраля 2000-го, собрав 305 действительных подписей и 9 действительных комментариев. После первоначального успеха системы, Комитет по публичным петициям предложил доработать ее и предоставить к ней доступ через официальный сайт парламента.

В последующие годы опыт Шотландского парламента стали перенимать и другие управляющие органы. В 2004-2006 годах усовершенствованные версии e-Petitioner появились в нескольких городских и областных администрациях Великобритании, а в 2005-м году немецкий Бундестаг в сотрудничестве с

Шотландским парламентом также ввел у себя аналогичную систему электронных петиций [4].

Изучение возможностей ИКТ и оформившихся цифровых социальных медиа привело к появлению в 2009 году меморандума ООН, в котором была официально зафиксирована (уже существовавшая до этого) терминология, обозначены основные принципы электронной демократии и даны рекомендации по ее развитию в странах-участницах.

Следует отметить, что изначально меморандум не разграничивал понятия «электронное правительство» и «электронная демократия». Но эти термины, хотя и взаимосвязаны, по своему назначению несколько отличаются. В первом случае речь идет о цифровой, т.е. представленной в онлайн-пространстве версии существующих органов власти: e-parliament, e-legislation, e-justice и т.д. (э-парламент, э-законодательство, э-правосудие, соответственно). А во втором — о различного рода взаимодействиях с этими органами власти: e-election (э-выборы), e-referendum (э-референдум), e-initiative (э-инициативы), e-voting (э-голосование), e-consultation (э-консультации), e-petitioning (э-петиции), e-campaigning (э-кампании) и пр [5].

Конкретизация взаимоотношения терминов произошла несколько позже: в резолюции ООН, принятой в 2015 году в качестве основной программы действий до 2030 года, электронное правительство было названо частью электронной демократии, а электронная демократия – частью ЦУР (целей устойчивого развития) [6].

Электронное правительство и электронное участие

Дальнейшие исследования Департамента по экономическим и социальным вопросам ООН позволили еще четче очертить аспекты и функции составляющих частей электронной демократии, а именно электронного правительства (e-government) и так называемого «электронного участия» (e-participation).

Под электронным правительством понимается создание специализированных сайтов государственных структур, обеспечения возможности подачи заявок на них и получения государственных услуг онлайн. Департамент ООН выделил четыре основных сферы электронного взаимодействия:

- между органами государственной власти (G2G – Government-to-government);
- между правительством и госслужащими (G2E – Government-to-employees);
- между правительством и гражданами (G2C – Government-to-citizens);
- между правительством и бизнесом (G2B – Government-to-business).

Постепенно был конкретизирован и принцип e-participation, т.е. принцип электронного взаимодействия граждан с органами власти. Его основными подпунктами были названы:

- e-inform – электронное информирование граждан о действиях органов власти;
- e-consult – действия консультативного характера, выясняющие мнение населения по поводу принимаемых органами власти решений;
- e-decision-making – вовлечение граждан в совместный с органами власти процесс разработки и принятия решений.

Более четкое понимание принципов построения э-государства позволило на их основе разработать Electronic Government Development Index (EGDI) — индекс развития электронного государства, который на данный момент используется в отчетах ООН для сравнительного анализа инициатив по интеграции э-демократии в 193 странах мира. Частью этих отчетов является и Electronic Participation Index (EPI), выстраиваемый на основе данных по степени вовлеченности граждан в электронное взаимодействие с органами власти. В 2018 году лидерами рейтинга были названы Дания, Австралия и Южная Корея [7]

Э-демократия в России

Развитие идей электронного правительства в России пошло по несколько иному пути. В его основу легли «Стратегия развития информационного общества в Российской Федерации», утвержденная Президентом Российской Федерации от 7 февраля 2008 г. № Пр-212, и государственная программа «Информационное общество (2011-2020 годы)», утвержденная распоряжением Правительства Российской Федерации от 20 октября 2010 г. № 1815-р, с учетом международных соглашений Российской Федерации [8].

Основной упор российской программы сделан на:

- обеспечение населения качественными и доступными услугами связи и доступа в Интернет;
- предупреждение угроз в информационном обществе и обеспечение информационной безопасности;
- реализацию в электронной форме полномочий государственных и муниципальных органов власти и предоставление их услуг гражданам, а также повышение качества государственного управления и оперативности взаимодействия органов государственной (муниципальной) власти, граждан и организаций [9].

С 2014 по 2019 годы в Государственную программу вносились изменения в соответствии с приоритетами, определенными указами Президента Российской Федерации, но основная ее структура и направленность сохранились.

Следует отметить, что выполнение данной программы принесло ощутимые результаты. Например, уровень представленности в электронном пространстве всех агентств федерального уровня власти был доведен до 100 %. Была создана и налажена единая система межведомственного электронного взаимодействия, которая объединила базы данных различных государственных служб, банков, внебюджетных организаций и прочих участников. Это позволило освободить граждан от необходимости самостоятельного сбора документов при обращении в государственные органы. Также были введены в широкую эксплуатацию государственные порталы, в том числе «Цифровое правительство» (digital.gov.ru) и «Госуслуги» (www.gosuslugi.ru).

отдавали созданию специализированных веб-ресурсов и сервисов. Примечателен пример профинансированной Европейской комиссией парламентской инициативы Africa4All, целью которой было вовлечение населения пяти стран Африки в обсуждение важных для континента ключевых проблем, таких как искоренение бедности, улучшение уровня здоровья наций, гендерное равноправие и т.д.

Эксперимент был признан неудачным ввиду слабой активности участников: некоторые разделы сайта, специально созданного для обсуждения проблем, получили лишь по одному или нескольким комментариям в течение трех лет работы инициативы [11]. Данный пример наглядно демонстрирует одну из ключевых проблем электронного участия, особенно остро стоявших в 2000-х годах и названных в литературе «электронным разделением» [12]. Суть ее заключается в неравномерном распределении среди населения доступа к средствам ИКТ и выхода в Интернет на основании социо-экономических и географических факторов.

С другой стороны, положительные примеры использования социальных медиа в странах с высоким уровнем проникновения ИКТ начали отмечаться уже на рубеже 2009–2010 годов. Например, использование Facebook, Twitter и YouTube помогли полиции Квинсленда (Австралия) быстро оповещать, собирать информацию и координировать свои действия с населением во время внезапных наводнений в 2010 году, вызванных серией тропических штормов [13]. Еще одним характерным примером можно назвать случай с эпидемией гриппа H1N1 в США, где медицинские службы использовали Twitter для распространения информации о болезни среди населения и взамен получали сообщения очевидцев, позволявшие точнее выявлять зону поражения и вовремя принимать меры [14]. Подобные случаи позволили однозначно признать социальные сети и микроблоггинг-платформы эффективными инструментами контроля и управления в экстренных ситуациях [15].

Еще одной сферой электронного участия, где на сегодняшний день социальные сети играют немаловажную роль, являются демократические выборы и политическое просвещение населения. Однако в данной статье мы не будем углубляться в особенности использования социальных медиа подобным образом и сосредоточимся на электронном участии, взаимодействующем со сформировавшимися органами власти.

Э-правительство и социальные сети: особенности использования

Одним из ключевых аспектов эффективной работы электронного правительства является доверие граждан к органам, осуществляющим управление, что, в свою очередь, требует прозрачности их структуры и действий для стороннего обывателя [16]. Для осуществления этой задачи практически все органы власти как на местном, так и на федеральном уровне в современных развитых государствах предпринимают усилия для полноценной презентации своих департаментов не только на специализированных государственных сайтах, но и в популярных мировых социальных медиа. Обусловлено это очевидными причинами – государственные сайты обладают гораздо более низким уровнем популярности по сравнению с социальными сетями, а также закрытой

структурой, зачастую не позволяющей вести открытый диалог с конкретными чиновниками по интересующим вопросам.

Существует немало случаев, когда активное присутствие представителей органов муниципальной власти в соцсетях позволило значительно повысить уровень эффективности принятия решений за счет прямой и постоянной связи с гражданами. Показательным является пример аккаунта мэра Сеула Пака Вонсуна в микроблоггинг-сервисе Twitter, где он напрямую общается с населением города и оперативно реагирует на сообщения от граждан о проблемах дорожного и бытового характера.



Твиттер мэра Сеула: отчет о посещении района города
в ответ на жалобы о ржавой воде из-под крана

Исследование с использованием алгоритмов анализа социальных сетей продемонстрировало, что его аккаунт выступил центром и связующим звеном между несколькими кластерами аккаунтов жителей города и представителей муниципальной власти. Это позволяет говорить о широком охвате его деятельности и прямом вовлечении населения в процесс совместного принятия решений [17].

Но активное использование социальных сетей представителями власти не всегда ведет к положительному результату. Нецелевое применение данного

ресурса, как правило, направлено не на решение рабочих вопросов и диалог с гражданами, а на укрепление личных политических позиций и прямой пиар. Исследование 2015 года, объектом которого были государственные структуры Косово, выявило, что две трети чиновников, стоящих во главе государственных департаментов страны, гораздо активнее использовали личные аккаунты для популистских кампаний и саморекламы, чем для реальной работы и общения с населением [18].

Существует и обратная проблема общения граждан и чиновников посредством онлайн-медиа: ожидание немедленных результатов после обращения к представителям власти. Для социальных сетей характерна очень высокая скорость обсуждения и реакции, а также нередкая поляриность мнений. Эти факторы зачастую приводят к ожиданию схожей реакции и от политических сил, с которыми ведется диалог. Невозможность представителей власти реагировать с такой же скоростью, с какой пользователи выдвигают им свои предложения, вызывает порой резкую критику действий чиновников. Характерным примером подобной неоднозначной реакции является ситуация в Египте в 2011 году, когда назначенный после февральской революции новый министр финансов подвергся жесткой критике в Twitter и обвинениям в бездействии после всего двух дней пребывания на своем посту. На четвертый день ему даже пришлось выступить на национальном телевидении, чтобы публично объяснить необходимость проанализировать сложившуюся ситуацию, прежде чем предпринимать какие-то действия [19].

В то же время, следует учитывать, что подобная разнородность и активность выражаемых мнений является прямым следствием свободы слова, предоставляемой большинством социальных онлайн-медиа. В условиях жестких рамок и прямой модерации уровень открытости дискуссии заметно снижается, как показывает изучение поведения респондентов в китайской социальной сети Weibo [20]. Данный факт также можно приписать к недостаткам использования специализированных государственных сайтов как базы для диалога и обсуждения предложений населения и совместного принятия решений. Однако свобода слова часто воспринимается как вседозволенность, которая будучи подкрепленной виртуальностью соцсетевой идентификации сильно понижает уровень дискуссии и уменьшает ее конструктивный уровень, что делает такую коммуникацию бесполезной для решения каких-либо вопросов.

Согласно Макинтош [4], электронное участие (e-participation) делится на три уровня: локальный, региональный и национальный. Текущий опыт использования онлайн-социальных сетей подтверждает, что эффективное взаимодействие с органами власти посредством этих сетей происходит именно на локальном и региональном уровнях. Это объясняется, прежде всего, физическим ограничением временного ресурса начальствующего органа по отношению к количеству обращений граждан, другими словами, эффективность одного человека при ведении диалога с гражданами обратно пропорциональна количеству граждан.

С другой стороны, чем локальнее проблема, тем меньшее количество людей требуется для обсуждения, выработки и принятия вариантов решения этой проблемы. Это второй момент, благодаря которому e-партисипейшн хорошо работает на муниципальном уровне.

Э-участие и социальные сети в Арктической зоне Российской Федерации (АЗРФ)

В силу своей специфики арктическая зона РФ и входящий в нее Кольский полуостров могут извлечь большую выгоду от использования возможностей онлайн-социальных сетей как инструмента для поддержки принятия решений на муниципальном и региональном уровнях. Суровые природные условия, затрудняющие перемещения, небольшое количество населения, его кластерное расположение и удаленность выводят на первый план методы дистанционной коммуникации, среди которых социальные сети играют очень важную роль.

Чтобы выявить потенциально перспективные варианты их использования, авторы статьи провели за прошедший год ряд предварительных разнонаправленных исследований.

В одном из них [21] была предпринята попытка представить способы, которыми можно использовать социальные сети (в частности, наиболее распространенную в России социальную сеть Вконтакте) в качестве источника данных при принятии органами управления решений на региональном уровне. Для этого был проведен эксперимент по сравнению количества пользователей соцсетей в городах и регионах страны с официальными данными Росстата по численности проживающего в них населения. Полученные результаты позволили говорить о тесной взаимосвязи данных показателей, а значит, и о возможности использовать получаемые из ВК цифры (после выработки соответствующих техник уточнения результата) в качестве оперативных субиндексов статистики по населению.

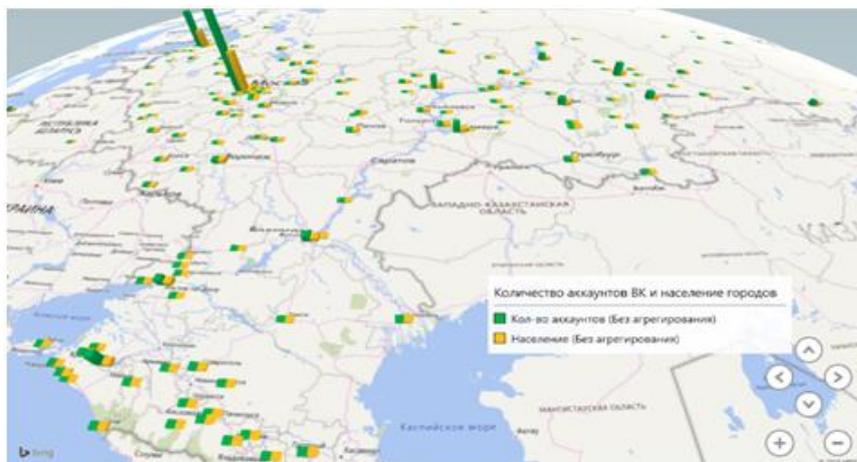


Рис. 3 Наглядное представление распределения и соотношения аккаунтов и населения в российских городах

В другом исследовании [22] основное внимание было уделено оценке представленности вышестоящих чиновников из органов регионального управления Мурманской области в социальных сетях Facebook, Вконтакте и Twitter. По результатам исследования выяснилось, что хотя сами по себе министерства областного правительства в полной мере представлены в соцсетях (т.е. цель электронного присутствия, поставленная в государственной программе

«Информационное общество (2011-2020 годы)», была достигнута), сами чиновники, в основном, либо не вели никакой серьезной онлайн-деятельности, связанной с работой, либо отсутствовали в социальных сетях вовсе. Единственным активно действующим лицом, предоставлявшим на ежедневной основе отчеты о проделанной работе и предлагавшим темы для народного обсуждения, оказался губернатор области Чибис А. В.

С именем губернатора также связана интересная в рамках исследования способов электронного участия инициатива - созданный весной 2019 года интернет-портал «Наш север». Портал представляет собой «сервис совместного управления, позволяющий всем жителям Мурманской области, не тратя времени на заполнение бумажных документов и длительный поиск инстанций, эффективно взаимодействовать с властью и реально влиять на развитие родного города и региона» [23]. На сайте присутствует возможность вносить жалобы и предложения по улучшению дорожных и жилищно-коммунальных условий в городах Мурманской области, а также голосовать в предлагаемых муниципальной властью проектах. Несмотря на отсутствие прямой связи с непосредственными начальниками городских служб, выставленные на рассмотрение вопросы и жалобы получают заключения созданных комиссий и, при необходимости, вносятся в список задач муниципальных органов управления.

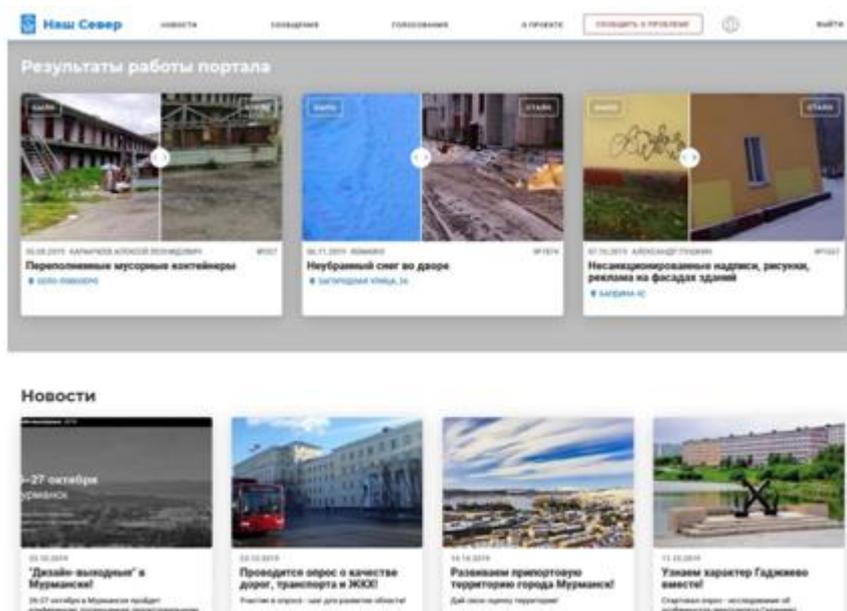


Рис 4. Результаты работы коммунальных служб на сайте «Наш Север»

Несмотря на ряд полезных функций портала, его основными недостатками, как и в случае с другими специализированными государственными сайтами, являются малая известность и отсутствие прямого диалога с принимающими решения лицами. За полгода работы на сайте зарегистрировались лишь 4700 человек при общей численности населения Мурманской области около 750 тысяч.

В то же самое время упомянутая выше деятельность самого губернатора в социальной сети ВКонтакте имеет более широкий отклик. Имея 30 тысяч подписчиков, страница губернатора под каждой публикацией получает в среднем от 50 тысяч просмотров и более 200 комментариев. Коммуникационная доступность и привлекательность социальной сети могла бы неплохо дополнить возможности официального портала «Наш север».

Надо отметить, что еще одним недостатком подобных ресурсов является слабая проработка подсистемы контроля и верификации исполнения предложений, полученных исполнительными органами в рамках электронного взаимодействия с гражданами. Например, в качестве отчетных результатов по одной из проблем с дорожным покрытием были представлены рекомендации обратившимся гражданам по самостоятельной подаче документов в суд и понуждению ответственных лиц к решению заявленной проблемы, что можно отнести к консультативным действиям, но нельзя назвать полноценным решением вопроса.

Очевидно, что такие результаты взаимодействия не могут устраивать граждан, обратившихся к власти ни электронными, ни какими-либо другими способами. Это позволяет говорить о необходимости учета при реализации технологий электронного взаимодействия, как и других информационных технологий, морально-этических принципов [22].

Заключение

На сегодняшний день, электронное участие на уровне совместного принятия решений официально не обозначено среди основных направлений развития информационного общества России. Однако инициативы, направленные на прозрачность принятия решений и вовлечение населения в процесс управления на муниципальном и региональном уровнях, продолжают разрабатываться и совершенствоваться. Данная тенденция представляется закономерным развитием уже достаточно распространенных функций электронной демократии: э-информирования и э-консультирования.

Онлайновые социальные сети, будучи уже устоявшимся репрезентативным элементом общества, могут выступать полноправным инструментом для органов управления, позволяющим не только оперативно получать отклик на принятые решения, но и предоставлять разноплановую социологическую информацию, необходимую в фазе определения путей решения проблемы, а также обеспечивать обществу прозрачность и инклюзивность в сам процесс принятия решений.

Более детальное изучение взаимосвязи между электронными формами управления и онлайн-социальными сетями позволит выявить наиболее эффективные способы взаимодействия как со стороны государственных представителей, так и со стороны простых граждан.

Вероятным решением проблемы совместного управления могут стать модели, методы, технологии и основывающиеся на них информационные системы, которые позволят выстроить «мост» между органами управления муниципального и регионального уровня (администрациями города, области) и социальными сетями. Что позволит гражданам не только четче видеть действия власти, но и активнее взаимодействовать с ней в удобной для обеих сторон форме.

Литература

1. Obama's win means future elections must be fought online // The Guardian. 7 ноября 2008 г. – URL: <https://www.theguardian.com/technology/2008/nov/07/barackobama-uselections2008>.
2. Как социальные сети помогли Трампу стать президентом. URL: <https://www.cossa.ru/152/145969/>.
3. Coleman S., Macintosh A., Schneeberger A. (2007) DEMO-net: Deliverable 12.3: eParticipation Research Direction based on barriers, challenges and needs. DEMO-net Consortium. University of Leeds, UK.
4. Macintosh, A. (2004) Characterizing E-Participation in Policy-Making. In: Proceedings of the 37 Annual Hawaii International Conference on System Sciences (HICSS-37), January 5-8, 2004, Big Island, Hawaii.
5. Electronic democracy (“e-democracy”): Recommendation CM/Rec(2009)1 and explanatory memorandum. Council of Europe Publishing, F-67075 Strasbourg Cedex, 2009, ISBN 978-92-871-6647-0.
6. Преобразование нашего мира: Повестка дня в области устойчивого развития на период до 2030 года. Резолюция Генеральной Ассамблеи ООН. 25 сентября 2015 года, A/RES/70/1.
7. United Nations E-Government Survey 2018. Department of Economic and Social Affairs, UNITED NATIONS, New York, 2018. URL: publicadministration.un.org.
8. Белов В. Г. Электронная демократия в современной России // Ценности и смыслы. 2012. №4 (20). URL: <https://cyberleninka.ru/article/n/elektronnaaya-demokratiya-v-sovremennoy-rossii> (дата обращения: 25.11.2019).
9. Государственная программа «Информационное общество». Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации. URL: <https://digital.gov.ru/ru/activity/programs/1/> (дата обращения: 25.11.2019).
10. Российская общественная инициатива (РОИ). URL: <https://www.roi.ru> (дата обращения: 19.11.2019).
11. МЕТЕР: Инструмент измерения и оценки готовности электронного правительства. URL: <https://publicadministration.un.org/en/Capacity-Building/Tools/METER>.
12. Compaine B. M. (2001) The Digital Divide: Facing a Crisis or Creating a Myth? Cambridge and London: MIT Press, 374 pp.
13. Ehnis C., Bunker D. (2012) Social Media in Disaster Response: Queensland Police Service - Public Engagement During the 2011 Floods. In ACIS 2012: Proceedings of the 23rd Australasian Conference on Information Systems, 3-5 Dec. 2012, [Geelong, Vic.], pp. 1-10.
14. Näkki, Pirjo et al. (2011) Social media for citizen participation. Report on the Somus project. VTT Technical Research Centre of Finland. VTT Publications, Espoo, v. 755., 131 pp.
15. Magro M. (2012) A Review of Social Media Use in E-Government // Administrative Sciences, 2012, vol. 2, 148-161 pp. doi:10.3390/admsci2020148.
16. Khan, S., Rahim N., Maarop N. (2019) A Review on Antecedents of Citizen’s Trust in Government Social Media Services, 3C Tecnología. Special Issue – January 2019, 28-2. DOI: 10.17993/3ctecno.2019.specialissue.10.

17. Eom S.-J., Hwang H., Kim J.(2016) How do social media make government more responsive?: Evidence from network analysis of Twitter use in Seoul Metropolitan City in Korea. *Government Information Quarterly*. Vol. 35, Issue 1. pp. 109–122
DOI: 10.1016/j.giq.2017.10.002.
18. Rexhepi A., Filiposka S., Trajkovic V. (2016) The Social Networking for eGovernment: An initial assessment of Web 2.0 tools used by the Kosovo Central Government. In: RTA-CSIT 2016, Proceedings of the 2nd International Conference on Recent Trends and Applications in Computer Science and Information Technology. Tirana, Albania, November 18–19, 2016. CEUR Workshop Proceedings. pp. 80–89.
19. McGrath K., Elbanna A., Hercheui M., Panagiotopoulos P., Saad E. (2012) Exploring the Democratic Potential of Online Social Networking: The Scope and Limitations of e-Participation. *Communications of the Association for Information Systems*. – Vol. 30, Article 16. – DOI: 10.17705/1CAIS.03016.
20. Medaglia R., Zhu D. (2017) Public deliberation on government-managed social media: A study on Weibo users in China. *Government Information Quarterly*. Vol. 34, Issue 3, September 2017, pp. 533-544.
21. Fedorov A.M., Datyev I.O., Shchur A.L., Oleynik A.G. (2019) Online Social Networks Analysis for Digitalization Evaluation. In: Silhavy R. (eds) *Software Engineering Methods in Intelligent Algorithms*. CSOC 2019. *Advances in Intelligent Systems and Computing*, vol 984. Springer, Cham.
22. Datyev I.O., Fedorov A.M., Shchur A.L., Bystrov V.V. (2019) Social Networking Services as a Tool for State and Regional Governance (On the Example of the Arctic Region). In: Silhavy R., Silhavy P., Prokopova Z. (eds) *Computational Statistics and Mathematical Modeling Methods in Intelligent Systems*. CoMeSySo 2019. *Advances in Intelligent Systems and Computing*, vol 1047. Springer, Cham.
23. Портал «Наш север». URL: <http://nashsever51.ru/>.

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

Датъев Игорь Олегович

к.т.н., старший научный сотрудник

e-mail: datyev@iimm.ru

Igor O. Datyev

PhD (Tech. Sci.), senior researcher

Щур Андрей Леонидович

аспирант, инженер по НТИ

e-mail: shchur@iimm.ru

Andrey L. Shchur

postgraduate, research assistant engineer

DOI: 10.25702/KSC.2307-5252.2019.9.172-186

УДК 622.276+550.8+519.6+66.01.001

В. К. Каржавин

Геологический институт ФИЦ КНЦ РАН

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ГЕОХИМИЧЕСКИХ СИСТЕМ И ПЕРСПЕКТИВНЫХ СХЕМ ПЕРЕРАБОТКИ МИНЕРАЛОВ

Аннотация

В работе представлен необходимый и достаточный набор зависимых и независимых компонентов для положительного решения сложных систем. Это позволило оценить всё многообразие исследуемого природного процесса, а также предложить варианты технологической схемы переработки минерального сырья. Основой таких исследований является чёткое формулирование и грамотная постановка задачи моделирования, корректный выбор математической модели и т.д. Математическое моделирование служит не просто способом наглядного и достоверного отображения фактических данных, но и является инструментом получения новой информации. На этом фоне проявляются основные тенденции физико-химического моделирования природных геохимических и химико-технологических систем и оценка условий протекания исследуемых процессов

Ключевые слова:

математическое моделирование, минимизации свободной энергии Гиббса, платиноиды, гранат, кианиты, наноалмазы, апатиты, эвдиалит, минеральное сырьё, pH, Eh, равновесные условия, давление, температура.

V. K. Karzhavin

MATHEMATICAL MODELING OF GEOCHEMICAL SYSTEMS AND PERSPECTIVE SCHEMES FOR PROCESSING MINERALS

Abstract

The paper presents the necessary and sufficient set of dependent and independent components for the positive solution of complex systems. This made it possible to evaluate the entire diversity of the investigated natural process, and also to propose options for a technological scheme for processing mineral raw materials. The basis of such research is a clear formulation and competent formulation of the modeling problem, the correct choice of a mathematical model, etc. Mathematical modeling is not just a way of visual and reliable display of actual data, but it is also a tool for obtaining new information. Against this background, the main trends of physicochemical modeling of natural geochemical and chemical-technological systems and the assessment of the conditions of the processes under study are manifested

Keywords:

Mathematical modeling, minimization of the Gibbs free energy, platinumoids, garnet, kyanites, nanodiamonds, apatite, eudialyte, mineral raw materials, pH, Eh, equilibrium conditions, pressure, temperature.

Введение

Математическое моделирование позволяет воспроизводить определённым образом подобранные условия физико-химической ситуации для изучения взаимоотношений между компонентами в сложных системах на основе термодинамических методов исследования. Ярким примером этому является

математическое моделирование на ЭВМ природных процессов, базирующихся на основе экспериментальной, эмпирической и термодинамической информации.

Термодинамическим исследованиям отводится важная роль при анализе процессов, сопровождаемых химическими и фазовыми превращениями в широком интервале температур и давлений. Для равновесных или локально равновесных процессов расчёт параметров состояния позволяет с относительно высокой точностью моделировать реальные условия и представлять информацию, которую пока невозможно произвести экспериментальным путём в лабораторных условиях. Равновесные расчёты (численный эксперимент) отражают лишь качественные тенденции. Условие равновесия в системе, приводящее к задаче условной минимизации, может быть представлено в эквивалентной форме, как условие определения степени протекания реакций в сложной системе. Для исследования промышленного процесса необходимо использовать термодинамику необратимых процессов через решение систем нелинейных уравнений с ограничениями в виде равенств и неравенств, которые часто не имеют точного аналитического решения.

Принято считать, что состав и термодинамические свойства равновесных продуктов однозначно зависят от элементарного состава, температуры и давления [1,2]. Автором показано, что система уравнений констант равновесия и материального баланса имеет одно единственное решение, отвечающее физическому смыслу. Это позволяет широко использовать термодинамические методы исследования для решения многочисленных задач. Например, оценка температуры и давления протекающих процессов в природной среде, при которых происходило образование минералов, является важной задачей для прогнозирования протекающих в природных условиях постмагматических (метаморфических) процессов. Для данной цели исследуется химический состав совместно присутствующих в природном образце минералов, и осуществляют расчёт на основании констант равновесия химических реакций. Наличие равновесия в исследуемой природной системе позволяет оценить исходные Р-Т параметры процессов образования минералов при магматических, метаморфических, метасоматических, гидротермальных и рудообразующих процессах методом TWQ [3].

Исследование природы метаморфических и метасоматических преобразований пород рудоносных интрузивов предполагается проводить при комплексном использовании современных методов оценки термодинамических и физико-химических условий их формирования. Обычно для оценки фазового состава, твёрдых, летучих компонентов и флюидного режима в целом используется метод минимизации свободной энергии Гиббса [4,5]. Такой подход в планируемых исследованиях представляется перспективным для поиска равновесных составов в природной системе, а также позволяет более обосновано оценивать Р-Т условия, в которых формировался исследуемый объект.

Компьютерный расчёт химических взаимодействий действительно имеет ряд несомненных преимуществ перед трудоёмким модельным экспериментом. Основой физико-химического моделирования является описание объекта и процессов, протекающих в нем с изменением его свойств, а также состава компонентов. При этом термодинамическое моделирование позволяет проследить поэтапный (например, в зависимости от Р-Т параметров) механизм преобразования в гомогенных и гетерогенных системах компонентов твердой

фазы и состава флюида. Именно поэтому математический аппарат химической термодинамики позволяет, на основе корректных термодинамических величин [6–8], проследить химические взаимодействия в многокомпонентных и многофазных системах.

Известно, что превращение исходных веществ в продукты химических реакций большинства исследуемых процессов осуществляется через ряд элементарных реакций, о механизме которых можно делать лишь различные предположения. Поэтому для расчёта векторного состава независимых компонентов создаваемых мультисистем проводится путём грамотного перерасчёта результатов химического анализа исследуемого объекта. На данном этапе подготовки необходимо учитывать наличие возможных реакций окисления-восстановления, растворения, гидролиза, комплексообразования, ионного обмена и т.д. Накопленный нами за годы исследований фактический материал, позволяет по новому осветить ряд вопросов: по минералогии, геохимическим процессам, термодинамическим условиям формирования минеральных парагенезисов в расслоенных горизонтах и т.д.

I. Платиноиды^X. Оценка состава флюида, равновесного с твёрдой платино-палладиевой фазой, осуществлялась путём реализации мультисистем для каждой метаморфической зоны. Для теоретических исследований были созданы физико-химические модельные мультисистемы. Например, исходная мультисистема для платино-палладиевых образцов Панского массива содержала 18 независимых компонентов (элементы):

Al-C-Ca-Cu-Fe-H-K-Mg-Na-Ni-O-Pd-Pt-S-Se-Si-Te-Ti.

Расчётная матрица мультисистемы была составлена из 34 зависимых компонентов силикатно-сульфидных фаз и флюидной фазы, включающей следующие компоненты: H₂O, H₂, O₂, CO, CO₂, S₂, SO₂, H₂S, Se₂, SeO₂, H₂Se, Te₂, TeO₂, H₂Te, CH₄, C₂H₆ [9].

В результате проведённого исследования (метод TWQ) было установлено, что для начального этапа раннего автометаморфического преобразования пород с пироксен-куммингтонитовыми парагенезисами (Орх-Cum №1) определены следующие термодинамические* параметры: T=623°C и P=6.15 кбар. Протекающие в данной системе химические реакции являются равновесными, например, при 623°C и давлении 6.15 кбар (рис. 1).

Последующая оценка P-T параметров метаморфических преобразований пород определялась аналогично при помощи программы TWQ.

Так ранняя плагиоклаз-актинолитовая ассоциация (Pl-Act № 2) осуществлялась при T=382°C и P=1.65 кбар. А парагенезис (Pl-Act-Ts №3) при T = 398°C и P = 2.6 кбар (рис. 2).

* Используемые в работе обозначения соответствуют R.Kretz [10].

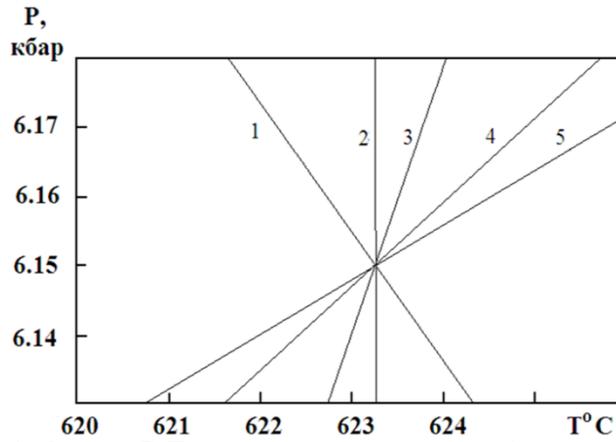


Рис. 1. Оценка P-T параметров автометаморфических ассоциаций (Orx – Cum)

- 1) $9 An + 3 Cum + 3 Di + 2 Mgt = 12 En + 3 Fs + 6 Czo + O_2$
- 2) $6 Cum + 2 Mgt = 21 En + 3 Fs + 6 H_2O + O_2$
- 3) $Cum + 2 Czo = An + Di + 3 En + 2 H_2O$
- 4) $21 An + 7 Di + 2 Mgt + 8 H_2O = Cum + 3 Fs + 14 Czo + O_2$
- 5) $18 An + 6 Di + 2 Mgt + 6 H_2O = 3 En + 3 Fs + 12 Czo + O_2$

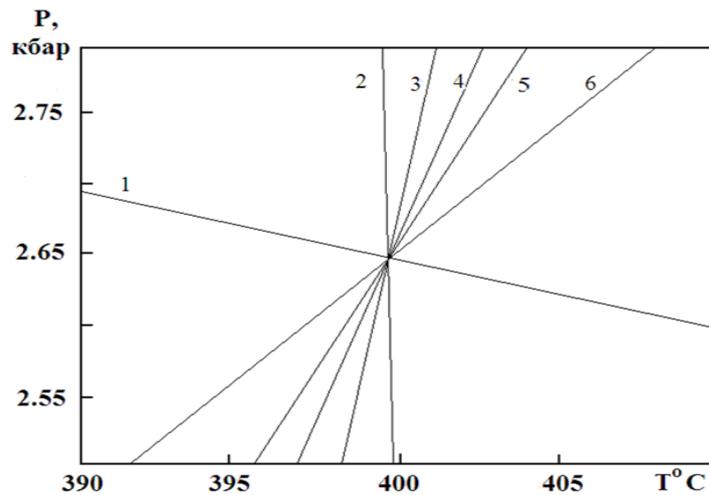


Рис. 2. Оценка P-T параметров автометаморфических ассоциаций (Pl-Act-Ts)

- 1) $5 Ann + 14 aQz + 5 Tsc + 12 cZo = 5 Phl + 28 An + 3 Act + 8 H_2O$
- 2) $4 An + 2 Mt + 2 Phl + 4 aQz + 2 H_2O = 2 Tsc + 2 Ann + O_2$
- 3) $12 cZo + 30 aQz + 3 Phl + 8 Mt = 3 Act + 3 Ann + 12 An + 3 Tsc + 4 O_2$
- 4) $12 cZo + 42 aQz + 9 Phl + 14 Mt + 6 H_2O = 3 Act + 9 Ann + 9 Ts + 7 O_2$
- 5) $24 cZo + 48 aQz + 10 Mt = 6 Act + 36 An + 6 H_2O + 5 O_2$
- 6) $6 Act + 84 An + 14 Mt + 24 Phl + 30 H_2O = 24 cZo + 24 Ts + 24 Ann + 7 O_2$

Преобразование пород с более поздними плагиоклаз-роговообманковыми парагенезисами (Pl-Hbl №4) протекало при температуре 473°C и давлении 4.3 кбар, а последующее образование диавторированных пород (Pl-Act-Chl №5) при температуре 370°C и давлении 1.16 кбар.

Физико-химическими модельными исследованиями установлено, что при достижении определенной величины парциального давления кислорода компонентный состав мультисистемы претерпевает некоторые изменения. Полученная информация о компонентном составе твердой фазы и флюида для различных стадий метаморфизма подробно представлена на рисунках 3 и 4.

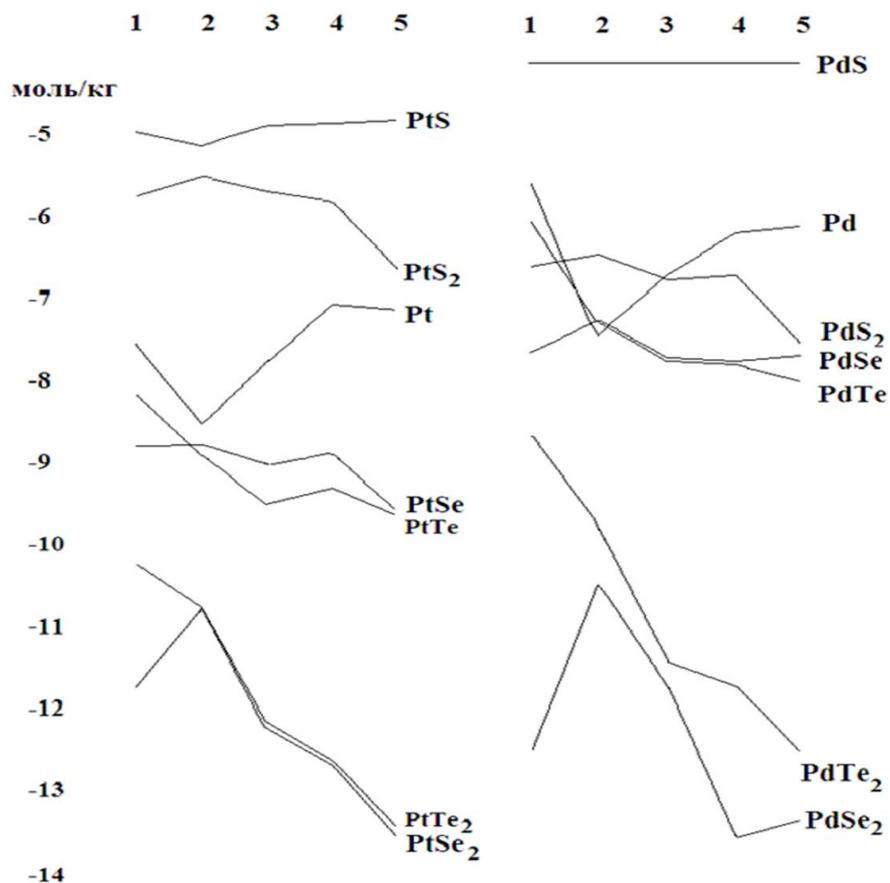


Рис. 3. Поведение селенидов и теллуридов платиноидов в процессе метаморфизма

1-5 — минеральные парагенезисы: 1 — Pl-Cum;
2 — Pl-Act; 3 — Pl-Act-Ts; 4 — Pl-Hbl; 5 — Pl-Act-Chl

Действительно существование зоны резкой смены окислительных условий на восстановительные оказывает определенное влияние на состав флюида и твердой фазы элементов платины, палладия и их халькогениды. Кроме того, из установленного интересного явления на определенной стадии метаморфизма (зеленосланцевой) при условии постоянного содержания серы в природной системе в составе ЭПГ присутствуют только элементы палладия,

платины и их сульфиды, а дисульфиды исчезают. Как показали результаты исследования, для образования дисульфидов элементов платиновой группы на этом этапе метаморфизма требуется более высокая концентрация серы в природной системе.

Этот факт, как представляется, может являться одним из поисковых признаков платиноидов.

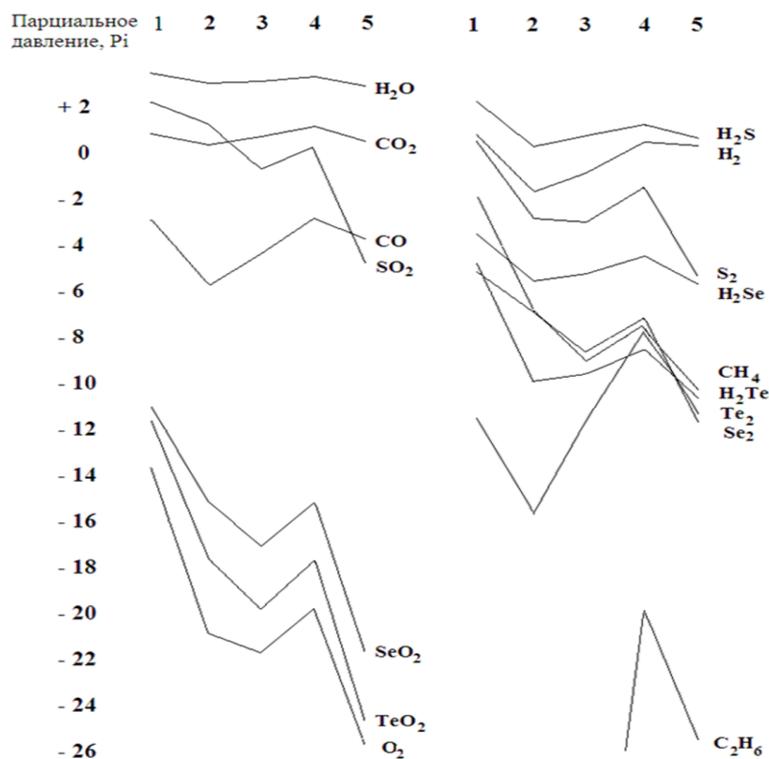


Рис. 4. Поведение компонентов флюида в процессе метаморфизма
1-5 – минеральные парагенезисы: 1 – Pl-Cum; 2 – Pl-Act; 3 – Pl-Act-Ts;
4 – Pl-Hbl; 5 – Pl-Act-Chl

II. Алмазы. Форма нахождения и количественное содержание углеродсодержащих соединений в природных образованиях различно. Например, в кианитовых сланцах методом ДТА установлен факт одновременного присутствия двух различных модификаций углерода (аморфная и кристаллическая) до 3.5 мас.% [11]. В кианит-силлиманитовых сланцах Больших Кейв наряду с углеводородными газами парафинового ряда установлено среднее содержание углерода до 1.18 мас.%.

Исходными данными для модельных исследований, с целью определения состава твёрдой и флюидной фаз при различных P-T параметрах, использованы результаты химического анализа образца 352/10 [12], представленные в таблице (табл. 1, вектор b₁).

Таблица 1. Химический и векторные составы исследуемого образца

Хим. состав	Мас. %	Вектора			
		b ₁	b ₂	b ₃	b ₄
		моль/кг			
SiO ₂	61.89	10.5050	10.3891	10.089	9.96472
TiO ₂	1.75	0.2192	0.21562	0.21461	0.21197
Al ₂ O ₃	28.67	5.6275	5.53291	5.50797	5.44036
Fe ₂ O ₃	0.62	-	-	-	-
FeO	1.25	0.2518	0.24994	0.24642	0.24344
MgO	0.49	0.1217	0.01204	0.1191	0.11761
CaO	0.26	0.0464	0.04530	0.04542	0.04485
Na ₂ O	1.31	0.4229	0.41805	0.4140	0.40894
K ₂ O	1.16	0.2464	0.24168	0.24122	0.23827
∑H ₂ O	0.06	2.4658	2.44865	2.41384	2.38422
S _{общ.}	0.07	0.0218	0.0193	0.01542	0.02112
CO ₂	0.1	0.0227	0.9032	1.78366	2.76262
Ni	0.01	0.0017	0.0017	0.00167	0.00165
Cu	0.01	0.0025	0.0016	0.00154	0.00152
P ₂ O ₅	0.13	0.0184	0.0180	0.01794	0.00830
П.п.п.	2.16	-	-	-	-
∑O		30.6593	30.3882	30.3313	30.3174

В наших исследованиях каждая мультисистема содержала 15 независимых компонентов (элементы):

Al-C-Ca-Cu-Fe-H-K-Mg-Na-Ni-O-P-S-Si-Ti.

Расчётная матрица мультисистемы была составлена из 45 зависимых компонентов минеральных фаз и флюидной фазы, включающей десять газообразных компонентов: H₂O, H₂, O₂, CO, CO₂, H₂S, SO₂, S₂, CH₄, C₂H₆. За основу минеральных фаз в мультисистемах были использованы результаты изучения минерального состава основных, второстепенных и аксессуарных компонентов: силикаты, алюмосиликаты, оксиды, карбонаты, апатиты, сульфиды и др. [11]. Для проведения модельных физико-химических теоретических исследований была использована справочная литература [6-8].

За основу модельного исследования были использованы мольные содержания элементов всех векторов (табл. 1). Теоретическое исследование было проведено в интервале температуры 300-900°C и при P = 5, 10 и 15 кбар. Результаты расчёта для всех точек, отмеченных на рисунке (рис. 5), позволили установить, что с повышением давления в системе увеличивается концентрация кианита при относительном стабильном содержании анортита и альбита. При этом постепенно исчезают из системы мусковит, ставролит и хлорит, а рутил замещается ильменитом при температуре около 700°C. Практически во всем интервале давления и температуры отмечается присутствие карбонатапатита, за исключением высоких P и T, где вместо него появляется гидроксилпатит. Была установлена восстановительная среда с характерными изменениями в составе твердой фазы в присутствии флюида, который содержит незначительные концентрации CO₂, CH₄, H₂S и других компонентов. Состав флюидной фазы в основном водноуглекислый при относительно высоком содержании метана и этана.

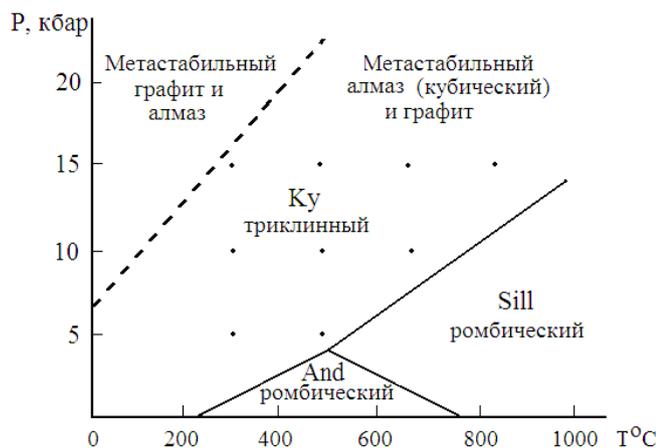


Рис. 5. Совмещенная P-T диаграмма Al_2SiO_5 с фазовой P-T диаграммой состояния углерода

В модельных системах было предусмотрено оценить влияние различного содержания углерода на конечные результаты. В результате теоретического исследования было установлено, что образование алмаза возможно только при определенных величинах T, P, концентрации углерода и, главное, парциального давления кислорода (табл. 2). Выявлены характерные изменения в составе твердой фазы в присутствии восстановительных условий среды с флюидом, который содержит незначительные концентрации CO_2 , CH_4 , H_2S и другие компоненты.

Таблица 2. Содержание алмазов в зависимости от T, P и P_{O_2} в исследуемой системе

T, °C	P, бар	моль/кг			P_{O_2} , бар
500	1000	1.332	2.663	3.995	$2.4e^{-24}$
	5000	-	0.256	0.786	$9.0e^{-24}$
	10000	-	0.263	0.836	$1.2e^{-23}$
700	1000	-	0.215	0.845	$3.4e^{-19}$
	5000	-	-	0.436	$1.2e^{-18}$
	10000	-	-	0.804	$1.6e^{-18}$
900	1000	-	-	0.388	$4.1e^{-16}$
	5000	-	-	сл.	$4.0e^{-16}$
	10000	-	-	-	$6.0e^{-15}$

На основании проведенных теоретических физико-химических исследований была показана возможность кристаллизации и существования обеих полиморфных модификаций углерода в алюмосиликатной системе кианит-силлиманит-андалузит (Ky – Sill – And) и в области метастабильного алмаза с графитом. Это позволило считать, что совместная кристаллизация алмаза и графита в алюмосиликатной системе свидетельствует об относительно низкотемпературных условиях их образования (при температуре ниже $1000^\circ C$).

Согласно многочисленным публикациям следует, что условия образования алмазов следующие: октаэдрические (крупные) – высокотемпературные, а тетраэдрические и, особенно, кубические (мелкие) относят к условиям кристаллизации на 200° ниже. Это означает, что при низких температурных условиях кристаллизации преобладают в основном кристаллы алмазов кубической формы, т.е. образование морфологических разновидностей алмазов зависят от Р-Т параметров кристаллизации.

При обсуждении результатов физико-химического моделирования процесса совместного образования графита и алмаза в широком интервале температуры и давления в области существования кианита было высказано предположение о возможном присутствии в данной сложной алюмосиликатной природной системе минерала карбида кремния – муассанита – SiC.

На основании проведенных исследований была показана возможность кристаллизации и существования карбида кремния совместно с обеими полиморфными модификациями углерода в системе Ky-Sill-C и области метастабильного алмаза с графитом (рис. 5). Это позволило считать, что именно совместная кристаллизация графита, муассанита и алмаза в алюмосиликатной системе свидетельствует, что их образование возможно при определённой концентрации углерода в условиях восстановительной среды и определенных величин давления и температуры (рис. 6).

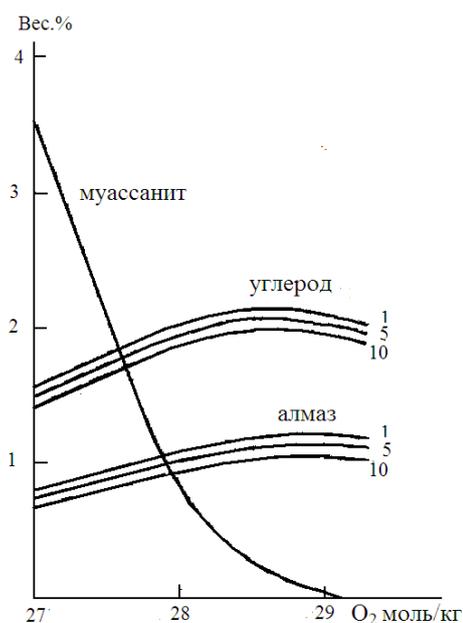


Рис. 6. Распределение концентраций углерода, алмаза и муассанита в зависимости от концентрации кислорода и разных величин давления

III. Апатит. Апатит широко используется для производства фосфорных удобрений, для получения фосфора и его соединений, а также в металлургии и др. Увеличение объема фосфорных удобрений и расширение ассортимента многие исследователи связывают с получением плавящихся кальциево-магниевого фосфатов [13-15]. В качестве примера ими предлагались использовать

комплексные руды Ковдорского месторождения, а также отходы апатитовой флотации и магнитной сепарации. Условия температуры на процесс и важность влияния различных добавок на образование кальциево-магниевого фосфата авторами не установлено.

Нами создана математическая модель процесса разложения сложной фосфорсодержащей минеральной смеси с учётом результатов изложенных выше экспериментальных исследований. Модель для исследования процесса получения плавящихся магниевого фосфата описана системой из 11 независимых компонентов:

Al-C-Ca-F-Fe-H-K-Mg-O-P-Si,

в которую включены 39 наиболее вероятных минералов, химических соединений и ряда газообразных компонентов [7,8]. В качестве объекта исследования был принят апатит Ковдорского месторождения. Данный природный объект представляет собой сложную смесь, содержащую фтор-, гидроксил-, карбонат- и оксиапатит, т.е. гидроксифторапатит с незначительной долей его карбонатсодержащей. Целью исследования явилось определение условий образования лимоннорастворимого α - $\text{Ca}_3(\text{PO}_4)_2$. В качестве основы для проведения математического моделирования использовались искусственные составы шихты Ковдорского апатита с присутствием различных добавок: доломита, флогопита, диоксида и их смесей. Исследуемый температурный интервал нагрева составил 800-1400°C. В результате проведённых физико-химических расчётов при помощи ЭВМ было установлено, что лимоннорастворимые α - и β - $\text{Ca}_3(\text{PO}_4)_2$ действительно образуются за счёт разложения $\text{Ca}_{10}(\text{PO}_4)_2\text{CO}_3$, $\text{Ca}_{10}(\text{PO}_4)_2\text{O}$, и $\text{Ca}_{10}(\text{PO}_4)_2(\text{OH})_2$ (рис. 7). Однако, в этом же температурном интервале фторапатит, присутствующий в сложной шихте, не разлагается, что подтвердилось и результатами проведённых лабораторных экспериментов. Поэтому для получения из апатита плавящихся магниевого фосфата необходимы иные добавки в шихту (например, углерод, диоксид кремния, которые будут способствовать разложению апатита).

По результатам проведённого нами термодинамического моделирования процесса получения плавящихся магниевого фосфата (прототип предлагаемого комплексного использования минерального сырья Ковдорского гока) из шихт разного состава следует, что полученные расчётным путём данные показали относительно высокую корреляцию с результатами экспериментальных исследований. Определены оптимальные температурные условия (свыше 1000°C) получения необходимого лимоннорастворимого фосфорсодержащего компонента. Результаты теоретического исследования подтверждают вывод экспериментаторов о постоянном присутствии фторапатита в продуктах термического воздействия на исходную шихту (свыше 12 вес.%).

Таким образом, получаемый по данной методике «плавящийся кальциево-магниевого фосфат» как удобрение будет засорять поля нерастворимым фторапатитом $\text{Ca}_{10}(\text{PO}_4)_2\text{F}_2$ (его ПР равен 115,6 [16]). В связи с этим данный метод возможного получения плавящихся кальциево-магниевого фосфата не был рекомендован для производства промышленностью.

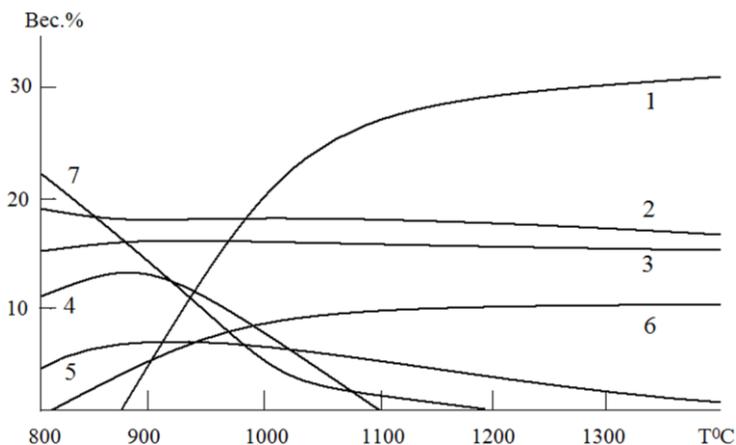


Рис. 7. Зависимость содержания компонентов от температуры нагрева смесей фосфорсодержащего соединения с диопсидом

Фазы: 1) α - и β - $\text{Ca}_3(\text{PO}_4)_2$; 2) Mg_2SiO_4 ; 3) $\text{Ca}_{10}(\text{PO}_4)_2\text{F}_2$; 4) $\text{Ca}_{10}(\text{PO}_4)_2\text{CO}_3$; 5) $\text{Ca}_{10}(\text{PO}_4)_2\text{O}$; 6) CaMgSiO_4 ; 7) $\text{Ca}_{10}(\text{PO}_4)_2(\text{OH})_2$

IV. Эвдиалит $\text{Na}_{12}\text{Ca}_6\text{Fe}_3\text{Zr}_3[\text{Si}_3\text{O}_9]_2[\text{Si}_9\text{O}_{24}(\text{OH})_3]_2$. Эвдиалит – известный минерал Хибинского и Ловозёрского щелочных массивов (Кольский полуостров), является концентратом многочисленных (более 16) редких и редкоземельных элементов: Zr, Nb, Hf, РЗЭ (TR_2O_3) и др. Их переменное количество в виде катионов и анионов содержится в его составе на уровне изовалентного и гетеровалентного изоморфизма. В нем может присутствовать до 17% ZrO_2 , а общее содержание РЗЭ (TR_2O_3) достигает 7%. Содержание оксидов некоторых РЗЭ (TR_2O_3) в эвдиалите достигает следующих концентраций: La_2O_3 (0.11-0.24), Ce_2O_3 (0.19-0.30), HfO_2 (0.15-0.36), Nd_2O_3 (0.07), Cd_2O_3 (0.12), Dy_2O_3 (0.1) [12]. Присутствие в эвдиалите значительного количества циркония, редкоземельных и других элементов для исследователя представляет определённый интерес с целью его использования в качестве перспективного минерального сырья для производства различных химических соединений и, соответственно, последующего использования как сырья в промышленности. По данным химического анализа концентрата эвдиалита был проведен расчёт его формульного выражения:



Для описания термодинамических равновесий в теоретической модели разложения кислотами эвдиалита использована 15-ти компонентная мультисистема, которая содержала в своем составе следующие независимые компоненты (элементы):



Расчетная матрица мультисистемы представлена 59 зависимыми компонентами (химические соединения) с жидкой фазой – водный раствор, свойства которого описаны растворенными частицами. В модели учтены наиболее важные растворенные формы элементов – катионы и анионы [8], в соответствии с используемой в лабораторных исследованиях для разложения эвдиалита соляной кислотой (HCl 37 мас.%).

При этом одновременно учитываются возможные реакции окисления-восстановления, растворения, гидролиза, комплексообразования, ионного обмена и т.д. При наличии в водном растворе компонентов с различной степенью окисления в создаваемую систему дополнительно вводится независимый компонент – заряд, т.к. система должна быть электронейтральной.

Из полученного предварительного расчётного материала следует, что эвдиалит растворим в воде и довольно хорошо разлагается соляной кислотой. Последующими детальными исследованиями установили некоторую характерную особенность влияния HCl на этот процесс. Так, с увеличением величины ж/тв, т.е. количества соляной кислоты в растворе, система претерпевает смещение по координате рН в сторону снижения, в кислую область, а значения Eh соответствуют восстановительной среде. Окислительно-восстановительный потенциал является термодинамической характеристикой, которая зависит от химического состава и соотношения концентраций окисленной и восстановленной форм вещества.

В составе раствора из компонентов, содержащих цирконий, установлена значительная концентрация оксихлорида циркония ($ZrOCl_2$), хлорионов циркония ($ZrCl_2^{+2}$, $ZrCl^{+3}$) и Zr^{+4} . Содержание последних ионов имеет как прямую, так и обратную корреляцию с концентрацией соляной кислоты. В растворе отмечается значительная концентрация хлоридов натрия, калия, магния, кальция и железа, равновесных с их хлорионами. Из кремнийсодержащих компонентов в растворе прослеживается высокая концентрация только положительного иона ортокремнёвой кислоты. Соединение ниобия в виде компонента $Nb(OH)_5$ установлено только в растворе. В продуктах исследуемого процесса твёрдая фаза представлена диоксидом кремния, цирконом, диоксидом и хлоридом титана, а также гидроксидхлоридом марганца и хлоридом стронция равновесного с его ионом Sr^{+2} в жидкой фазе.

Следует отметить несколько неожиданное влияние 1,5–2 кг соляной кислоты на процесс разложения 1 кг эвдиалита. В растворе отмечается резкое возрастание содержания HCl, что сопровождается скачком величин рН и Eh (рис. 8).

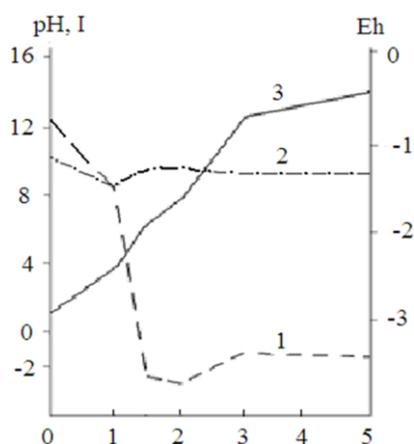


Рис. 8. Зависимость кислотно-основных свойств, рН (1), окислительно-восстановительного потенциала, Eh (2) и ионной силы раствора, I (3) при разложении эвдиалита соляной кислотой в зависимости от величины отношения ж/тв

Этот факт нашел отражение в продуктах разложения эвдиалита – снижение в растворе содержания: CaCl^+ , FeCl^+ , K^+ , Na^+ , MgCl^+ , Sr^{+2} и увеличение (почти на порядок) концентрации ZrCl_2^{+2} , а в составе твердой фазы возросло содержание диоксида кремния и исчез циркон (ZrSiO_4). Анализ результатов процесса разложения эвдиалита соляной кислотой показал, что увеличение отношения в системе ж/тв не способствовало заметному возрастанию концентраций твердых продуктов и многих компонентов жидкой фазы (рис. 9,10).

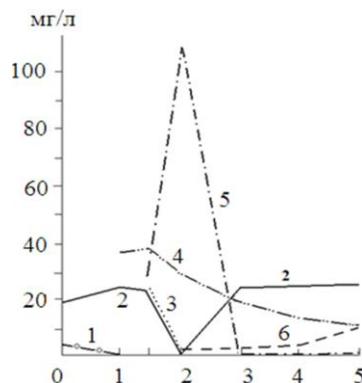


Рис. 9. Изменение содержания продуктов разложения эвдиалита соляной кислотой в зависимости от величины отношения ж/тв:
 1 – HZrO_3^- , 2 – ZrSO_4 (вес.%), 3 – Zr^{+4} (x 10), 4 – $\text{ZrOCl}_2 \cdot 10^{-3}$, 5 – $\text{ZrCl}_2^{+2} \cdot 10^{-3}$, 6 – ZrCl^{+3} .

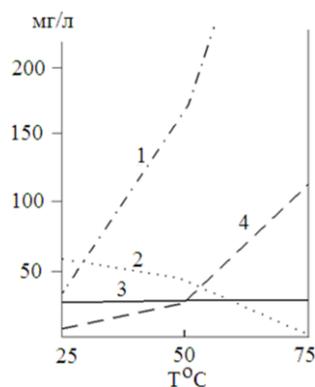


Рис. 10. Влияние температуры на концентрацию компонентов – продуктов разложения эвдиалита соляной кислотой (ж/тв = 3):
 1 – ZrCl_2^{+2} , 2 – MgCl^+ , 3 – ZrSiO_4 (мас.%), 4 – ZrCl^{+3} .

Данной работой предполагалось получить путём кислотной переработки эвдиалита технологическую схему с целью получения циркона и побочных продуктов за счёт присутствия различных добавок (с учётом комплексности использования минерального сырья). Анализируя полученный расчетный материал, следует отметить:

1. В процессе разложения эвдиалита соляной кислотой образуется твердая фаза, в составе которой установлены: ZrSiO_4 , TiO_2 , TiCl_2 , MnOHCl , SiO_2 и $\text{SrCl}_2 \cdot \text{H}_2\text{O}$.

2. В жидкой фазе продуктов разложения эвдиалита присутствует значительная концентрация $ZrOCl_2$, хлорионов циркония $ZrCl_2^{+2}$ и $ZrCl^{+3}$, а также следовое содержание ионов Zr^{+4} , ZrO^{+2} . Содержание последних ионов непосредственно коррелирует с содержащейся в системе соляной кислоты. В жидкой фазе установлено присутствие значительной концентрации ионов ортокремниевой кислоты.

3. Хлориды натрия, калия и хлорсодержащие ионы железа, кальция и магния в жидкой фазе находятся в равновесном состоянии с ионами натрия, калия и хлора.

4. Среди продуктов разложения эвдиалита ниобий, в виде компонента $Nb(OH)_5$, и ионы алюминия присутствуют только в жидкой фазе.

5. Высокое содержание иона стронция (Sr^{+2}) в жидкой фазе равновесно с $SrCl_2 \cdot H_2O$ (твёрдая фаза) только при определенном отношении жидкости (соляная кислота) к твёрдому (эвдиалит).

6. Изменение стандартных условий процесса разложения эвдиалита за счет повышения температуры в реагирующей системе практически оказало незначительное влияние на состав и содержание образующихся твёрдых продуктов.

7. Полученный расчетный материал позволяет считать нецелесообразным использование для разложения эвдиалита больших величин отношения жидкости (соляная кислота) к твёрдому (эвдиалит).

Заключение

В результате произведенных исследований следует, что образование алмазов в алюмосиликатной системе происходит при относительно низких величинах давления, температуры, восстановительных условиях среды и наличия углерода как результата разложения присутствующих в ней карбонатов.

На конкретных результатах численного моделирования показана целесообразность предварительной теоретической проработки любой планируемой экспериментальной работы. В качестве примера рассмотрены примеры химико-технологической переработки некоторых видов минерального сырья Кольского полуострова. Такой подход позволит, например, значительно снизить затраты на поиски оптимальных (экспериментальных) путей переработки данного минерального сырья. Вычислительный эксперимент обходится дешевле, чем проведение лабораторного или полупромышленного реального процесса. Такой подход к решению проблемы легко управляем и даёт возможность теоретически прогнозировать поведение объектов в неизвестных ситуациях (давление, температура, концентрации кислот и др.).

Поэтому одной из внутренних задач данной работы является дальнейшее более интенсивное использование общей методологии экспериментального и математического моделирования подобных систем с использованием вычислительной техники для решения многочисленных задач науки.

Литература

1. Зельдович Я. Б. О единственности решения уравнений закона действующих масс // Ж. физ. химии. 1938. Т. 11, Вып. 5. С. 685–687.
2. Зельдович Я. Б., Овчинников А. А. Асимптотика приближения к равновесию и флуктуация концентрации // Письма в ЖЭТФ. 1977. Т. 26. С. 588–591.

3. Berman R. G. Thermobarometry Using Multi-Equilibrium Calculations: a New Technique, with Petrological Applications // *Canad. Miner.* 1991. Vol. 29, № 4. pp. 833–855.
4. Казьмин Л. А., Халиуллина О. А., Карпов И. К. Расчет химических равновесий поликомпонентных гетерогенных систем методом минимизации свободной энергии // *Информ. Бюл. «Алгоритмы и программы» / ВНИИЦ.* 1975. № 3. С. 8–19.
5. Карпов И. К. Физико-химическое моделирование на ЭВМ в геохимии. Новосибирск: Наука. 1981. 248 с.
6. ИВТАНТЕРМО, 2006. Режим доступа: <http://www.chem.msu.su/cgi-bin/tkv2.pl>
7. Yokokawa H. Tables of thermodynamic properties of inorganic components. Kagaku gidziutsu kenkiudze kokaku // *J. Nat. Chem. Lab. Indust. Spec. Iss.* 1988. Vol. 83. 121 p.
8. Каржавин В. К. Термодинамические величины химических элементов и соединений. Примеры их практического применения. Апатиты: Изд-во КНЦ РАН. 2011. 160 с.
9. Каржавин В. К., Волошина З. М. Модельные исследования условий метаморфизма и флюидного режима рудоносного горизонта Панского массива в связи с ЭПГ оруденением // *Геохимия.* 2006. № 5. С. 522–531.
10. Kretz R. Symbols for rock-forming minerals // *Amer. Mineral.* 1983. Vol. 68. № 1/2. pp. 277–279.
11. Гинзбург И. В., Горшков А. И. О графите кианитовых сланцев Кейв (Кольский полуостров) // *Тр. Минералог. музея им. А .Е.Ферсмана. М.: Изд-во АН СССР.* 1961. Вып. 12. С. 171–176.
12. Нерадовский Ю. Н., Войтеховский Ю. Л. Атлас структур и текстур кристаллических сланцев Больших Кейв. Апатиты: Изд-во К&М. 2013. 116 с.
13. Брицке Э. В., Ионасс Л. А. Плавленные фосфаты // *Исследования по прикладной химии Э. В. Брицке, Л.А Ионасс. М.-Л.: Изд-во АН СССР.* 1955. С. 58–66.
14. О влиянии диоксида кремния на процесс получения плавленных кальциево-магниевых фосфатов электротермическим методом / Н. Н. Дербунович и др. // *Ж. прикл. химии.* 1986. Т. 59. № 8. С. 1670–1673.
15. Влияние минерального состава шихты на процесс получения фосфорно-магниевых удобрений / А. П. Процюк и др. // *Ж. прикл. химии.* 1989. Т. 62. № 9. С. 2035–2040.
16. Лурье Ю. Ю. Справочник по аналитической химии. М.: Химия. 1971. 456 с.
17. Каржавин В. К. Термодинамическое моделирование сернокислого разложения эвдиалита // XVI Международная конференция по химической термодинамике в России. т.1, Москва: Изд-во МГУ. 2007. С. 73–74.

Сведения об авторе

Каржавин Владимир Константинович

к.х.н., старший научный сотрудник

e-mail: karzhavin@geoksc.apatity.ru

Vladimir K. Karzhavin

PhD in Chemistry, senior researcher

В. М. Нифантов

Филиал ФГБОУ ВО МАГУ в г. Апатиты

ДИАГНОСТИКА, ОЦЕНКА И ПРОГНОЗИРОВАНИЕ ТЕХНИЧЕСКОГО СОСТОЯНИЯ ТЕХНОЛОГИЧЕСКОГО ОБОРУДОВАНИЯ ПРИ ПОМОЩИ НЕЧЕТКОЙ ЭКСПЕРТНОЙ СИСТЕМЫ В ЦЕНТРАЛИЗОВАННЫХ СИСТЕМАХ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ И РЕМОНТА

Аннотация

В статье рассмотрены вопросы, связанные с решением задач диагностики, оценки и прогнозирования технического состояния технологического оборудования при использовании централизованных систем технического обслуживания и ремонта. Описывается возможность построения информационной системы для определения параметров распределения ресурса оборудования. Подход обусловлен применением нечетко-логических процедур для управления прогнозированием состояния оборудования, по результатам экспертной оценки специалистами предприятия. Описываются возможности экспертной системы как инструмента практического решения задач при использовании централизованных систем технического обслуживания и ремонта оборудования.

Ключевые слова:

техническое состояние, техническая диагностика, технологическое оборудование, прогнозирование состояния, нечеткая экспертная система, централизованная система, техническое обслуживание и ремонт.

V.M. Nifantov

DIAGNOSTICS, ASSESSMENT AND FORECASTING OF TECHNICAL CONDITION OF TECHNOLOGICAL EQUIPMENT BY USING FUZZY EXPERT SYSTEM IN CENTRALIZED MAINTENANCE AND REPAIR SYSTEMS

Abstract

The article discusses issues related to solving the problems of diagnosis, assessment and prediction of the technical condition of technological equipment when using centralized maintenance and repair systems. The possibility of constructing an information system for determining the parameters of the equipment resource distribution is described. The approach is due to the use of fuzzy-logical procedures to control the prediction of the state of equipment, according to the results of an expert assessment by the company's specialists. The capabilities of an expert system are described as a tool for practical solving problems when using centralized systems for maintenance and repair of equipment.

Ключевые слова:

technical condition, technical diagnostics, process equipment, condition forecasting, fuzzy expert system, centralized system, maintenance and repair.

Введение

От технического состояния оборудования зависят не только экономические показатели предприятия, но и безопасность работы эксплуатирующего персонала. Совершенствование системы технического обслуживания и ремонта, и последующий переход к обслуживанию по фактическому техническому состоянию (ОФС) является залогом безопасной и эффективной работы предприятий. Неотъемлемой составной частью

эксплуатации технологического оборудования является определение износа оборудования путем мониторинга и диагностики его технического состояния.

Сегодня большинство крупных производств используют в своей структуре централизованные системы технического обслуживания и ремонта (ТО и Р) оборудования, которые разделены по таким областям как: механика, энергетика, автоматизация и управление технологическими процессами, диагностика технологических машин. Соответственно, каждое структурное подразделение централизованной системы ТО и Р несет ответственность за конкретные узлы и агрегаты оборудования в рамках своей зоны ответственности.

Основным недостатком таких систем является некоторое снижение оперативности принятия решений при проведении ремонтных работ. Но их можно устранить и компенсировать, например, с помощью сокращения доли внеплановых ремонтов за счет лучшего планирования ремонтных работ и раннего выявления неисправностей, внедрения специализированных информационных систем управления ТО и Р оборудования [1].

Для эффективного использования технологического оборудования необходимо обладать информацией о его работоспособности не только в процессе эксплуатации, но и в некоторый момент времени в будущем, что составляет задачу прогнозирования технического состояния. Решение данной задачи связано с определением текущего технического состояния оборудования, путём оценки всех систем и опираясь на опыт эксплуатации специалистами предприятия (история оборудования). Оценка обычно производится по каким-либо параметрам работы оборудования, и имеет детерминированное значение. Основная проблема состоит в том, что сложная техническая система (технологическое оборудование) может находиться в бесконечном множестве состояний.

1. Существующие методики и нормативные документы в области оценки и прогнозирования состояния технологического оборудования

При диагностике определяемое множество состояний ограничено возможностями контрольных и измерительных средств. Основным стандартом в области контроля и диагностики машин является ГОСТ Р ИСО 17359-2015[2]. Стандарт устанавливает рекомендации в отношении процедур, используемых при реализации программ контроля состояния и диагностирования машин (систем мониторинга). Приведенные рекомендации распространяются на машины всех видов, стандарт является основополагающим для комплекса стандартов в области контроля состояния и диагностики оборудования.

Основным недостатком [2] является то что, данным документом не учитываются особенности организационных структур, где используются централизованные системы ТО и Р оборудования. Также реализация в реальных производственных условиях зачастую затруднена, т.к. требуется соответствующая информационная инфраструктура, а существующие ЕАМ и ERP системы предприятий используются для решения задач в бизнес-процессах.

Существующие методики [3] и стандарты [4] в области прогнозирования остаточного ресурса технологических машин и оборудования требуют наличия большого объема вводных и статистических данных для расчета. Информационное наполнение документов [3,4] требует высокой квалификации

специалиста для правильного применения методик в реальных производственных условиях и правильной интерпретации результатов.

2. Технологическое оборудование и его информационное пространство

Основными составляющими элементами различного технологического оборудования, влияющими на выполняемые им функции (технологический процесс), представлены на рисунке 1.



Рис. 1. Основные элементы технологических машин и оборудования

Каждая составляющая из систем напрямую или косвенно может влиять на бесперебойную работу оборудования. В централизованных системах ТО и Р для этого привлекаются специалисты предприятия, в зоне ответственности которых лежит оценка и принятие решений о воздействии на оборудование (его составляющие элементы), или же его дальнейшая эксплуатация и мониторинг состояния. Различают субъективные и объективные методы оценки технического состояния оборудования [5]:

– под **субъективными** (органолептическими) методами подразумеваются такие методы оценки технического состояния оборудования, при которых для сбора информации используются органы чувств человека. При этом для анализа собранной информации используется аналитико-мыслительный аппарат человека, базирующийся на полученных знаниях и имеющемся опыте. К таким методам оценки технического состояния относят визуальный осмотр, контроль температуры, анализ шумов и другие методы;

– под **объективными** (приборными) методами подразумеваются такие методы оценки технического состояния, при которых для сбора и анализа информации используются специализированные устройства и приборы, компьютеры, а также соответствующее программное и нормативное обеспечение. К объективным методам оценки технического состояния относятся вибрационная диагностика и методы неразрушающего контроля (магнитный, электрический, вихретоковый и др.).

Информация, необходимая для оценки и прогнозирования — это совокупность сведений о состоянии оборудования, его узлов и агрегатов, а также параметрах, тенденциях и закономерностях происходящих процессов. Для каждого оборудования существует информационное пространство, представляющее единое целое и которое развивается со временем. Это информационное пространство (рис. 2) формируют специалисты предприятия, которые используют собственные специализированные онтологии и тезаурус в конкретных производственных условиях.

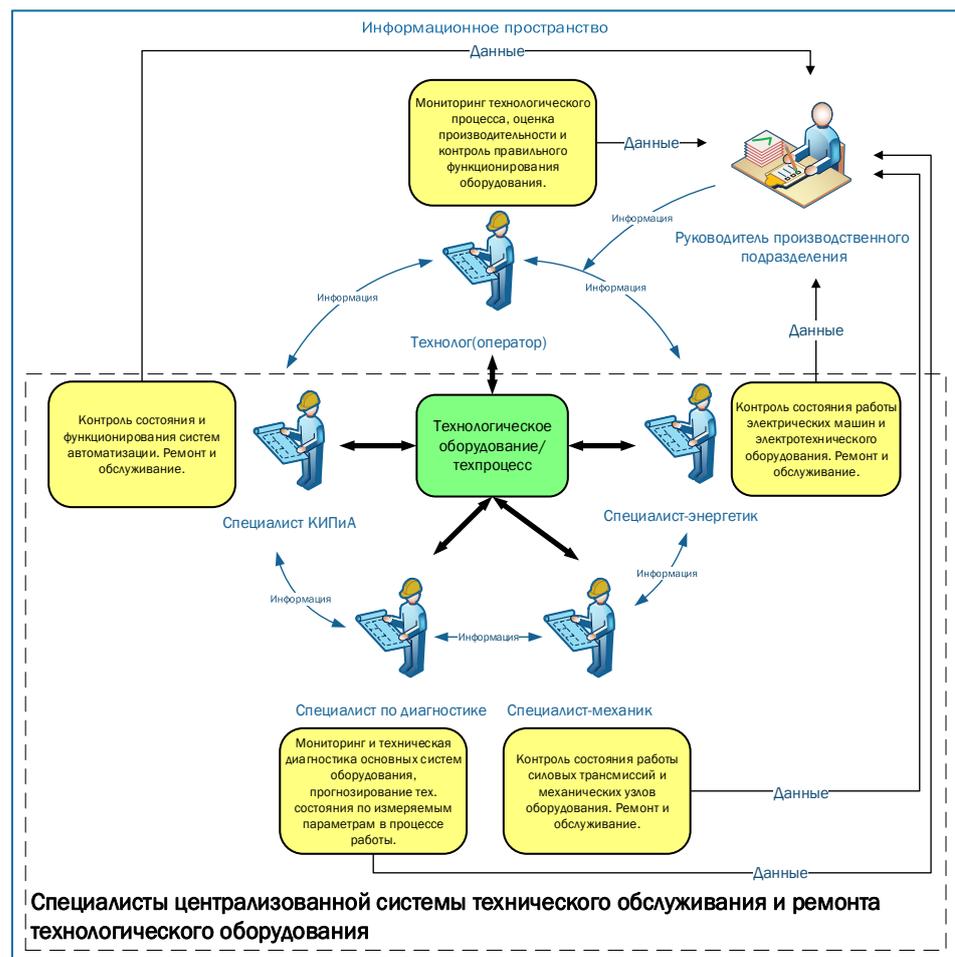


Рис. 2. Информационное пространство, формируемое специалистами предприятия (подразделения) в процессе эксплуатации оборудования

В общем виде информационное пространство можно представить как сложную информационную систему со своей инфраструктурой. В данном случае общую оценку технического состояния оборудования можно получить путем получения структурированных данных от каждого специалиста (по направлению деятельности) о состоянии отдельно взятых подсистем и узлов технологического оборудования.

Значимость информации [6] которая используется для информационного описания состояния технологического оборудования, и энтропию информации можно определить по формуле К. Шеннона:

$$H = - \sum_{i=1}^n p_i \log_2 p_i \quad (1)$$

где H — энтропия; p_i — вероятность нахождения системы в i -том состоянии; n — число возможных состояний системы.

Качественные характеристики и количественные показатели информационного описания зависят от уровня декомпозиции сложной системы, а значит, и от соответствующего уровня контроля. В количественном отношении информация о состоянии объекта, которую получают при диагностике, равна разности энтропии объекта до и после очередной проверки. Объектом диагностики может быть как технологическое оборудование в целом, так и его отдельные подсистемы.

При разработке современных систем оценки и прогнозирования технического состояния различного технологического оборудования необходимо использовать основные принципы моделирования процесса обработки информации по субъективным и объективным методам оценки. Главная особенность информационных систем в области диагностики и прогнозирования состояния технических систем связана с обработкой больших объемов анализируемой информации и ограниченным периодом времени получения пользователем рекомендаций.

В процессе работы технологических машин и оборудования происходят сложные нелинейные взаимодействия между входящими в его состав элементами, сложные цепочки (сценарии) причинно-следственных связей между опасными, имеющими вероятностный характер событиями и процессами, протекающими в течение всего срока эксплуатации.

Современное понимание измерений существенно шире только количественных измерений. Есть наблюдаемые явления, в принципе не допускающие числовой меры, но которые можно фиксировать в определённых шкалах или опросных листах. Расплывчатость некоторых наблюдений признана их неотъемлемым свойством, которому придана строгая математическая форма, и разработан формальный аппарат работы с такими наблюдениями [7]. Оценка технического состояния оборудования также не равноценна по количеству экспертных оценок, поэтому в задачах оценки и прогнозирования существенным является выбор наиболее информативных признаков для описания текущего состояния оборудования.

Одной из главных задач по результатам экспертной оценки технологического оборудования является оптимизация информационного обеспечения процесса принятия решений. Оптимизация в данном случае заключается в минимизации объема результирующей информации. При этом в качестве основного критерия оптимизации выступает требование к однозначности выходной информации, получаемой на основании оценки экспертной группы.

Для оптимизации процесса принятия решений о техническом состоянии оборудования необходимо располагать сведениями о состоянии всех составляющих узлов, агрегатов и всех систем. Поэтому при оценке состояния важным вопросом является выбор информативных диагностических параметров для описания, и значимости мнения того или иного эксперта в зоне своей ответственности.

3. Информационное обеспечение для диагностики, оценки и прогнозирования технического состояния технологического оборудования

Наиболее эффективным методом диагностирования силового привода и механизмов с вращающимися узлами является виброакустический. Диагностические признаки, определяющие техническое состояние, вызываемые различного рода дефектами, следует оценивать по следующим состояниям: структурному (геометрия, взаимосвязь деталей, состояние материала деталей), функциональному (эксплуатационные показатели и характеристики рабочего процесса и процесса регулирования), вибрационному (совокупность колебательных процессов).

Значительное число дефектов составляют прочностные дефекты, что объясняется сложностью и недостаточным совершенством динамических расчетов, а также упрощенными представлениями о физике колебательных процессов и вызываемых ими напряжениях. Важную роль в обнаружении подобных дефектов играет виброакустическая диагностика, так как возникающие при работе машины вибрации достаточно полно отражают относительное состояние детали или узла, позволяют судить о возникающих динамических нагрузках, а также обнаруживать причины дефектов и прогнозировать техническое состояние исследуемых элементов. Отказы проявляются постепенно в изменении одного или нескольких выходных параметров. Контролируя изменение общего (интегрального) уровня вибрации, можно прогнозировать момент наступления следующего отказа путем интерполяции и экстраполяции результатов измерений.

Оценить состояние других основных систем технологического оборудования (систем АСУ ТП, силового электропитания, перегрузочных устройств, состояние исполнительных органов гидравлических и пневматических систем при наличии и т.п.) можно путем получения данных от специалистов-экспертов, в зоне ответственности которых находится контроль исправного состояния данных элементов технологического оборудования.

Например, применение *шкалы оценки состояния* (исправное, работоспособное, ограниченно работоспособное, недопустимое, аварийное) является средством адекватного сопоставления и определения численных значений отдельных свойств и качеств различных систем технологического оборудования. *Чек-лист (чек-кейс)* как метод содержит перечень контрольных вопросов/заданий/критериев, с помощью которых определяется перечень узлов и систем оборудования, подверженных оценке по определенным критериям. Комплексная экспертная оценка состояния позволяет произвести синтез информации и получить данные о текущем состоянии технологического оборудования.

4. Экспертные информационные системы в задачах диагностики, оценки и прогнозирования технического состояния оборудования

Нечеткая экспертная система – экспертная система, которая для вывода решения использует вместо булевой логики совокупность нечетких функций принадлежности и набор правил. Сегодня в различных областях промышленности разработаны комплексы программ с использованием нечеткой экспертной системы: «КвиХар» (металлургия), «КОМПАКС» (Энергетика), АПК

для диагностики в движении роторных узлов вагонной тележки пригородного электропоезда. Все эти системы имеют ряд обобщенных функций:

- определение общего статуса и диагноза состояния объекта;
- определение дат наступления событий изменения состояния объекта по тренду или по приращениям диагностических параметров;
- мониторинг и диагностика состояния объекта на протяжении всего срока его службы, полное использование ресурса, заложенного в оборудовании при одновременном сохранении его ремонтпригодности и безопасности работы.

Основным недостатком вышеописанных систем является то, что они разработаны под конкретные производственные задачи и не учитывают особенности организационных структур.

Применительно к решению задач оценки технического состояния технологического оборудования и прогнозирования его работоспособности при использовании централизованных систем ТО и Р может быть использована экспертная система и применены алгоритмы нечеткой логики.

Выбор в качестве способа построения экспертной системы обработки информации с применением нечеткой логики в централизованных системах ТО и Р для диагностики, оценки и прогнозирования технического состояния оборудования обусловлен по следующим причинам:

- нечеткая логика дает возможность достаточно просто заложить информацию об объекте в виде нечетких правил, которые более близки к естественному языку и понятны;
- при привлечении экспертов (специалистов, связанных с эксплуатацией и ремонтом оборудования) увеличится и вероятность того, что диагностическая информация и прогноз состояния оборудования будут ближе к истине;
- нечеткие модели оказываются более простыми для реализации по сравнению с классическими алгоритмами оценки и прогнозирования технического состояния сложных технических систем;
- правила нечеткой логики могут быть использованы для построения структуры нейроподобной сети;
- результаты вывода нечеткой экспертной системы в задачах прогнозирования и диагностирования объекта могут быть использованы в качестве обеспечивающей системы ТО и Р для принятия решений, а также интегрированы в ЕАМ/ERP системы и др.

5. Общая концепция разрабатываемой информационной системы

На стадии эксплуатации жизненного цикла технологического оборудования преобладают стохастические процессы. Как показывает практика, усложнение систем автоматизации ограничено технически или нецелесообразно, а иногда приводит к увеличению затрат на ТО и Р. Концептуальная модель разрабатываемой информационной системы диагностики, оценки и прогнозирования технического состояния оборудования представлена на рисунке 3.

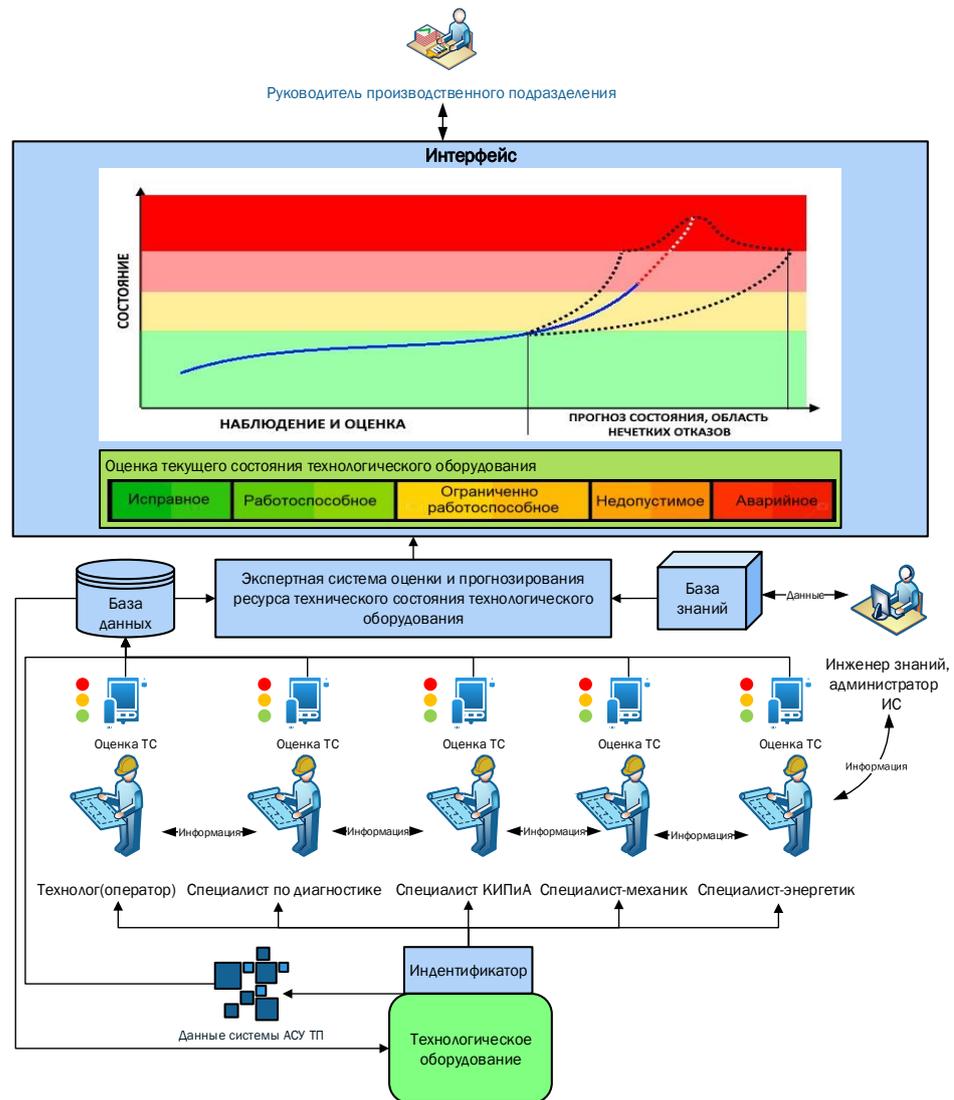


Рис. 3. Концепция разрабатываемой информационной системы

В разрабатываемой экспертной системе подразумевается использование параметров вибрации и теплового режима работы как основных критериев для диагностики и оценки состояния технологического оборудования. Наиболее достоверные такие данные возможно получить от специалистов службы диагностики централизованной системы ТО и Р, или из стационарных систем мониторинга при работе оборудования.

В свою очередь этого недостаточно для достоверной оценки и прогнозирования, т.к. отдельные системы (узлы, агрегаты, элементы и т.п.) технологического оборудования подвергаются замене, изменяются технологические параметры работы (история оборудования). Для корректной оперативной оценки требуется дополнительная информация об оборудовании от других специалистов. Зачастую такая информация представляет собой неструктурированные данные и специфичный тезаурус по областям знаний.

Графически отобразив тренд изменения вибрации (среднеквадратичное значение изменения виброскорости) в критической точке возможно прогнозировать дальнейшие изменения состояния оборудования (рис. 4).



Рис. 4. Тренд изменения вибрации машины

В данном случае (рис. 4) прогнозируемое значение состояния не является детерминированным и имеет недостаточную достоверность. В классических теориях надёжности [8] зачастую рассматривается нормальное распределение плотности распределения ресурса, которому подчиняется наработка на отказ многих восстанавливаемых и невосстанавливаемых объектов (рис. 5).

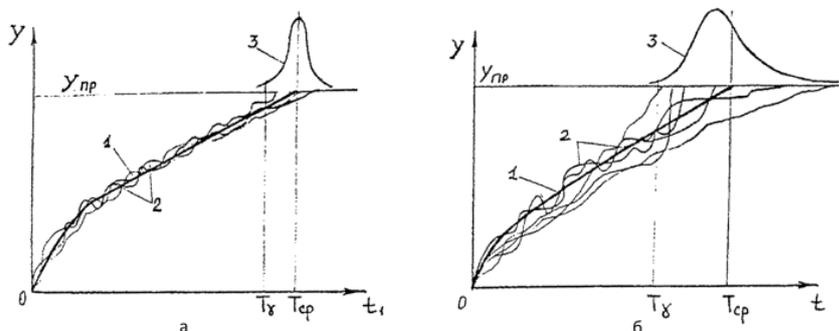


Рис. 5. Y при постоянной дисперсии (а) и непостоянной дисперсии (б); кривые: 1 – математического ожидания $Y(t)$; 2 – отдельных реализаций; 3 – плотность распределения ресурса; $T_{ср}$ – средний ресурс оборудования; T_{γ} – гарантированный ресурс оборудования; $Y_{пр}$ – предельное значение технического состояния

На практике распределение остаточного ресурса оборудование имеет ассиметричную форму (логнормальное или гамма распределения). Разрабатываемая экспертная система оценки специалистами позволит определить параметры плотности распределения ресурса оборудования на основе нечеткого вывода системы (область нечетких отказов), рисунок 6.

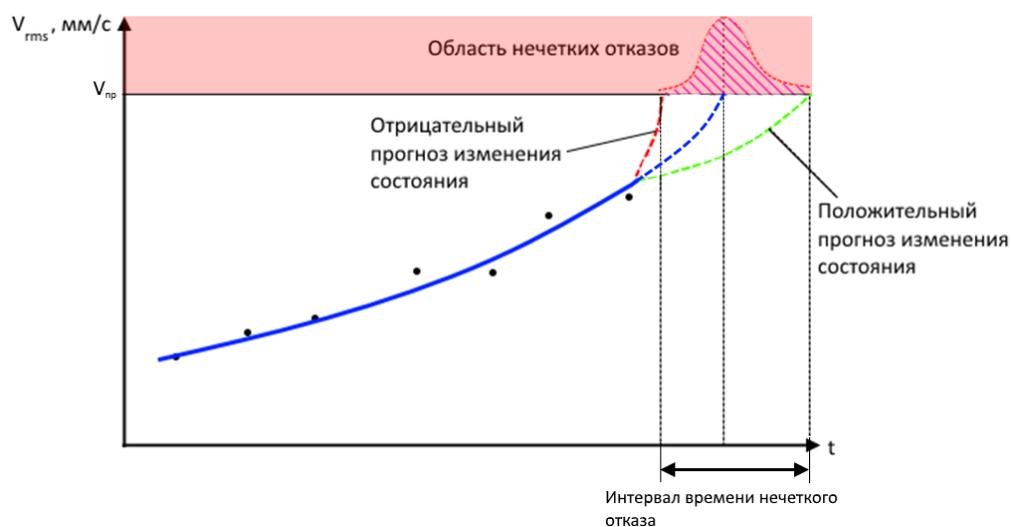


Рис. 6. Область нечетких отказов оборудования

Развитие и доступность информационной инфраструктуры предприятий (промышленные Wi-Fi технологии, мобильный интернет) и применение мобильных устройств для решения производственных задач сегодня являются распространённым явлением. В разрабатываемой системе предполагается использование специалистами шкалы для оценки состояния при помощи мобильного интерфейса, а для идентификации оборудования – системы штрих-кодов.

База знаний представляет собой набор правил для конкретных производственных условий и учитывает влияние тех или иных систем оборудования. В данном случае **задача системы нечеткого вывода** состоит в том, чтобы, руководствуясь базой правил для фактических значений входных переменных (экспертных оценок, системы АСУ ТП и др.), определить основные параметры плотности распределения ресурса (область нечетких отказов). Таким образом, возможно сузить или расширить интервал нечеткого отказа. Распределение Вейбулла применяют при описании надежности сложных технических систем [8, 9]. Это распределение является двухпараметрическим универсальным законом, так как при изменении параметров оно в пределе может описывать нормальное распределение, логнормальное распределение, экспоненциальное распределение и др.

Наиболее трудными являются вопросы обоснованного назначения критерия выбор прогнозирующих параметров. Теоретически обоснованные ответы на эти вопросы удается получить далеко не всегда и только для простых объектов [8].

Заключение

В статье рассматривается возможность применения нечеткой экспертной системы для диагностики, оценки и прогнозирования технического состояния технологического оборудования в виде информационной системы. Такая система позволит на основе экспертной информации проводить предварительную оценку общего состояния технологического оборудования, и получить достоверный прогноз ресурса. За счет оценки специалистами централизованной системы ТО и

Р предприятия возможно вычислить параметры распределения ресурса (область нечетких отказов) и тем самым повысить достоверность прогноза.

Экспертная система оценки состояния специалистами предприятия и обработка информации при помощи компьютерных и телекоммуникационных устройств, реализованная на базе современных технологий позволят:

- повысить качество оценки технического состояния технологического оборудования и достоверность прогнозирования;
- повысить эффективность работы оборудования за счет снижения количества отказов и заблаговременного принятия мер по предотвращению аварий;
- осуществить переход системы ТО и Р оборудования предприятия к обслуживанию по фактическому техническому состоянию;
- произвести анализ компетенций и навыков специалистов как экспертов в зоне ответственности направления деятельности, различных структур централизованной системы ТО и Р предприятия.

Литература

1. Попов, В. Н. Реорганизация системы технического обслуживания и ремонта оборудования на предприятиях // Организатор производства, 2011. № 4. С. 30–32.
2. ГОСТ Р ИСО 17359-2015 Контроль состояния и диагностика машин. Общее руководство.
3. РД 26.260.004-91 Методические указания. Прогнозирование остаточного ресурса оборудования по изменению параметров его технического состояния при эксплуатации.
4. ГОСТ Р ИСО 13381-1-2016 Контроль состояния и диагностика машин. Прогнозирование технического состояния. Часть 1. Общее руководство.
5. Зданевич В., Сидоров В.А. Осмотр механического оборудования как метод технической диагностики // Техническое обслуживание и ремонт, 2010. № 4. С. 12–18.
6. Трененков Е. М., Дведенидова С. А. Диагностика в антикризисном управлении. Режим доступа: <https://www.cfin.ru/press/management/2002-1/01.shtml>.
7. Диагностика и мониторинг состояния сложных технических систем: учебное пособие / Н. А. Махутов., В. Н. Пермяков, Р. С. Ахметханов и др. Тюмень: ТИУ, 2017. 632 с.
8. Федотов, А. В. Основы теории надежности и технической диагностики: конспект лекций / А. В. Федотов, Н. Г. Скабкин. Омск: Изд-во ОмГТУ, 2010. 64 с.
9. Абиев, Р.Ш. Надежность механического оборудования и конструкций: учебник. СПб: Изд-во СПбГТИ (ТУ), 2016. 299 с.

Сведения об авторах

Нифантов Вячеслав Михайлович

начальник группы диагностики оборудования, цеха централизованного ремонта и технического обслуживания технологического оборудования АО «Ковдорский горно-обогатительный комбинат»
e-mail: slava_niffon@mail.ru

Vyacheslav M. Nifantov

head of the equipment diagnostics group, centralized repair and maintenance shop of technological equipment of Kovdor Mining and Processing Plant JSC

