Югорский научно-исследовательский институт информационных технологий
Федеральный исследовательский центр «Информатика и управление» РАН
Челябинский государственный университет
Уфимский государственный авиационный технический университет»
Уральский федеральный университет»

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И СИСТЕМЫ

Труды Седьмой Всероссийской научной конференции с международным участием

Ханты-Мансийск, Россия,
12–16 марта 2019 года

Научное электронное издание

Ханты-Мансийск,
2019

ББК 381
И741

ОРГАНИЗАТОРЫ КОНФЕРЕНЦИИ:

Югорский научно-исследовательский институт информационных технологий

Челябинский государственный университет

Федеральный исследовательский центр «Информатика и управление» РАН

Уфимский государственный авиационный технический университет

Уральский федеральный университет

Подготовлено по решению редакционно-издательского совета

Информационные технологии и системы [Электронный ресурс]: тр. Седьмой
И741 Междунар. науч. конф., Ханты-Мансийск, Россия, 12–16 марта 2019 г. (ИТиС – 2019): науч. электрон. изд. (1 файл 28,0 Мб)/ отв. ред. Ю.С. Попков, А.В. Мельников. Ханты-Мансийск 2019.

ISBN 978-5-6042173-7-5

Тематика конференции ориентирована на прикладные исследования в области информационных технологий и систем с использованием математического моделирования, теории управления, компьютерной лингвистики, интеллектуальных интернет-технологий и семантического web, инженерии знаний, онтологии, управления знаниями, цифровой обработки изображений и сигналов, разработки интеллектуальных и мультиагентных систем, разработки систем поддержки принятия решений, распределенных и облачных вычислений, системного анализа и моделирования бизнес-процессов.

Предназначено для ученых, специалистов и преподавателей, работающих в области информатики и вычислительной техники.

Мероприятие поддержано Российским фондом фундаментальных исследований, проект № 19-07-20047.

ББК 381я431+397я431

Подписано в печать 15.04.2019 г.
Формат 60x80/16. Печ. л. 1,0. Печать цифровая.
Тираж 100 Заказ №1591
Отпечатано ООО «Печатный мир г. Ханты-Мансийск»
г. Ханты-Мансийск, ул. Мира, 46.
Тел. +7 (3467) 33-49-91

ISBN 978-5-6042173-7-5

МЕЖДУНАРОДНЫЙ ПРОГРАММНЫЙ КОМИТЕТ

Председатель (сопредседатели) программного комитета:

Юрий Соломонович Попков, академик РАН (ФИЦ «Информатика и Управление», Институт системного анализа РАН, Россия).

Заместитель председателя программного комитета:

Андрей Витальевич Мельников, доктор технических наук, профессор (Югорский научно-исследовательский институт информационных технологий, Россия).

Программный комитет:

B. Blessing (University of Applied Sciences, Austria);

F.-M. Suchanek (Telecom ParisTech University, France);

V. Kober (Center for Scientific Research and Higher Education at Ensenada, Mexico);

J. Ovtcharova (Karlsruhe Institute of Technology, Germany);

B. Wenzel (University of Applied Sciences, Austria);

В.Л. Арлазаров, чл.-корр РАН (ФИЦ «Информатика и Управление», Институт системного анализа РАН, Россия);

В.И. Васильев (Уфимский государственный авиационный технический университет, Россия);

В.Э. Вольфенгаген (Национальный исследовательский ядерный университет «МИФИ», Россия);

А.В. Вохминцев (Челябинский государственный университет, Россия);

С.В. Матвеев, академик РАН (Челябинский государственный университет, Россия);

А.В. Маслобоев (Кольский научный центр, Россия);

Л.В. Массель (Институт систем энергетики РАН, Россия);

А.Г. Олейник (Кольский научный центр, Россия);

Г.С. Осипов (ФИЦ «Информатика и Управление» Институт системного анализа РАН, Россия, Россия);

Ю.М. Полищук (Югорский научно-исследовательский институт информационных технологий, Россия);

В.А. Путилов (Кольский научный центр, Россия);

А.А. Петунин (Уральский федеральный университет, Россия);

С.А. Петунин (ФГУП «Всероссийский научно-исследовательский институт автоматики» им. Н.Л. Духова, Россия);

Ю.В. Петриченко (Челябинский государственный университет, Россия);

С.Г. Пятков (Югорский государственный университет, Россия);

С.В. Смирнов (Институт проблем управления сложными системами РАН, Россия);

Н.И. Юсупова (Уфимский государственный авиационный технический университет);

О.Н. Христодуло (Уфимский государственный авиационный технический университет).

СОДЕРЖАНИЕ

СЕКЦИЯ 1. МЕТОДОЛОГИЯ СИСТЕМНОГО АНАЛИЗА. ТЕХНОЛОГИИ БОЛЬШИХ ДАННЫХ. МЕТОДЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ 6

Попков Ю.С. Метод сжатия матриц для рандомизированного машинного обучения регрессионных моделей..... 6
Захаров А.А. Захарова И.Г., Котельников А.А., Панфиленко Д.В., Решетникова Ю.С. Инфраструктура управления данными электронных историй болезни для создания цифрового фенотипа пациента..... 12
Маслов С.Г., Бельтюков А.П. Системный анализ в формировании физико-антропно-технических систем 16
Наймушина А.Г., Чекардовская И.А., Бакановская Л.Н. Системный анализ и обработка экспертной информации для определения степени значимости факторов риска..... 20
Тагирова К.Ф., Кощеев В.А., Вульфин А.М., Миронов К.В. Разработка обобщенного алгоритма обработки данных различной физической природы на основе системного анализа 26

СЕКЦИЯ 2. МАТЕМАТИЧЕСКИЕ МОДЕЛИ И АЛГОРИТМЫ..... 32

Бычин И.В., Гореликов А.В. Ряховский А.В. Алгоритм численного решения начально-краевой задачи о размагничивании проводящего шара..... 32
Куприянов М.А. Оценка погрешности измерения периметра площадного объекта по его цифровому изображению 36
Ротко В.В. О некоторых классах обратных задач для систем конвекции-диффузии и их обобщений..... 39
Сафонов Е.И, Шергин С.Н. Численное решение обратных задач для уравнения квазистационарных электромагнитных волн в анизотропных неметаллических средах..... 44
Тагирова К.Ф., Талипов Д.А., Вульфин А.М., Миронов К.В. Интеллектуальная система управления режимами работы камеры сгорания газотурбинной установки 49
Шориков А.Ф., Калёв В.И. Алгоритм решения задачи минимаксного программного терминального управления расходом топлива средств выведения при наличии фазовых ограничений 54
Юлдашева А.В. О разрешимости одной некорректной задачи для уравнения колебания балки..... 61

СЕКЦИЯ 3. ЦИФРОВАЯ ЭКОНОМИКА 66

Blessing B. Logistics education..... 66
Максименко З.В., Лакман И.А, Бахитова Р.Х. Информационно-аналитическая поддержка оценки и планирования расходов на выхаживание и реабилитацию недоношенных детей, родившихся с очень низкой и экстремально низкой массой тела 69
Мельников В.А., Харченко Н.Д. Оценка предсказательной способности временных рядов на валютном рынке с использованием методов машинного обучения..... 74
Мещеряков Р.В. Инфокоммуникационные технологии управления оказания услуг населению с использованием телевидения..... 79
Петриченко Ю.В., Сушкова О.И. Внутренняя миграция молодежи в современной экономике: исследование причин с целью создания мониторинговой системы 83

СЕКЦИЯ 4. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭНЕРГЕТИКЕ. ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ 88

Деркач Е.В. Функциональное тестирование исполнительного ядра программно реализованного контроллера на этапе разработки программной составляющей..... 88
Костюченко С.В., Черемисин Н.А., Емельянов А.Н. Локализация текущих извлекаемых запасов нефти и прямой расчет охвата вытеснением в геолого-технологических моделях нелинейной фильтрации 94
Массель Л.В. Массель А.Г. Цифровая трансформация энергетики: методы и интеллектуальные технологии научного обоснования стратегических решений..... 101
Миронов К.В., Тришин С.В, Махмутов А.Р., Картак В.М. О применении систем распределенного реестра в энергетике..... 107
Муратов И.Н., Терентьева М.С. Экспериментальное исследование статистических свойств полей термокарстовых озер в криолитозоне Западной Сибири 112
Семенов А.А., Вылков М.А. Использование подсистемы пространственных данный в АИС «Инновации для Югры» 116
Семенов С.П., Ташкин А.О. Оценка эффективности ГИС для маломобильных групп населения..... 121
Юсупова Н.И., Сметанина О.Н., Сазонова Е.Ю., Агадуллина А.И. Data Mining для поддержки принятия решений по совершенствованию психофизической готовности человека к успешной профессиональной деятельности..... 126

СЕКЦИЯ 5. МАШИННОЕ ОБУЧЕНИЕ И ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ 132

Юсупова Н.И., Нургаянова О.С. Обучение нейронной сети для прогнозирования свойств никелевых сплавов на основе генетического алгоритма 132

Аббаси М.-М., Бельтюков А.П. Анализ эмоций из текста на русском языке с использованием синтаксических методов 137

Ботыгин И.А., Шерстнёв В.С. Шерстнёва А.И., Солтаганов Н.А., Тартаковский В.А. Использование рекуррентных нейронных сетей для построения прогнозных моделей метеопараметров 143

Егоров Ю.А. Исследование эффективности применения классических подходов для решения задачи классификации поз человека с использованием скелетной модели 148

Еримбетова А.С., Мурзин Ф.А. Разработка и анализ технологий определения релевантности текста поисковому запросу 152

Загорулько Ю.А., Боровикова О.И. Проблемы построения онтологий научных предметных областей на основе паттернов онтологического проектирования 157

Карпов М.А. Распознавание шума в ответах оракула в процессе активного обучения 162

Кленин Ю.Д. Обзор применения подходов глубокого обучения в задаче структурирования текстовой информации 167

Пачганов С.А. Создание модели для распознавания запрещенной символики на изображениях при ограниченной обучающей выборке 172

Смирнов С.В. Методика разработки нечетких когнитивных карт на основе онтологического анализа экспертных данных 176

Тузов А.В. Использование интеллектуальных методов прогнозирования загруженности процессора в высоконагруженной WEB системе 180

Якимчук А.В. Реализация модели оценки и прогнозирования рисков состояния лесного фонда нефтедобывающего региона 184

СЕКЦИЯ 6. ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ. ОБЛАЧНЫЕ ВЫЧИСЛЕНИЯ И ГРИД-ТЕХНОЛОГИИ 189

Боровский А.С., Боровская Я.А. Обеспечение комплексной безопасности как основа устойчивого социально-экономического развития бизнеса 189

Васильев В.И., Кириллова А.Д., Сулавко А.Е., Жумажанова С.С. Анализ защиты конфиденциальных данных в критической информационной инфраструктуре и использовании биометрической, нейронной сети и криптографических алгоритмов 193

Молородов Ю.И., Загорулько Г.Б. Архитектура и особенности функционирования системы доступа к удаленным хранилищам данных 198

СЕКЦИЯ 7. МЕТОДЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ. ПРИМЕНЕНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ДЛЯ ЖИЗНИ В АРКТИЧЕСКОЙ ЗОНЕ И РАЙОНАХ, ПРИРАВНЕННЫХ К РАЙОНАМ КРАЙНЕГО СЕВЕРА 205

Богданова Д.Р., Котельников В.А. Юсупова Н.И. Модели и методы оценки эффективности процесса оказания услуг в социально-экономических системах 205

Загорулько Г.Б. Методология разработки интеллектуальных СППР и ее применение для задач медицинской диагностики 211

Куликов Г.Г., Антонов В.В., Фахруллина А.Р., Родионова Л.Е. Программно-аналитический комплекс оперирования с множеством реальных и виртуальных объектов по правилам Декартово замкнутой логики в информационном пространстве предметной области 217

Ризванов Д.А., Юсупова Н.И. Алгоритмы интеллектуальной обработки данных при управлении ресурсами в условиях неопределенности и динамичности внешней среды 223

СЕКЦИЯ 8. ДОКЛАДЫ МОЛОДЫХ УЧЕНЫХ 229

Балдунов М.Б. Методические основы количественного анализа систем заводнения нефтяных месторождений 229

Гаянов Р.Ч., Гаянова М.М., Вульфин А.М. Исследование и разработка энергоэффективного метода взаимодействия в гетерогенной сети Интернета вещей с использованием спутниковой связи 235

Завалич К.И., Мельников В.А. Прогнозирование спроса на продукцию производственного предприятия методами машинного обучения 240

Карадобри К.М., Петриченко Ю.В., Полторак С.Н. Автоматизация бизнес-процессов ведения рекламных кампаний как инновация в сфере интернет-маркетинга 245

Леонтьев З.Д., Башаров И.И. Современные методы и технологии создания виртуальных лабораторных работ для технических специальностей 250

Сичинский Д.А., Мельников А.В. Модель оценки IT-сервисов умного города. Точки влияния на общественно-политическую оценку деятельности органов государственной власти 254

УДК 004.852

Ю.С. Попков

МЕТОД СЖАТИЯ МАТРИЦ ДЛЯ РАНДОМИЗИРОВАННОГО МАШИННОГО ОБУЧЕНИЯ РЕГРЕССИОННЫХ МОДЕЛЕЙ

Аннотация

Предлагается двух-этапный метод сжатия матрицы, имитирующей входные данные (прецеденты) в задачах рандомизированного машинного обучении (РМО). Ядро метода составляют операции «прямого» и «обратного» проектирования, оптимизация которых осуществляется посредством максимизации информационной относительной энтропии. Соответствующие этим операциям модули могут применяться в виде последовательной или параллельной процедуры. Разработаны алгоритмы максимизации информационной относительной энтропии, использующие по-координатную модификацию метода проекций градиента.

Ключевые слова: относительная информационная энтропия, операторы проектирования, матричные производные, градиентный метод, прямые и обратные проекции.

Введение

Процедуры рандомизированного машинного обучения (РМО) являются эффективным инструментальным средством решения задач классификации, регрессии и кластеризации в условиях неопределенности. Источниками ее являются модели и данные, входящие в состав обучающих коллекций. Обучающие коллекции включают две компоненты. Входные данные, характеризуемые матрицами определенных размеров $(m \times s)$, где m – количество объектов в обучающей коллекции, и s – размерность пространства признаков, в терминах которых характеризуются объекты. Выходные данные характеризуются m-мерным вектором, компонентами которого являются параметры принадлежности классу, либо параметры состояния объектов.

Основу концепции рандомизированного машинного обучения составляет гипотеза о случайной природе данных. Вообще говоря, обучающая коллекция (обе ее компоненты) содержит ошибки, но в процедурах обучения предполагается, что основные ошибки, порождающие неопределенность, сосредоточены в выходных данных.

По разным причинам в перечисленных выше задачах возникает необходимость «сжать» обучающую коллекцию, т.е. входную компоненту описывать ($n \times r$)-матрицей. Содержательно это сводится к редукции пространства признаков и уменьшению количества объектов, на массиве которых проводится обучение модели.

Данная проблема вложена в более общую: приближение заданного набора многомерных точек маломерным аффинным многообразием [4]. Наиболее распространенным для ее решения является метод главных компонент (МГК) [5, 6], его робастные версии [7] и многочисленные эмпирические модификации. Следует также отметить, что МГК представляет собой комбинацию математического метода получения упорядоченной последовательности собственных чисел матрицы данных и эвристических соображений по поводу «обрезания» этой последовательности.

Однако, часто можно задать числа *n, r*. Здесь развиваются энтропийный метод сжатия матрицы данных и его реализации в виде последовательной и параллельной процедур. При последовательном процедуре сначала производится сжатие по размерности признакового пространства ($r < s$), а затем по количеству объектов во входной обучающей коллекции ($n < m$). При параллельном процедуре производится одновременное сжатие по обеим размерностям. Технологически в обеих методах реализуется процедура «прямого» и «обратного» проектирования с помощью матриц-«проекторов», соответствующих размеров. Поскольку метод сжатия ориентирован на приложения в РМО, то матрицы-«проекторы» определяются путем максимизации информационной относительной энтропии [10].

Энтропийное «прямое» и «обратное» проектирование

Рассмотрим пространство R^{NM} матриц $A_{(N \times M)}$ с вещественными элементами, которые принимают случайные значения [1]. Преобразуем матрицу $A_{N \times M}$ к матрице $B_{N \times m}$, где $m < M$. Это можно сделать с помощью двух-шаговой процедуры. Сначала проектируем пространство R^{NM} на пространство R^{Nm} ($R^{NM} \Rightarrow R^{Nm}$) с помощью матрицы-«проектора» $S_{(N \times m)}$. Получаем матрицу $D_{(N \times m)} = A_{(N \times M)} S_{(M \times m)}$. Эту операцию будем назвать «прямым» проектированием.

Переход от пространства R^{NM} к пространству меньшей размерности R^{Nm} связан с потерей информации, заключенной в элементах исходного пространства. Для оценки этих потерь нужно вернуться в исходное пространство, т.е. осуществить операцию «обратного» проектирования: $R^{Nm} \Rightarrow R^{NM}$. Эта операция осуществляется с помощью матрицы-«проектора» $J_{(m \times M)}$.

[1] При этом используется *векторизация* матрицы [12]

Получаем матрицу $X_{(NxM)}= D_{(Nxm)}\, J_{(mxM)} \in R^{NM}$.

Матрицы $X_{(N\times M)}, A_{(N\times M)}$ являются элементами одного пространства, что дает возможность привлечь подходящий критерий оценки качества осуществляемого сжатия исходной матрицы.

Поскольку предполагается, что матрица $A_{(N\times M)}$ имеет элементы со случайными значениями, то целесообразно оценивать качество ее сжатия, которое зависит от матриц-«проекторов» $S_{(N\times m)}, J_{(m\times M)}$, в терминах энтропии, которая является характеристикой неопределенности, свойственной задачам РМО. Здесь будет использоваться информационная энтропия [10] матрицы $X_{(N\times M)}$ относительно реализации матрицы $A_{(N\times M)}$, которую далее будем называть *информационной относительной энтропией*:

$$H(X \mid A) = -\sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij} \ln \frac{x_{ij}}{a_{ij}} = tr\left(X^{\acute{u}} Y\right), \quad (1)$$

где tr – след матрицы, заключенной в скобках; Y – $(N\times M)$ – матрица с элементами:

$$y_{ij} = \ln \frac{x_{ij}}{a_{ij}}. \quad (2)$$

Описанный метод «прямого» и «обратного» проектирования составляет ядро метода сжатия матрицы со случайными значениями элементов.

Последовательная процедура сжатия

Рассмотрим матрицу со случайным происхождением данных U размерности $(m\times s)$ и неотрицательными элементами. Ее нужно «сжать» до матрицы Z, размером $(n\times r), r < s, n < m$. Будем решать эту задачу последовательно в два этапа. Сначала редуцируем матрицу $U_{(m\times s)}$ до матрицы $Y_{(m\times r)}$, а затем последнюю «сожмем» до матрицы $Z_{(n\times r)}$. Указанные операции будем осуществлять с помощью матриц-«проекторов», упомянутых в предыдущем разделе выше.

1. Первый этап. Введем матрицу $Q_{(s\times r)}$ и осуществим операцию «прямого» проектирования для получения сжатой по количеству столбцов матрицы

$$Y_{(m\times r)} = U_{(m\times s)} Q_{(s\times r)}. \quad (3)$$

Введем матрицу $T_{(r\times s)}$, и осуществим операцию «обратного» проектирования. Получим матрицу

$$X_{(m\times s)} = U_{(m\times s)} Q_{(s\times r)} T_{(r\times s)}. \quad (4)$$

Матрица $X_{(m\times s)}$ имеет ту же размерность, что и исходная матрица $U_{(m\times s)}$, и является характеристикой последствий (потерь) операций «прямого» и «обратного» проектирования на первом этапе.

Редукция матрицы данных сопровождается потерей информации, которая заключена в исходной матрице, или увеличением относительной информационной энтропии [10]. Так, для первого этапа данного метода относительная информационная энтропия матрицы $X_{(m\times s)}$ при фиксированной исходной матрице $U_{(m\times s)}$ определяется следующим равенством:

$$H_1(X \mid U) = -\sum_{i=1}^{m} \sum_{j=1}^{s} x_{ij}(Q,T \mid U) \ln \frac{x_{ij}(Q,T \mid U)}{u_{ij}} = H_1(Q,T \mid U) \Rightarrow \max_{(Q,T)\geq 0}, \quad (5)$$

где согласно (4)

$$x_{ij}(Q,T \mid U) = \sum_{\mu=1}^{r} \sum_{\nu=1}^{s} t_{\mu,j} q_{\nu,\mu} u_{i\nu}. \quad (6)$$

В результате максимизации функции (5) получаем энтропийно-оптимальную матрицу-«проектор» $Q^*_{(sxr)}$ и «сжатую» матрицу первого этапа

$$Y^*_{(m\times r)} = U_{(m\times s)} Q^*_{(s\times r)} \quad (7)$$

2. Второй этап. Матрица данных второго этапа $Y_{(m\times r)}$ имеет сокращенное количество столбцов $r < s$, и требуется сжать ее по количеству строк до заданной величины $n < m$. Введем матрицы $B_{(n\times m)}, A_{(m\times n)}$, с помощью которых реализуем снова операции «прямого» и «обратного» проектирования:

$$Z_{(n\times r)} = B_{(n\times m)} Y_{(m\times r)},$$
$$G_{(m\times r)} = A_{(m\times n)} Z_{(n\times r)}. \quad (8)$$

Матрица $G_{(m\times r)}$ имеет ту же размерность, что и исходная матрица второго этапа $Y_{(m\times r)}$, и является характеристикой последствий (потерь) операций «прямого» и «обратного» проектирования на втором этапе. Используем информационную относительную энтропию для характеризации этих потерь:

$$H_2(G \mid Y^*) = -\sum_{i=1}^{m} \sum_{j=1}^{r} g_{ij}(B,A \mid Y) \ln \frac{g_{ij}(B,A \mid Y^*)}{y_{ij}} = H_2(B,A \mid Y^*) \Rightarrow \max_{(B,A)\geq 0}, \quad (9)$$

где

$$g_{ij}(B,A \mid Y^*) = \sum_{\beta=1}^{n} \sum_{\alpha=1}^{m} a_{i,\beta} b_{\beta,\alpha} y^*_{\alpha,j}. \quad (10)$$

$$y^*_{\alpha,j} = \sum_{eta=1}^{s} u_{\alpha,\eta} q^*_{\eta,j}. \quad (11)$$

Энтропийная оптимизация последовательной процедуры

Итак, первый этап последовательной процедуры состоит в определении оптимальных элементов неотрицательных матриц Q, T:

$$(Q^*, T^*) = \arg\max_{(Q,T)\geq 0} H_1(Q,T \mid U). \quad (12)$$

Численное решение этой задачи будем осуществлять с помощью итерационного алгоритма, основанного на методе проекций градиента (комбинация градиентного спуска и отсечения отрицательных элементов).

Компоненты градиента функции $H_1(Q,T \mid U)$ по элементам матриц Q, T имеют следующий вид:

$$\frac{\partial H(Q,T \mid U)}{\partial q_{\nu,\mu}} = \sum_{i=1}^{m} \sum_{j=1}^{s} \frac{\partial h_{ij}(Q,T \mid U)}{\partial x_{ij}} \frac{\partial x_{ij}(Q,T \mid U)}{\partial q_{\nu,\mu}},$$
$$\nu = \overline{1,s}, \mu = \overline{1,r}, \quad (13)$$

$$\frac{\partial H(Q,T \mid U)}{\partial t_{\mu,j}} = \sum_{i=1}^{s} \sum_{j=1}^{m} \frac{\partial h_{ij}(Q,T \mid U)}{\partial x_{ij}} \frac{\partial x_{ij}(Q,T \mid U)}{\partial t_{\mu,j}},$$

$$\mu = \overline{1,r}, j = \overline{1,s}, \quad (14)$$

где

$$h_{ij}(T,Q \mid X) = x_{ij}\ln\frac{x_{ij}}{u_{ij}},$$

$$\frac{\partial h_{ij}(T,Q \mid X)}{\partial x_{ij}} = \ln\frac{x_{ij}}{u_{ij}} + 1, \qquad i = \overline{1,m}, j = \overline{1,s}; \quad (15)$$

$$\frac{\partial x_{ij}((Q,T \mid U))}{\partial q_{\nu,\mu}} = t_{\mu,j}\, u_{i,\nu}, \qquad \nu = \overline{1,s}, \mu = \overline{1,r};$$

$$\frac{\partial x_{ij}((Q,T \mid U))}{\partial t_{\mu,j}} = \sum_{\nu=1}^{s} q_{\nu,\mu} u_{i,\nu}, \qquad \mu = \overline{1,r}; \quad (16)$$

$$i = \overline{1,m}, \quad j = \overline{1,s}.$$

Произведем *векторизацию* [12] матриц Q,T и введем соответствующие векторы q,t. Обозначим $\nabla_Q(q,t)$ – вектор градиента функционала относительной энтропии $H(Q,T \mid U)$ с компонентами (18), $\nabla_T(q,t)$ – вектор градиента функционала относительной энтропии $H(Q,T \mid U)$ с компонентами (13–??).

Алгоритм **max** $\boldsymbol{H_1}$ имеет вид:

а) начальный шаг

$$q^0 > 0, \quad t^0 > 0;$$

б) n-ый итерационный шаг

$$X^n = U\, Q^n\, T^n;$$

$$H(Q^n, T^n \mid U) = \sum_{i=1}^{m}\sum_{j=1}^{r} x_{ij}^n \ln\frac{x_{ij}^n}{u_{ij}} = H^n.$$

$$q^{(n+1)} = \begin{pmatrix} q^n + \gamma_q \nabla_Q(q^n, t^n), \chi Ъ \end{pmatrix}$$

$$q^{(n+1)} \Rightarrow Q^{(n+1)};$$

$$t^{(n+1)} = \begin{pmatrix} t^n + \gamma_t \nabla_T(q^n, t^n), \chi Ъ \\ 0 \end{pmatrix}$$

$$t^{(n+1)} \Rightarrow T^{(n+1)};$$

$$X^{(n+1)} = U\, Q^{(n+1)}\, T^{(n+1)};$$

$$H(Q^{(n+1)}, T^{(n+1)} \mid U) = \sum_{i=1}^{m}\sum_{j=1}^{r} x_{ij}^{(n+1)} \ln\frac{x_{ij}^{(n+1)}}{u_{ij}} = H^{(n+1)}.$$

в) Условие остановки

$$ЪЪ \quad H^{(n+1)} - H^{(n)} \le \delta, \qquad Ъ \; STOP.$$

Аппроксимация функционала относительной энтропии. Иногда может оказаться достаточным для определения элементов редуцированной матрицы Q приближенное выражение *H*-функционала, построенное на аппроксимации логарифмической функции в окрестности точки $x_0 = a$, где $a \ge a_{min}$:

$$\ln x < lna + \frac{1}{a_{min}}(x - a). \quad (17)$$

Применяя эту аппроксимацию для $a = x_{ij}$, получим следующее приближенное выражение:

$$H_1(Q,T \mid U) \approx \sum_{i=1}^{s}\sum_{j=1}^{m} \left[x_{ij}^2(Q,T \mid U) - x_{ij}(Q,T \mid U)u_{ij}\right] = \widetilde{H}(Q,T \mid U). \quad (18)$$

Представим это равенство в матричной форме:

$$\widetilde{H}_1(Q,T \mid U) = tr(XX^{ú}) - tr(XU^{ú}) = ⍰X(Q,T), X(Q,T)⍰ - ⍰X(Q,T), U⍰, \quad (19)$$

где $⍰A,B⍰$ – фробениусово скалярное произведение [12],

$$tr(AB^{ú}) = tr(BA^{ú}) = ⍰A,B⍰ = ⍰B,A⍰. \quad (20)$$

Элементы матриц Q,T будем определять минимизируя функционал $\widetilde{H}_1(Q,T \mid U)$ на множестве неотрицательных матриц Q,T при фиксированной матрице данных U, т.е.

$$(\tilde{Q}^*, \tilde{T}^*) = \arg\min_{(Q,T)\ge 0} \widetilde{H}_1(Q,T \mid U). \quad (21)$$

Алгоритм решения этой задачи использует компоненты градиента функционала $\widetilde{H}_1(Q,T \mid U)$ по матрицам Q и T.

Напомним правила дифференцирования некоторых скалярных функций, зависящих от матриц [12]. Рассмотрим линейное пространство $(m\times n)$-матриц A. Скалярное произведение в этом пространстве:

$$⍰A,B⍰ = tr(AB^{ú}) = tr(BA^{ú}).$$

Норма (фробениусова):

$$\| A \|^2 = ⍰A,A⍰.$$

Из этих определений видно, что матрица рассматривается как вектор, составленный их примыкающих строк (*векторизация* матрицы). Для скалярной функции от матрицы градиент определяется следующим образом. Рассмотрим скалярную функцию $f(A)$. Ее приращение представимо в виде:

$$f(A + \Delta) = f(A) + ⍰H, \Delta⍰ +$$

Здесь – остаток ряда Тейлора. Градиент функции $f(A)$ по матрице A

$$\Delta_A f(A) = H.$$

Приведем простые, но полезные примеры скалярных функций от матрицы:

$$f(A) = ⍰C,A⍰, \qquad \Delta_A f(A) = C.$$

$$f(A) = \| A \|^2, \qquad \Delta_A f(A) = 2A.$$

Используя правилам матричного дифференцирования, изложенным выше, получим следующие выражения для градиентов функционала $\widetilde{H}_1(Q,T \mid U)$ (21):

$$\Delta_Q(Q,T) = \frac{\partial \widetilde{H}_1(Q,T \mid U)}{\partial X}\frac{\partial X}{\partial Q} = 2TPQT - TP. \quad (22)$$

$$\Delta_T(Q,T) = \frac{\partial \widetilde{H}_1(Q,T \mid U)}{\partial X}\frac{\partial X}{\partial T} = 2Q^{ú}PQT - Q^{ú}P, \quad (23)$$

Здесь $P = XX^{ú}$. В этих равенствах приняты следующие обозначения: $\Delta_Q(Q,T), \Delta_T(Q,T)$ – градиенты по матрицам Q и T, соответственно. Согласно (22, 23) градиенты функции $\widetilde{H}(Q,T \mid U)$ по матрицам Q и T представляет собой матрицы размером $(s\times r)$.

Для максимизации функционала $\widetilde{H}_1(Q,T \mid U)$ воспользуемся градиентным спуском с контролем условий неотрицательности матриц U,T:

Алгоритм **max** $\widetilde{\boldsymbol{H}}_{\boldsymbol{1}}$ имеет вид:

а) начальный шаг

$$T^0 > 0, \quad Q^0 > 0;$$

б) n-ый итерационный шаг

$$X^n = U\,Q^n\,T^n;$$

$$H(Q^n, T^n \mid U) = \sum_{i=1}^{m}\sum_{j=1}^{r} x_{ij}^n \ln\frac{x_{ij}^n}{u_{ij}} = H^n.$$

$$Q^{(n+1)} = \left(Q^n + \gamma_Q \Delta_Q(\widetilde{H}(Q^n, T^n \mid U) \geq 0, \quad \chi Ъ \right. \qquad (n+1) \geq 0,$$

$$T^{(n+1)} = \left(T^n + \gamma_T \Delta_T(\widetilde{H}(Q^n, T^n \mid U) \geq 0, \quad \chi Ъ \right. \qquad (n+1) \geq 0,$$

$$X^{(n+1)} = U\,Q^{(n+1)}\,T^{(n+1)};$$

$$H(Q^{(n+1)}, T^{(n+1)} \mid U) = \sum_{i=1}^{m}\sum_{j=1}^{r} x_{ij}^{(n+1)} \ln\frac{x_{ij}^{(n+1)}}{u_{ij}} = H^{(n+1)}.$$

Условие остановки:

$$ЪЪ \quad H^{(n+1)} - H^n \leq \delta, \quad Ъ \; STOP.$$

В результате применения алгоритмов **max** $\boldsymbol{H}_{\boldsymbol{1}}$ или **max** $\widetilde{\boldsymbol{H}}_{\boldsymbol{1}}$ получаем энтропийно-оптимальную матрицу-«проектор» $Q^*_{(s\times r)}$. Следовательно, можем определить матрицу $Y^*_{(m\times r)}$ (7) для перехода ко второму этапу сжатия матрицы U, который характеризуется матрицами-«проекторами» $B_{(n\times m)}$ и $A_{(m\times n)}$.

Оптимальные элементы неотрицательных матриц B, A определяются следующим выражением:

$$(B^*, A^*) = \arg\max_{(B,A)\geq 0} H_2(B, A \mid Y^*). \qquad (24)$$

Численное решение этой задачи будем осуществляеть также, как и на первом этапе, с помощью итерационного алгоритма, основанного на методе проекций градиента (комбинация градиентного спуска и отсечения отрицательных элементов).

Компоненты градиента функции $H_2(B, A \mid Y^*)$ по элементам матриц B, A имеют следующий вид:

$$\frac{\partial H_2(B,A \mid Y^*)}{\partial a_{i,\beta}} = \sum_{i=1}^{m}\sum_{j=1}^{r} \frac{\partial s_{ij}(B,A \mid Y^*)}{\partial g_{ij}} \frac{\partial g_{ij}(B,A \mid Y^*)}{\partial a_{i,\beta}},$$
$$i = \overline{1,n}, j = \overline{1,r}, \beta = \overline{1,n}, \qquad (25)$$

$$\frac{\partial H_2(B,A \mid Y^*)}{\partial b_{\beta,\alpha}} = \sum_{i=1}^{m}\sum_{j=1}^{r} \frac{\partial s_{ij}(B,A \mid Y^*)}{\partial g_{ij}} \frac{\partial g_{ij}(B,A \mid Y^*)}{\partial b_{\beta,\alpha}},$$
$$i = \overline{1,n}, j = \overline{1,r}, \beta = \overline{1,n}, \alpha = \overline{1,m}. \qquad (26)$$

где

$$\frac{\partial s_{ij}(B,A \mid Y^*)}{\partial g_{ij}} = \ln\frac{g_{ij}}{y^*_{ij}} + 1, \qquad i = \overline{1,n}, j = \overline{1,r}; \qquad (27)$$

$$\frac{\partial g_{ij}((B,A \mid Y^*))}{\partial a_{i,\beta}} = \sum_{\alpha=1}^{m} b_{\beta,\alpha} y^*_{\alpha,j}, \qquad \beta = \overline{1,n};$$

$$\frac{\partial g_{ij}((B,A \mid Y^*))}{\partial b_{\beta,\alpha}} = a_{i,\beta} y^*_{\alpha,j}, \qquad \beta = \overline{1,n}, \alpha = \overline{1,m};$$
$$i = \overline{1,m}, \quad j = \overline{1,r}. \qquad (28)$$

Произведем *векторизацию* матриц B, A и введем соответствующие векторы *b, a*. Обозначим $\nabla_B(b,a)$ – вектор градиента функционала относительной энтропии $H_2(B, A \mid Y^*)$ с компонентами (26–??), $\nabla_A(q,t)$ – вектор градиента функционала относительной энтропии $H_2(B, A \mid Y^*)$ с компонентами (??).

Алгоритм **max** $\boldsymbol{H}_{\boldsymbol{2}}$ имеет вид:

а) начальный шаг

$$b^0 > 0, \quad a^0 > 0;$$

б) n-ый итерационный шаг

$$G^n = A^n B^n Y^*;$$

$$H(B^n, A^n \mid Y^*) = \sum_{i=1}^{n}\sum_{j=1}^{r} g_{ij}^n \ln\frac{g_{ij}^n}{y^*_{ij}} = H^n.$$

$$b^{(n+1)} = \left(b^n + \gamma_b \nabla_B(b^n, a^n), \chi Ъ \; b^{(n+1)} < 0; \right.\{$$

$$b^{(n+1)} \Rightarrow Q^{(n+1)};$$

$$a^{(n+1)} = \left(a^n + \gamma_a \nabla_A(b^n, a^n), \chi Ъ \right.$$

$$a^{(n+1)} \Rightarrow A^{(n+1)};$$

$$G^{(n+1)} = A^{(n+1)} B^{(n+1)} Y^*;$$

$$H(B^{(n+1)}, A^{(n+1)} \mid Y^*) = \sum_{i=1}^{n}\sum_{j=1}^{r} g_{ij}^{(n+1)} \ln\frac{g_{ij}^{(n+1)}}{y^*_{ij}} = H^{(n+1)}.$$

в) Условие остановки

$$ЪЪ \quad H^{(n+1)} - H^{(n)} \leq \delta, \qquad Ъ \; STOP.$$

Энтропийная оптимизация параллельной процедуры

Параллельная реализация «прямого» и «обратного» проектирования приводит к следующим матричным равенствам:

$$U_{(m\times s)} Q_{s\times r} = Y_{(m\times r)}, \quad B_{(n\times m)} Y_{(m\times r)} = Z_{(n\times r)},$$
$$Z_{(n\times r)} W_{(r\times s)} = D_{(n\times s)}, \quad E_{(m\times n)} D_{(n\times s)} = X_{(m\times s)}. \qquad (29)$$

Первая цепочка «прямой» проекции с матрицей $Q_{(s\times r)}$ и «обратной» проекции с матрицей $B_{(n\times m)}$ дает редуцированную матрицу $Z_{(n\times r)}$, а вторая – с матрицей $W_{(r\times s)}$ «прямой» проекции и с матрицей $E_{(m\times n)}$ «обратной» проекции дает матрицу $X_{(m\times s)}$.

Из равенств (??) имеем:

$$X_{(m\times s)} = E_{(m\times n)}\left\{\left[B_{(n\times m)}\left(U_{(m\times s)} Q_{(s\times r)}\right)\right] W_{(r\times s)}\right\}. \qquad (30)$$

Скобки в этом равенстве указывают на последовательность операций проектирования.

Элементы матрицы $X_{(m\times s)}$ имеют вид:

$$x_{ij} = \sum_{\mu=1}^{n} e_{i,\mu} \sum_{\nu=1}^{r} w_{\nu,j} \sum_{\beta=1}^{m} b_{\mu,\beta} \sum_{\alpha=1}^{s} u_{\beta,\alpha} q_{\alpha,\nu},$$
$$i = \overline{1,m}, j = \overline{1,s}. \quad (31)$$

Для измерения отклонения преобразованной матрицы $X_{(m\times s)}$ от исходной $U_{(m\times s)}$ воспользуемся относительной энтропией

$$\mathcal{H}(X \mid U) = -\sum_{i=1}^{m} \sum_{j=1}^{s} s_{ij}(X \mid U), \quad (32)$$

где $s_{ij} = x_{ij} \ln \frac{x_{ij}}{u_{ij}}$. (33)

С учетом равенства (31) не трудно видеть, что относительная энтропия (32) есть скалярная функция от матриц $(Q, B, W, E) \geq 0$, т.е.

$$\mathcal{H} = \mathcal{H}(Q, B, W, E). \quad (34)$$

Оптимальные значения элементов этих матриц определяются максимумом функции $\mathcal{H}$:

$$(Q^*, B^*, W^*, E^*) = \arg \max_{(Q,B,W,E)\geq 0} \mathcal{H}(Q, B, W, E). \quad (35)$$

Для численного решения этой задачи воспользуемся по-координатной схемой метода проекций градиента.

В параллельной процедуре выделяется пара матриц Q, B, участвующие в редукции исходной матрицы U по одной переменной ($r < s$), и пара матриц W, E, с помощью которых осуществляется «сжатие» матрицы U по второй переменной ($n < m$). Итерационный шаг по-координатной схемы метода проекций градиента состоит из двух под-шагов: на одном осуществляется итерация по (Q,B)-градиентам, а на другом – по (W,E)-градиентам.

Согласно (31–33) компоненты градиента функции $\mathcal{H}(Q, B, W, E)$ имеют вид:

$$\frac{\partial \mathcal{H}}{\partial q_{\alpha,\nu}} = \sum_{i=1}^{m} \sum_{j=1}^{s} \frac{\partial s_{ij}}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial q_{\alpha,\nu}},$$
$$\frac{\partial \mathcal{H}}{\partial b_{\mu,\beta}} = \sum_{i=1}^{m} \sum_{j=1}^{s} \frac{\partial s_{ij}}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial b_{\mu,\beta}},$$
$$\frac{\partial \mathcal{H}}{\partial w_{\nu,j}} = \sum_{i=1}^{m} \sum_{j=1}^{s} \frac{\partial s_{ij}}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial w_{\nu,j}},$$
$$\frac{\partial \mathcal{H}}{\partial e_{i,\mu}} = \sum_{i=1}^{m} \sum_{j=1}^{s} \frac{\partial s_{ij}}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial e_{i,\mu}}, \quad (36)$$

где

$$\frac{\partial s_{ij}}{\partial x_{ij}} = \ln \frac{x_{ij}}{u_{ij}} + 1,$$
$$\frac{\partial x_{ij}}{\partial q_{\alpha,\nu}} = w_{\nu,j} \sum_{\mu=1}^{n} e_{i,\mu} \sum_{\beta=1}^{m} b_{\mu,\beta} u_{\beta,\alpha}, \quad \alpha = \overline{1,s}, \nu = \overline{1,r};$$
$$\frac{\partial x_{ij}}{\partial b_{\mu,\beta}} = e_{i,\mu} \sum_{\nu=1}^{r} w_{\nu,j} \sum_{\alpha=1}^{s} u_{\beta,\alpha} q_{\alpha,\nu}, \quad \mu = \overline{1,n}, \beta = \overline{1,s};$$
$$\frac{\partial x_{ij}}{\partial w_{\nu,j}} = \sum_{\mu=1}^{n} e_{i,\mu} \sum_{\beta=1}^{m} b_{\mu,\beta} \sum_{\alpha=1}^{s} u_{\beta,\alpha} q_{\alpha,\nu},$$
$$\frac{\partial x_{ij}}{\partial e_{i,\mu}} = \sum_{\nu=1}^{r} w_{\nu,j} \sum_{\beta=1}^{m} b_{\mu,\beta} \sum_{\alpha=1}^{s} u_{\beta,\alpha} q_{\alpha,\nu},$$
$$\mu = \overline{1,n}; \quad (37)$$

Применяя *векторизацию* матриц, перейдем к соответствующим векторам q, b, w, e. Обозначим компоненты градиента по этим векторам $\nabla_q, \nabla_b, \nabla_w, \nabla_e$.

Алгоритм максимизации **max** $\mathcal{H}$ имеет следующий вид:

а) начальный шаг

$$q^0 > 0, \quad b^0 > 0, \quad w^0 > 0, \quad e^0 > 0;$$

б) n-ый итерационный шаг

$$X^n = E^n\{[B^n(UQ^n)]W^n\}$$

$$\mathcal{H}^n = \sum_{i=1}^{m} \sum_{j=1}^{s} x_{ij}^n \ln \frac{x_{ij}^n}{u_{ij}}.$$

$$q^{(n+1)} = \left(q^n + \gamma_q \nabla_q(q^n, b^n, w^n, e^n), \chi\right.Ъ$$

$$q^{(n+1)} \Rightarrow Q^{(n+1)};$$

$$b^{(n+1)} = \left(b^n + \gamma_b \nabla_b(q^n, b^n, w^n, e^n), \chi\right.Ъ\ b^{(n+1)} < 0;\{$$

$$b^{(n+1)} \Rightarrow B^{(n+1)};$$

$$w^{(n+1)} = (w^n + \gamma_w \nabla_w(q^{(n+1)}, b^{(n+1)}, w^n, e^n), \chi Ъ$$

$$w^{(n+1)} \Rightarrow W^{(n+1)};$$

$$e^{(n+1)} = (e^n + \gamma_e \nabla_e(q^{(n+1)}, b^{(n+1)}, w^n, e^n), \chi Ъ$$

$$e^{(n+1)} \Rightarrow E^{(n+1)};$$

$$X^{(n+1)} = E^{(n+1)}\left\{\left[B^{(n+1)}\left(UQ^{(n+1)}\right)\right]W^{(n+1)}\right\};$$

$$\mathcal{H}^{(n+1)} = \sum_{i=1}^{m} \sum_{j=1}^{s} x_{ij}^{(n+1)} \ln \frac{x_{ij}^{(n+1)}}{u_{ij}};$$

в) Условие остановки.

$$ЪЪ\ \ \mathcal{H}^{(n+1)} - \mathcal{H}^{(n)} \leq \delta, \qquad Ъ\ STOP.$$

Выводы

Предлагается двух-этапный метод сжатия матрицы со случайными значениями элементов, имитирующей входные данные (прецеденты) в задачах рандомизированного машинного обучении (РМО). Ядро метода составляют операции «прямого» и «обратного» проектирования, оптимизация которых осуществляется посредством максимизации информационной относительной энтропии. Соответствующие этим операциям модули могут применяться в виде последовательной или параллельной процедуры. Разработаны алгоритмы максимизации информационной относительной энтропии, использующие по-координатную модификацию метода проекций градиента

Список литературы

1. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Классификация и снижение размерности. М., Фининсы и статистика, 1989.

2. Мерков А.Б. Распознавание образов. Введение в методы статистического обучения. М., Едиториал УРСС, 2010.

3. Имельбаев Ш.С., Шмульян Б.Л. Моделирование стохастических коммуникационных систем. // Вильсон А.Дж. Энтропийные методы моделирования сложных систем. М., Наука, 1978, с. 170–234.

4. Bruckstein A.M., Donoho D.L., Elad M. From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images. SIAM Rev. 2009, v.51, No.1, p. 34–81.

5. Кендал М.Дж., Стюарт А. Статистические выводы и связи./ Пер. с англ. М.. Наука, 1973.

6. Jollife I.T. Principal Component Analysis. N.Y. Springer-Verlag, 2002.

7. Поляк Б.Т., Хлебников М.В. Метод главных компонент: робастные версии. Автоматика и Телемеханика, 2017, № 3, с. 130–148.

8. Воронцов К.В. Математические методы обучения по прецендентам. Курс лекций МФТИ, 2006.

9. Jaynes E.T. Information Theory and Statistical Mechanics. Physics Review, 1957, v.106, p. 620–639.

10. Popkov Yu.S. Macrosystem Theory and Applications. Springer, Lecture Notes in Informatics and Control, 1995.

11. Kullback S., Leibler R.A. On information and Sufficiency. Ann. of Math. Statistics, 1951, v.22(1), p. 79–86.

12. Magnus J., Neudecker H. Matrix Differential Calculus with Applications in Statistics and Econometrics. John Willey and Sons, 2007, thied Edt.

УДК 57.087

**А.А. Захаров, И.Г. Захарова, А.А. Котельников,
Д.В. Панфиленко, Ю.С. Решетникова**

ИНФРАСТРУКТУРА УПРАВЛЕНИЯ ДАННЫМИ ЭЛЕКТРОННЫХ ИСТОРИЙ БОЛЕЗНИ ДЛЯ СОЗДАНИЯ ЦИФРОВОГО ФЕНОТИПА ПАЦИЕНТА

Аннотация

Совершенствование системы здравоохранения требует эффективного использования оцифрованных биомедицинских данных, уже накопленных и постоянно пополняемых благодаря широкому внедрению прикладных информационных систем. В качестве одной из приоритетных задач здесь выступает создание цифрового фенотипа пациента на основе данных электронной истории болезни. Для решения этой задачи необходимо опираться на экосистему, обеспечивающую защищенное хранение, обработку и анализ больших объемов разнородной информации (таблицы, изображения, тексты на естественном языке). Цель статьи – исследование возможностей использования методов и технологий Big Data для повторного использования оцифрованных биомедицинских данных на примере проектирования и реализации инфраструктуры управления данными электронных историй болезни, предоставляющей инструменты для создания цифрового фенотипа пациента. На основе обобщенного подхода NoSQL и модели «хранилище документов» разработано информационное хранилище (технология реализации PostgreSQL). Разработана система сервисов для решения задач безопасной выгрузки данных из медицинских информационных систем, защищенного хранения, проверки корректности информации, предварительного анализа и визуализации данных, извлечения значений признаков из неструктурированных документов и формирования выборок для использования методов машинного и глубокого обучения. На базе созданной инфраструктуры исследовалась проблема прогнозирования степени риска болезней системы кровообращения. Для построения модели были сформированы цифровые фенотипы более 10 000 пациентов.

Ключевые слова: электронная история болезни, цифровой фенотип, информационное хранилище, Big Data, NoSQL, XML документ, база данных, медицинская информационная система.

Введение

Одним из естественных следствий широкого внедрения медицинских информационных систем (МИС) является накопление значительных объемов оцифрованных биомедицинских данных (БМД): результатов лабораторных исследований, показателей и изображений электрокардиограмм, описаний и изображений ультразвуковых исследований, протоколов осмотра пациента и др. Однако использование только учетного функционала МИС абсолютно неоправданно с точки зрения всех стейкхолдеров – пациентов, врачей-практиков и исследователей, руководителей системы здравоохранения и страховых компаний. Поэтому повторное использование БМД, как уже накопленных, так и постоянно пополняемых, выступает в качестве одного из перспективных направлений исследований. Особое место здесь занимают данные электронной медицинской карты (ЭМК), в совокупности определяющие цифровой след пациента, на основании которого можно выявить закономерности изменений состояния здоровья, особенности течения заболеваний, качества медицинского обслуживания и др.

При этом собственно новые знания из этих данных можно извлечь, рассматривая значительное число ЭМК и опираясь на методологию Data Mining (в частности, методы машинного и глубокого обучения) и технологии Big Data для кластеризации, классификации пациентов по различным основаниям, построения прогнозных моделей для предсказания риска возникновения и развития различных заболеваний и т.д. В качестве одной из приоритетных задач здесь выступает создание «умного» цифрового фенотипа пациента (Smart ЦФП), представленного именно теми признаками, выделенными из данных ЭМК, которые являются ключевыми в контексте конкретной медицинской проблемы.

Для решения этой задачи необходимо опираться на экосистему, обеспечивающую защищенное хранение, обработку и анализ больших объемов разнородной информации (таблицы, изображения, тексты на естественном языке. Цель настоящего исследования заключается в изучении возможностей методов и технологий Big Data для повторного использования оцифрованных биомедицинских данных на примере проектирования и реализации инфраструктуры управления данными ЭМК, предоставляющей инструменты для создания цифрового фенотипа пациента.

Обзор исследований в области интеллектуального анализа больших биомедицинских данных

В работах [1, 2] представлены результаты исследования вопросов представления, хранения и обработки данных в МИС в контексте специфических для этих систем ограничений, определяющих требования к технологиям организации распределенных информационных хранилищ. Исследованию возможных

структур хранилищ БМД посвящены работы [3, 4], в которых предложено развитие реляционной структуры хранилища за счет интеграции реляционного подхода с объектным и нереляционным, а также использование процессной модели данных и технологий Big Data с учетом специфики лечебно-диагностического процесса и, соответственно, характером порождаемых им информационных потоков.

Решение проблемы нормализации и стандартизации данных ЭМК (в англоязычной терминологии – EHR, Electronic Health Record), поставляемых различными МИС, представлено в работах по проекту SHARP (The Strategic Health IT Advanced Research Projects) [5, 6], в которых подчеркивается принципиальная значимость повторного использования накопленных БМД не только для улучшения качества лечения, но и для проведения биомедицинских исследований и эпидемиологического мониторинга на уровне системы национального здравоохранения. Авторы определяют в качестве важнейшей задачи стандартизацию БМД и предлагают архитектуру платформы, предоставляющей сервисы обмена данными между различными организациями, хранения, выделения структурированных данных из текстов ЭМК различного формата с помощью инструментов Apache clinical Text Analysis and Knowledge Extraction System (cTAKES), а также приведения к стандартной терминологии. Особое значение для постановки и решения исследовательских проблем имеет характер БМД (большие объемы, разноформатность и слабая структурированность), открывающий широкое поле для применения технологий Big Data [7, 8]. Вместе с тем, как отмечается в работах [9, 10], существует целый ряд проблем, связанных как с особенностями перехода к новым методам анализа данных (требования к безопасному распределенному хранению и обработке данных), так и с возможными отклонениями и некорректностью исходных БМД, в конечном счете определяющими точность возможных классификационных или прогнозных моделей и других выводов. Кроме того, в отличие от таких сфер биомедицинской информатики, как компьютерное моделирование работы органов, обработка и анализ изображений, исследования генома, приложение технологий искусственного интеллекта и Big Data для решения задач здравоохранения на основе ЭМК имеет свою специфику [11]. В этом контексте в качестве важнейшей исследователи выделяют проблему извлечения объективной информации из текстов протоколов обследования, входящих в ЭМК, отмечая объективность существования множества систем обработки текстов на естественном языке (NLP, Natural Language Processing), ориентированных на медицинские записи и возможность учета этого опыта при разработке аналогичных систем для текстов на других языках [12–14].

Другой принципиальной особенностью извлечения данных ЭМК является внутренняя структурная сложность, определяющая взаимосвязи между данными обследований и назначениями с учетом хронологии и др. [15, 16]. Это определяет востребованность методов машинного и глубокого обучения уже на этапе извлечения и предварительного анализа данных ЭМК [17, 18]. Наконец, важнейшую роль для корректного извлечения данных и использования фундаментальных предметных (медико-биологических) связей между сущностями при интерпретации БМД, их классификации и построении прогнозных моделей играет система онтологий, характеризующих различные группы заболеваний [19].

Методология исследования

Настоящее исследование выполнялось в рамках междисциплинарного пилотного научно-исследовательского проекта «Умный цифровой фенотип пациента». Проект направлен на повышение эффективности здоровьесбережения в рамках парадигмы «5П Медицина» (профилактика, прогнозирование, персонализация, парсипативность, практичность) на основе использования результатов интеллектуальной обработки медицинских данных и текстов из учетных медицинских систем с помощью методов и технологий машинного обучения и больших данных.

Выбор архитектуры информационного хранилища является ключевым для управления данными. Необходимость объединения разнотипных (и, возможно, потоковых) данных из различных источников, возможности не только для интеллектуального анализа, но и для дальнейшего масштабирования и другие особенности задач повторного использования БМД говорят о том, что информационное хранилище должно разрабатываться как озеро данных (Data Lake) [20, 21].

Исходя из этих предпосылок в основу архитектуры была положена мультислойность информационного хранилища, когда на первом слое обеспечено распределенное хранение БМД в исходном виде (распределенные сырые данные, РСД), в общем случае отличающимся для различных МИС. Эти данные загружаются с сопутствующей обработкой (извлечение метаданных, обезличивание и др.) в следующий слой (предварительно обработанные данные, ПОД). Для слоя ПОД необходима определенность с точки зрения выбора реляционного или не только реляционного (Not only SQL, NoSQL) характера связей. Особенности оцифрованных данных, представленных в ЭМК, а именно, их разнородность, слабая структурированность (наличие различных таблиц и нешаблонизированных текстов на естественном языке) и, наконец, значительные объемы, позволяют говорить о больших биомедицинских данных (BBD, Big Biomedical Data [12]. Именно это определило выбор подхода NoSQL [22–23] для разработки слоя ПОД информационного хранилища. Другое, очень важное для масштабируемости преимущество NoSQL заключается в том, что можно форми-

ровать структуру данных в соответствии со сценарием ее использования. В дальнейшем отдельные данные, необходимые для решения конкретной задачи, извлекаются из данных слоя ПОД и передаются в табличном виде гибкому Smart-слою хранилища – для проверки корректности, предварительного анализа и визуализации, а также формирования наборов данных для исследования медицинских проблем с помощью методов машинного и глубокого обучения.

Инфраструктура управления данными ЭМК в информационном хранилище

Для создания прототипа озера данных «5П Медицина – Big Data» были использованы данные ЭМК более 10 000 пациентов, хранящиеся в МИС SAP for Healthcare (https://www.sap.com/cis/industries/healthcare.html) и 1С: Медицина. Поликлиника (https://solutions.1c.ru/catalog/clinic).

Информационное хранилище реализовано в соответствии с предложенным мультислойным подходом. При выгрузке данных слоя РСД было выполнено обезличивание и в слое ПОД каждому пациенту определен каталог, в котором медицинские записи (результаты лабораторных исследований, описания ультразвуковых обследований и электрокардиограмм, протоколы осмотров и др.) сохранены в отдельных XML файлах. С каждым каталогом также связана извлеченная из ЭМК статическая информация о пациенте (пол, дата рождения). При этом данные слоя ПОД представлены файлами-объектами, которые можно условно разделить на 3 вида: табличные, частично структурированные текстовые, неструктурированные текстовые. Табличные файлы имеют четкую структуру с определенными названиями полей и формируются, как правило, по результатам лабораторных исследований (анализы крови, мочи и др.). Их можно хранить в реляционной базе, позволяющей использовать тип данных XML с возможностью извлечения данных, но в контексте методологии NoSQL (использована СУБД PostgreSQL). Частично структурированные текстовые файлы наряду с однозначно определяемыми и обязательными полями (показатели электрокардиограммы, ультразвукового исследования и т.п.) содержат текстовые заключения в произвольной форме. Наконец, файлы третьего вида формируются на основе протоколов осмотра пациента и содержат неструктурированный текст, в котором наряду с заключением и рекомендациями врача могут содержаться и дополнительные признаки пациента (вес, курение, употребление крепких алкогольных напитков и т.п.). При этом диагноз также может быть записан не только в соответствии с Международной классификацией болезней МКБ-10 (http://mkb-10.com), но и достаточно произвольно (например, I25, ИБС, хроническая ишемическая болезнь сердца). В общем случае такие файлы будут отличатся для различных МИС. Эти особенности определили выбор гибкой структуры хранения данных типа «хранилище документов» и ее интеграции с масштабируемым сервисом излечения информации.

Выводы

В рамках мультислойного подхода был создан прототип озера данных, на основе которого, в качестве ключевой, исследована задача построения прогнозной модели на основе данных ЭМК для ведения пациентов с ишемической болезнью сердца. Дополнительно решались такие частные задачи, как выявление закономерностей в динамике посещений пациентом поликлиники после первого упоминания диагноза категорий I20-I25 (стенокардия, инфаркт миокарда, ишемическая болезнь сердца) в соответствии с МКБ-10, а также построение «умного» цифрового фенотипа пациента с диагнозами I20-I25 с учетом сопутствующих заболеваний категорий I10-I15 (гипертензия) и E10-E14 (сахарный диабет).

Список литературы

1. Белышев Д.В., Кочуров Е.В., 2016. Анализ методов хранения данных в современных медицинских информационных системах. Программные системы: теория и приложения, 7(2 (29)): 85–103.

2. Белышев Д.В., Кочуров Е.В., 2016. Перспективные методы работы с данными в медицинских информационных системах. Программные системы: теория и приложения, 7(3 (30)): 79–97.

3. Малых В.Л., Калинин А.Н., Юсуфов Т.Ш., 2018. Объектно-реляционный подход к построению хранилища данных. Программные системы: теория и приложения, 8(3 (30)): 180–187.

4. Малых В.Л., Михеев А.Е., Рудецкий С.В., 2018. Проблемно-ориентированная модель банка клинических данных. Программные системы: теория и приложения, 9(4): 219–237.

5. Rea, S., Pathak, J., Savova, G., Oniki, T.A., Westberg, L., Beebe, C. E. Chute, C. G. (2012). Building a robust, scalable and standards-driven infrastructure for secondary use of EHR data: the SHARPn project. Journal of biomedical informatics, 45(4): 763–771.

6. Pathak, J., Bailey, K.R., Beebe, C.E., Bethard, S., Carrell, D.S., Chen, P.J., ... & Huff, S.M. (2013). Normalization and standardization of electronic health records for high-throughput phenotyping: the SHARPn consortium. Journal of the American Medical Informatics Association, 20(e2): e341–e348.

7. Raghupathi W., Raghupathi V. (2014). Big data analytics in healthcare: promise and potential. Health information science and systems. 2(1(3) Date Views 30.01.2019 https://link.springer.com/content/pdf/10.1186/2047-2501-2-3.pdf.

8. Manogaran, G., & Lopez, D. (2017). A survey of big data architectures and machine learning algorithms in healthcare. International Journal of Biomedical Engineering and Technology, 25(2–4), 182–211.

9. Peek N., Holmes J.H., Sun J. (2014) Technical challenges for big data in biomedicine and health: data sources, infrastructure, and analytics //Yearbook of medical informatics. 9(1), 42–47.

10. Hripcsak G., & Albers D.J. (2017). High-fidelity phenotyping: richness and freedom from bias. Journal of the American Medical Informatics Association, 25(3), 289–294.

11. Luo J., Wu M., Gopukumar D., & Zhao Y. (2016). Big data application in biomedical research and health care: a literature review. Biomedical informatics insights, 8, BII-S31559. Date Views 30.01.2019 https://europepmc.org/backend/ptpmcrender.fcgi?accid=PMC4720168&blobtype=pdf.

12. Luo L., Li L., Hu J., Wang X., Hou B., Zhang T., & Zhao L.P. (2016). A hybrid solution for extracting structured medical information from unstructured data in medical records via a double-reading/entry system. BMC medical informatics and decision making, 16(1), 114. Date Views 30.01.2019 https://bmcmedinformdecismak.biomedcentral.com/track/pdf/10.1186/s12911-016-0357-5.

13. Kreimeyer K., Foster M., Pandey A., Arya N., Halford G., Jones S.F., ... & Botsis T. (2017). Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. Journal of biomedical informatics, 73, 14–29.

14. Névéol A., Dalianis H., Velupillai S., Savova G., & Zweigenbaum P. (2018). Clinical natural language processing in languages other than english: opportunities and challenges. Journal of biomedical semantics, 9(1), 12. Date Views 30.01.2019 https://jbiomedsem.biomedcentral.com/articles/10.1186/s13326-018-0179-8.

15. Meystre S.M., Lovis C., Bürkle T., Tognola G., Budrionis A., & Lehmann C.U. (2017). Clinical data reuse or secondary use: current status and potential future progress. Yearbook of medical informatics, 26(01), 38–52.

16. Wang Y., Wang L., Rastegar-Mojarad M., Moon S., Shen F., Afzal N., ... & Liu H. (2018). Clinical information extraction applications: a literature review. Journal of biomedical informatics, 77, 34–49.

17. Luo Y. (2017). Recurrent neural networks for classifying relations in clinical notes. Journal of biomedical informatics, 72, 85–95.

18. Wang L., Wang Y., Shen F., Rastegar-Mojarad M., & Liu H. (2018, July). Predicting Practice Setting Using Topic Modeling. In 6th IEEE International Conference on Healthcare Informatics Workshops, ICHI-W 2018 (pp. 62–63). Institute of Electrical and Electronics Engineers Inc.

19. Haendel M.A., Chute C.G., & Robinson P.N. (2018). Classification, ontology, and precision medicine. New England Journal of Medicine, 379(15), 1452–1462.

20. Manogaran G., Thota C., Lopez D., Vijayakumar V., Abbas K.M., & Sundarsekar R. (2017). Big data knowledge system in healthcare. In Internet of things and big data technologies for next generation healthcare (pp. 133–157). Springer, Cham.

21. Maini E., Venkateswarlu B., Gupta A. Data Lake – An Optimum Solution for Storage and Analytics of Big Data in Cardiovascular Disease Prediction System. Date Views 30.01.2019 http://ijcem.org/papers112018/ijcem_112018_05.pdf.

21. Moniruzzaman A. B. M., Hossain S.A. (2013). NoSQL database: New era of databases for big data analytics-classification, characteristics and comparison. arXiv preprint arXiv:1307.0191. Date Views 30.01.2019 https://arxiv.org/ftp/arxiv/papers/1307/1307.0191.pdf.

22. Sharma S., Shandilya R., Patnaik S., Mahapatra A. (2016). Leading NoSQL models for handling Big Data: a brief review. International Journal of Business Information Systems, 22(1), 1–25.

23. Freire S.M., Teodoro D., Wei-Kleiner F., Sundvall E., Karlsson D., Lambrix P. (2016). Comparing the performance of NoSQL approaches for managing archetype-based electronic health record data. PloS one, 11(3), e0150069. Date Views 30.01.2019 https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0150069.

УДК 303.732.4

С.Г. Маслов, А.П. Бельтюков

СИСТЕМНЫЙ АНАЛИЗ В ФОРМИРОВАНИИ ФИЗИКО-АНТРОПНО-ТЕХНИЧЕСКИХ СИСТЕМ

Аннотация

Рассматриваются проблемы и методология системного анализа для построения физико-антропно-технических систем. Раскрывается организация системного мышления, которое гармонизирует субъективные и объективные точки зрения. Рассматриваются формы системного мышления, расслоения описаний и конструктивных процессов, непрерывность процесса понимания и построения систем, преобразования замысла в реализацию на основе информатизации, формализации, математизации и перехода к физическому воплощению. Глубокая дивергенция системных исследований привела к тому, что многие исследователи в силу, прежде всего, ограниченности временных и финансовых ресурсов инкапсулируются или попадают в иллюзию всемогущества. Фактически создается ситуация применения неадекватных средств решения сложных проблем, усложнению самих создаваемых инструментов и процессов их использования, происходит искажение смысла используемых понятий и целей деятельности. Неспособность понимать жизненные процессы рождает иллюзорные или мнимые цели, примитивные оценки и, как следствие, приводит к уничтожению в человеке творца, заменяя его торжествующим потребителем или эгоистичным разрушителем. Выход из сложившейся проблемной ситуации лежит в восстановлении гармонии человека с природной средой и их совместном ноосферном устойчивом развитии. Одним из вариантов такого развития является создание физико-антропно-технических систем, чему и посвящена данная статья.

Ключевые слова: системный анализ, физико-антропно-техническая система, эргатическая система, интерфейсы, когнитивная наука, конвергенция знаний, качество жизни, дескриптивное и конструктивное описание.

Введение

В нашей жизни значимую роль играют сложные системы, ещё не рассмотренные в достаточной мере проводимыми исследованиями. Цель настоящей работы – выявление таких систем, дающих альтернативные и более адекватные пути развития решений проблем, стоящих перед человечеством.

Одна из задач, решаемых для достижения этой цели – рассмотреть отдельный класс систем: физико-антропно-технических систем (ФАТ-систем) – в противовес сложным системам, таким, например, как самолёты, буровые платформы и т.д.

Вторая задача – выявление областей эффективного и надёжного использования таких систем.

Третья задача – разработка методов описания и построения таких систем.

Интересные и глубокие исследования ведутся в эргономике и когнитивных науках по взаимодействию человека и машины, по процессам интеллектуализации, созданию новых интерфейсов (неинвазивные интерфейсы мозг-мозг, мозг-компьютер и др.) [1, 2]. Естественно всё это важно для разработки ФАТ-систем. Но тематика таких систем шире и сложнее потому, что здесь сливаются два потока информации, идущие от абстрактных и физических систем.

Приведём некоторые примеры ФАТ-систем.

Один из примеров ФАТ-системы – инвалид (человек с ограниченными физическими возможностями), встроенный в жизненную среду с помощью подходящих технических средств.

Другой пример – ФАТ-системы для космических аппаратов, с замкнутыми экологическими контурами и внешним управляющим контуром.

Третий пример – ФАТ-системы в сельском хозяйстве (птицефабрика, гидропоника в растениеводстве, дрожжевое и грибное производство).

Говоря о системном анализе, следует иметь в виду, что существуют различные подходы к анализу и синтезу систем в различных областях деятельности. Можно привести следующие примеры:

- теория и методология в создании технических систем, по А.А. Коновалову [3–5];
- системно-объектный подход «Узел – Функция – Объект» для различных технических систем, предложенный группой исследователей под руководством С.И. Маторина [6–9];
- изучение сложных систем как упорядоченного множества элементарных систем в форме триад имеется в работах Б.Г. Ильясова и М.Б. Гузаирова [10–13];
- системный подход для разработки экономических и информационных систем и технологий развивается в работах группы исследователей под руководством В.Н. Волковой [14–16];
- исследование системы «Природа – Общество – Человек» с использованием тензорного подхода представлено в работах П.Г. Кузнецова, Б.Е. Большакова, А.Е. Петрова и других ученых [17, 18];
- системный подход к формированию среды программирования и построения эргатических систем, отражённый в работах А.П. Бельтюков и С.Г. Маслова [19–21].

Все эти подходы в основном описывают внешнюю составляющую мира, окружающего человека. Мы же хотим в ФАТ-системах больше внимания уделить человеческой компоненте. Здесь должна идти речь об объединении гетерогенных знаний.

Другие особенности ФАТ-подхода можно пояснить следующим рядом примеров.

Во-первых, такое понятие, как элемент, выполняющий опорную (одну из системообразующих) функцию, используется в основном в технических системах (короб, станина, рама и т.п.). А что такое опорная функция для человеко-машинной системы, пока не рассматривалось. Заметим, что все исследователи легко выделяют функциональное назначение системы, а тому, за счёт чего сохраняется целостность системы, внимания уделяется меньше.

Во-вторых, важно исследовать такую наиболее сложную связь, как связь между структурой и функцией. Эта связь зачастую ускользает от внимания. Например, такая структура, как колесо, может быть связана со многими функциями. Сравните колесо телеги и велосипеда. Нетрудно видеть, что у колеса велосипеда есть дополнительная функция – поддержка равновесия. Колесо может служить и амортизатором.

В-третьих, в настоящее время фактически не уделяется внимание структуре степеней свобод сложной системы, а также – смене группы свойств её элементов и подсистем, связанных со степенями свободы. Это – реструктуризация степеней свобод.

В-четвёртых, в построении системы выделяются альтернативы, и одна из них становится решением проблемы. В это же самое время другие альтернативы могут быть индикаторами обеспечения надёжности системы. Всё это реализуется в виде активных абстрактных образов не обязательно в материальной форме.

Методика

Для ФАТ-систем важен как материальный, так и идеальный результат, как суть ФАТ-системы, так и суть процессов, происходящих в ней при построении и управлении.

Требуется рассмотрение любой среды как естественного вычислителя – для естественного (отчуждённого или внешнего и внутреннего) компьютинга. Имеется в виду компьютинг, осуществляемый специально созданными искусственными объектами, привлечением для компьютинга природных объектов и выполнение компьютинга внутренними силами «мозговой активности» человека и активности всего его организма, направленной на переработку потоков информации.

В ФАТ-системах, в отличие от искусственных систем, важно не просто функциональное разделение между элементами системы, но и материализация абстрактных сущностей, которая приводит к разумному использованию имеющихся ресурсов, и уровень детализации имеющихся ресурсов. При принятии решения требуется также учесть уровень детализации постановки задачи (сравните, например, использование специально обученной птицы и беспилотного летательного аппарата, возможно, требующего слишком больших усилий при подготовке по сравнению с дрессировкой птицы).

Концептуальный анализ ФАТ-систем производится на основе пяти аспектов [19]:

- морфологического,
- функционального,
- атрибутивного,
- генетического (эволюционного),
- либернетического (основанного на степенях свободы).

Первые четыре аспекта – классические, последний используется реже и мало исследован. Эти аспекты можно проинтерпретировать как с логической, так и с наглядно-геометрической позиции.

Следует рассматривать различные виртуальное пространства ФАТ-систем: псевдо-физическое, абстрактное (отражающее процессы мышления и компьютинга) пространства. Можно также вводить дополненную и дополнительную реальность для ФАТ-систем.

Рассматриваем различные подходы построения систем (и их компьютинга):

- процессный,
- функциональный (аппликативный),
- объектный,
- автоматный,
- событийный.

Опора на различные парадигмы компьютинга даёт более богатую среду реализации взаимодействия искусственной и естественной систем.

Основная часть

О гармонизации процессов математизации, информатизации и физикализации

Гармонизация получается при разумном использовании ресурсов и полной используемости результатов (минимизируя проблему утилизации). Здесь идёт речь о логике ресурсов (их воспроизводства и утилизации отходов). С другой стороны, задачи нужно не просто решить эффективно, но надёжно, безопасно и своевременно. Здесь надо вовремя переключаться между уровнями для уточнения задачи.

Требуется соблюдение баланса между постановками задач, процессами их решения и процессами построения самих постановок задач. В плане физикализации важно рассматривать смену напряжений и расслаблений. Эта интерпретация – не просто аллегория, но некоторая аналоговая модель функционирования системы. Если некоторые информационные процессы ведут себя, как физические, то их полезно рассматривать соответствующими методами. Вообще говоря, ФАТ-системы могут иметь как физические,

так и биологические интерпретации, а не только технические. Например, при соответствующих интерпретациях, можно рассматривать такие расстройства ФАТ-систем, как «инсульт» и «инфаркт».

Необходимо также сохранять баланс между математизацией, информатизацией и физикализацией, не допуская полного ухода в идеальный или виртуальный мир, неявно отказываясь от природных законов и разумных ограничений.

Важные мысли и обсуждения этой проблемы имеются в работах Н. Б. Филимонова [22, 23].

2. Дескриптивно-конструктивные основы и проблемы описания и построения ФАТ-систем

Дескриптивное описание можно интерпретировать как компактную форму представления человеческого знания. Конструктивное же описание – некоторое развёрнутое (ожившее, активное) представление знаний. Следует, однако, заметить, что переход к конструктивному описанию, в силу ряда логико-алгоритмических причин, может оказаться сколь угодно сложным.

Столкновение постановки задачи и дескриптивного описания даёт импульс процессу перехода к конструктивному описанию. Однако, для осуществления этого процесса важны дополнительные составляющие, такие как компьютинг, метод и стратегия (и тактика) решения задачи, методы и способы оценки получаемых результатов. Компьютинг здесь по способу его осуществления может быть разделён на три класса:

• численный (конкретный, опирающийся на определённые меры),

• символьный (абстрактный, в том числе, логический),

• образно-геометрический (чаще ориентированный на взаимодействие с человеком).

Практическая значимость предлагаемых решений

Предлагаемые решения могут применяться при организации процессов образования, научного исследования, коллективного принятия решений и производства. С этой точки зрения надо полностью пересматривать технологическую среду взаимодействия искусственных и естественных систем.

Выводы

Фактически мы уходим от чисто технических и технологических решений и создаём новые основания для построения социально-экономических и социально-природных систем, гармонизируя искусственный и естественный мир. В конечном итоге это должно привести к повышению качества жизни.

К сожалению, системные исследования с трудом переходят к конвергенции систем знаний и их эффективного использования, то есть к синтезу парадигм компьютинга, реализуемого в различных направлениях системного исследования, к исследованию когнитивных оснований процессов и форм мышления, к их эргатическому базису. Явно недостаточно внимания уделяется взаимопроникновению идеальных и материальных форм при построении систем.

Учитывая, что основным системообразующим фактором становится информационный образ, необходимо развивать процессы дивергенции и конвергенции, которые обладают требуемой скоростью и непрерывностью трансформации идеальных образов в материальные. Необходимо перейти на новый формат распространения научно-технической информации на основе активных и живых компьютинговых изданий, которые являются частью ФАТ-систем.

Список литературы

1. Прохоров Н.Л., Баловнев Д.А., Рублев А.А., Красовский В.Е. 2016. Методы реализации неинвазивного интерфейса «мозг – мозг»: Вопросы радиоэлектроники. № 3: 96–103.

2. Сергеев С.Ф., Анохин А.Н., Падерно П.И., Обознов А.А. 2017. II международная научно-практическая конференция «ЭРГО 2016: Человеческий фактор в сложных технических системах и средах»: Психологический журнал. Т.38. № 1: 136–138.

3. Коновалов, А.А. 1993. Теория технических систем. Маркетинговый аспект. Екатеринбург: УИФ «Наука», 311 с.

4. Коновалов, А.А., Николаев, Ю.В., Вершинин Н.Н. 2004. Методология проектирования технических систем. Пенза: Изд-во Пензенск. гос. унта, 420 с.

5. Коновалов, А.А., Николаев, Ю.В. 2009. Методология проектирования технических систем. Ч.2. Ижевск: ИПМ Уро РАН, 2009, 319 с.

6. Маторин С.И., Жихарев А.Г., Зайцева Н.О., Четвериков С.Н. 2014. Системный подход «Узел-Функция-Объект» при моделировании транспортных потоков. Труды ИСА РАН. Т. 64. № 1: 9–18.

7. Petrosov D.A., Lomazov V.A. and Dobrunova A.I. and S.I. Matorin, 2015. Large discrete systems evolutionary synthesis procedure: Biosciences Biotechnology Research Asia. T. 12. № 2: 1767–1775.

8. Matorin, S.I., O.A. Zimovets and A.G. Zhikharev, 2015. The elements of general theory of the systems in terms of system-object approach of «unitfunction-object»: International Journal of Applied Engineering Research: T. 10. № 24: 44831-44837.

9. Matorin S.I., Zhikharev, A.G. 2017. Calculation of the function objects as the systems formal theory basis/ In the Proceedings of the Second International Scientific Conference “Intelligent Information Technologies for Industry” (IITI‟17) (September 14–16, 2017. Varna, Bulgaria). Springer International Publishing. V.1. 2017, pp. 182–191.

10. Гузаиров М.Б., Ильясов Б.Г., Герасимова И.Б. 2007. Системный подход к анализу сложных систем

и процессов на основе триад: Проблемы управления (5): 32–38.

11. Ильясов Б.Г., Закиева Е.Ш., Герасимова И.Б. 2013. Системный подход к построению когнитивной модели качества жизни: Вопросы современной науки и практики. Университет им. В.И. Вернадского. № 3 (47): 214–221.

12. Ильясов Б.Г., Герасимова И.Б., Макарова Е.А., Закиева Е.Ш. 2016. Объектно-субъективный подход к оценке качества жизни: Качество. Инновации. Образование. № 2 (129): 47–57.

13. Ильясов Б.Г., Исмагилова Л.А., Закиева Е.Ш. 2015. Системный подход к построению математической модели социально-экономической и политической системы как единого динамического комплекса: Информатизация образования и науки. №4 (28): 148–156.

14. Volkova V.N., Kudryavtseva A.S., Loginova A.V., Chernyy Y.Y., Leonova A.E. 2018. System analysis of innovative technologies of the industrial revolutions. In the Proceedings of the 3rd International Conference Ergo2018: Human Factors in Complex Technical Systems and Environments, pp. 57–61.

15. Волкова В.Н., Ефремов А.А., Козлов В.Н. 2018. Вклад школы «Системный анализ в проектировании и управлении» в развитие информатики, информационных систем и технологий: Прикладная информатика. Т. 13. № 2 (74): 120–140.

16. Васильев Ю.С., Волкова В.Н., Козлов В.Н. 2018. Основные результаты научно-педагогической школы «Системный анализ в проектировании и управлении». В сб. «Системный анализ в проектировании и управлении: науч. тр. XXII Междунар. науч.-практич. конф. 22–24 мая 2018 г.», Ч.1. Санкт-Петербург, 2018. С. 5–28.

17. Большаков Б.Е. 2011. Наука устойчивого развития. Книга I. Введение. М.: РАЕН, 272 с.

18. Петров А.Е 2011. Тензорный метод двойственных сетей. М.: ООО «Центр информационных технологий в природопользовании». 496 с.

19. Маслов С.Г. 2010. Системный подход к формированию среды программирования: Интеллектуальные системы управления / под ред. С.Н. Васильева. М.: Машиностроение, С. 41–47.

20. Beltiukov, A.P. and S.G. Maslov, 2015. Organizing understanding in the framework of ergatic system. In the CSIT,,2015: Proceeding of the 17th International Workshop on Computer Science and Information Technologies, Rome, Italy, September 22-26, 2015, V.1. Ufa, Russia. Ufa State Aviation Technical University, 2015, pp. 60–64.

21. Beltiukov, A.P. and S.G. Maslov, 2017. Deductive Ergatic Design of Constructive Tasks Solutions. In the CSIT ,,2017: Proceedings of the 17th International Workshop on Computer Science and Information Technologies CSIT,, 2017, Baden-Baden, Germany, 2017, Oct. 8–10. Vol. 1. Ufa, Russia.: Ufa State Aviation Technical University, 2017, pp. 233–240.

22. Н.Б. Филимонов, 2014. «Непостижимая эффективность математики» и кризис современной теории управления. Современные проблемы прикладной математики, информатики, автоматизации и управления // Материалы 4-го научно-технического семинара. 23–27 сентября 2014 г. Севастополь. – Москва: Изд-во ИПИ РАН. С. 16–28.

23. Филимонов Н.Б. 2016. «Непостижимая эффективность» математики в современной теории управления: Математические методы в технике и технологиях. № 11 (93): 3–13.

УДК 303.732.4 007.5

А.Г. Наймушина, И.А. Чекардовская, Л.Н. Бакановская

СИСТЕМНЫЙ АНАЛИЗ И ОБРАБОТКА ЭКСПЕРТНОЙ ИНФОРМАЦИИ ДЛЯ ОПРЕДЕЛЕНИЯ СТЕПЕНИ ЗНАЧИМОСТИ ФАКТОРОВ РИСКА

Аннотация

В статье предложена методика анализа и обработки экспертной информации, которая позволяет выявить наиболее значимые факторы риска в контексте персонифицированного подхода к исследованиям. Особое внимание авторами уделено способу ранжирования факторов и расчётам статистических параметров экспертного опроса. Проведённое исследование открывает возможности разработки системы интеллектуальной поддержки принятия решений обработки экспертной информации.

Ключевые слова: анализ и обработка информации, системы интеллектуальной поддержки принятия решений, персонифицированный подход, факторы риска, сердечно-сосудистые заболевания, артериальная гипертензия.

Введение

В настоящее время активно развиваются методики и инструменты проведения экспертных опросов [1–5] с использованием различных интеллектуальных компьютерных систем в предметных областях, где большую роль играют знания специалистов-экспертов [6–16]. Платформа медицинской науки «Профилактическая среда» позиционирует индивидуально-личностный подход оценки состояния здоровья человека и формирования приверженности здоровому стилю жизни в контексте персонификации факторов риска (ФР) социально значимых болезней [17].

В руководстве 2017 ACC к ФР существенно меняющим риск развития осложнений артериальной гипертензии отнесены: курение, сахарный диабет, гиперхолестеринемия, избыточная масса тела и ожирение, гиподинамия и несбалансированное питание. Факторы, которые трудно изменить – хронические заболевания почек, синдром обструктивных апноэ, низкий социально-экономический и образовательный статус. Возраст и инволюционные изменения, мужской пол и наследственность, по-прежнему остаются немодулируемыми ФР осложнений кардиоваскулярной патологии [18]. В отношении психоэмоционального стресса нет однозначного мнения: с одной стороны, коррекция стресс-индуцированных состояний улучшает качество жизни пациентов, а, с другой – не оказывает существенного значения в снижении риска ССЗ [19, 20].

Очевидное противоречие между персонифицированным подходом и стандартизацией оказания медицинской помощи актуализирует изучение значимости ФР развития сердечно-сосудистых заболеваний с использованием статистического и математического обеспечения систем анализа на основе экспертной информации. Перспектива реализации такого подхода открывает возможности построения системы интеллектуальной поддержки принятия решений как одного из методов ранней диагностики и факторной профилактики хронических неинфекционных заболеваний.

Методы

Для разработки методов получения, анализа и обработки информации, эксперты отвечали на представленную анкету дважды. Им было предложено указать значимость представленных факторов, влияющих на осложнения артериальной гипертензии, т.е. степень их влияния на эффективность ранжирования параметров. Оценивая фактор, специалист присваивает им определенный ранг. Из статистических данных выбраны мужчины, имеющие отдельные критерии выше нормы. Данные приведены в таблице 1.

Условные обозначения:
х – холестерин;
имт – индекс массы тела;
лж – увеличение левого желудочка;
мжп – межжелудочковая перегородка;
зслж – задняя стенка левого желудочка;
сад – систолическое артериальное давление;
дад – диастолическое артериальное давление.

Факторам, в зависимости от их значимости, присваивается значение в пределах «1 ÷ 7» баллов, то есть «1 балл – более значимый ÷ 7 баллов – менее значимый». Если специалист одинаково оценивает силу влияния нескольких факторов, то всем им присваивается одинаковый балл. Результаты опроса сведены в таблицу 2.

Общая сумма баллов по всем строкам (i) должна совпадать с суммой по столбцам (j):

$$\sum_{j=1}^{m}\sum_{i=1}^{k} x_{ij} = \sum_{i=1}^{k}\sum_{j=1}^{m} x_{ij}, \qquad (1)$$

где m = 7 – количество факторов, k = 14 – количество экспертов, x_{ij} – балл j-го фактора, присвоенный i-м экспертом. Далее выполняется ранжирование факторов (таблица 3), то есть их распределение, согласно набранной сумме баллов.

Таблица 1

Мужчины, имеющие заболевания выше нормы

Порядковые номера мужчин																							
1	2	3	6	7	8	9	11	12	14	15	16	17	19	20	22	23	24	25	26	28	30	…	37
Заболевания мужчин																							
х	х		х	х	х	х	х						х					х				…	х
имт	имт		имт		имт	имт	имт	имт					имт	имт			имт					…	имт
лж	лж	лж	лж		лж	лж		лж			лж	лж	лж				лж	лж	лж	лж	лж	…	лж
мжп			мжп					мжп			мжп			мжп			мжп					…	
			зслж					зслж						зслж								…	
				сад					сад	сад					сад	сад	сад					…	
				дад					дад	дад					дад	дад	дад					…	

Таблица 2

Экспертный анализ факторов, влияющих на здоровье мужчин со стресс-индуцированной артериальной гипертонией при превышении норм показателей

Факторы	Эксперты														
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	$\sum_{i=1}^{k} x_{ij}$
1. Холестерин	1	1	4	4	3	1	1	1	1	1	6	6	6	6	42
2. ИМТ	1	2	1	2	6	2	2	1	1	4	7	7	7	2	45
3. ЛЖ	2	2	6	1	6	3	2	5	1	4	6	4	3	2	47
4. МЖП	3	6	7	2	7	5	4	5	1	5	4	4	6	5	64
5. ЗСЛЖ	7	5	6	6	7	7	6	7	7	7	5	6	6	7	89
6. САД	6	5	7	6	5	7	5	6	4	5	5	5	6	7	79
7. ДАД	7	6	7	7	4	6	6	7	5	5	5	6	6	7	84
$\sum_{j=1}^{m} x_{ij}$	27	27	38	28	38	31	26	32	20	31	38	38	40	36	450

Данные таблицы 3 переформировываются таким образом, чтобы:

1) баллы каждого эксперта располагались в порядке возрастания «связанных и несвязанных баллов»;

2) для факторов, имеющих одинаковые баллы, рассчитываются «связанные ранги» (в примере А и таблице 3 обозначено курсивом). Например, двум факторам (1 и 2 ранга) эксперт № 1 присвоил балл 1, значит средний балл в переформированной таблице будет: (1 + 2) /2 = 1,5. Значение «несвязанного балла» должно быть равно соответствующему значению ранга;

3) распределяются значения баллов эксперта № 2, согласно действиям (1) и (2).

Результаты. Практическая значимость

Пример А:

Ранг факторов	1	2	3	4	5	6	7	
Баллы эксперта № 1								
1)	1	1	2	3	7	6	7	таблица 2
2)	*1*	*1*	2	3	6	*7*	*7*	таблица 3
3)	*1.5*	*1.5*	2	3	6	*6.5*	*6.5*	таблица 4

Вышеуказанные действия выполняются с баллами каждого эксперта. Далее рассчитываются статистические параметры экспертного опроса (таблица 4). Следует учитывать корректность работы экспертов. Для контроля составляется таблица 2 и сравниваются результаты ранжирования по таблице 3.Общая сумма баллов по всем строкам (i) должна совпадать с суммой по столбцам (j) в обеих таблицах.

Проверяется корректность заполнения таблицы 4. Очевидно следующее:

1) максимальный ранг оценки (α_{max}) по конкретному фактору не может быть больше числа факторов (m). В нашем случае α_{max}= 7, и ≤m = 7, то есть условие: $\alpha_{max} \leq m$ выполнено;

2) максимальное значение суммы рангов по любому фактору не может быть больше произведения максимально возможного ранга на число экспертов, т.е.

$$\sum_{i=1}^{k} x_{max} \le \alpha_{max} \cdot k \qquad (2)$$

где x_{ij} – балл j-го фактора, присвоенный i-м экспертом, α_{max}= 7, k = 14 – количество экспертов. Для проверки находим в таблице 4 максимальное значение суммы рангов по фактору 7:

$$\sum_{i=1}^{k} x_7 = 84{,}5 \le 7 \cdot 14 = 98$$

Значит второе условие выполнено;

3) минимально возможная сумма рангов по любому фактору не может быть меньше минимального ранга (α_{min}), умноженного на количество экспертов (k), т.е.

$$\sum_{i=1}^{k} x_{min} \ge \alpha_{min} \cdot k \qquad (3)$$

Для проверки находим в таблице 4 минимальное значение суммы рангов по фактору 1:

$$\sum_{i=1}^{k} x_1 = 16{,}5 \ge 1 \cdot 14 = 14$$

Так как три условия выполнены, то правильность заполнения таблицы 4 несомненна. Общая сумма рангов по экспертам и по факторам должна совпадать. В нашем примере общая сумма рангов

$$\sum_{j=1}^{m} \sum_{i=1}^{k} X_{ij} = 391 \text{ (таблица 4)}$$

и средняя сумма рангов:

$$\frac{\sum_{j=1}^{m} \sum_{i=1}^{k} X_{ij}}{m} = \frac{391}{7} = 55{,}86.$$

Отклонение суммы рангов всех экспертов по каждому фактору $\left(\sum_{i=1}^{k} x_{ij} \right)$ от средней суммы рангов определяется по уравнению:

$$\Delta = \left(\sum_{i=1}^{k} x_{ij} - \frac{\sum_{j=1}^{m} \sum_{i=1}^{k} x_{ij}}{m} \right) \qquad (4)$$

Таблица 3

Ранжирование факторов, влияющих на здоровье мужчин

Факторы	Эксперты														$\sum_{i=1}^{k} x_{ij}$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
	Ранги связанных (*обозначены курсивом*) и несвязанных баллов														
1. Холестерин	*1*	1	1	1	1	1	1	*1*	*1*	1	1	1	1	1	14
2. ИМТ	*1*	*2*	2	2	2	2	*2*	*1*	*1*	*4*	2	2	4	2	29
3. ЛЖ	2	*2*	*7*	3	*7*	3	*2*	*5*	*1*	*4*	*7*	*7*	*7*	5	62
4. МЖП	3	4	*7*	4	*7*	5	4	*5*	*1*	*5*	*7*	*7*	*7*	*7*	73
5. САД	6	5	*7*	5	*7*	6	5	6	4	*5*	*7*	*7*	*7*	*7*	86
6. ДАД	*7*	6	*7*	6	*7*	*7*	*6*	*7*	5	*5*	*7*	*7*	*7*	*7*	89
7. ЗСЛЖ	*7*	*7*	*7*	*7*	*7*	*7*	*6*	*7*	7	7	*7*	*7*	*7*	*7*	97
$\sum_{i=1}^{m} x_{ij}$	27	27	38	28	38	31	26	32	20	31	38	38	40	36	450

Таблица 4

Распределение значения баллов и статистические параметры экспертного опроса

Факторы m_i	Эксперты, k_j														Сумма рангов СР $\sum_{i=1}^{k} x_{ij}$	Отклонение СР Δ	$S = \Delta^2$	Занимаемое место M_i
	Ранги оценок																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14				
	Ранги оценок α_j																	
1.	*1,5*	1	1	1	1	1	1	*1,5*	*2,5*	1	1	1	1	1	16,5	-39,36	1549,21	**1**
2.	*1,5*	*2,5*	2	2	2	2	*2,5*	*1,5*	*2,5*	*2,5*	2	2	2	2	29	-26,86	721,46	**2**
3.	2	*2,5*	*5*	3	*5*	3	*2,5*	*3,5*	*2,5*	*2,5*	*5*	*5*	*5*	5	46,5	-9,36	87,61	**3**
4.	3	4	*5*	4	*5*	5	4	*3,5*	*2,5*	*5*	*5*	*5*	*5*	*5,5*	61,5	5,64	31,81	**4**
5.	6	5	*5*	5	*5*	6	5	6	5	*5*	*5*	*5*	*5*	*5,5*	73,5	17,64	311,17	**5**
6.	*6,5*	6	*5*	6	*5*	*6,5*	*6,5*	*6,5*	6	*5*	*5*	*5*	*5*	*5,5*	79,5	23,64	558,85	**6**
7.	*6,5*	7	*5*	7	*5*	*6,5*	*6,5*	*6,5*	7	7	*5*	*5*	*5*	*5,5*	84,5	28,64	820,25	**7**
$\sum_{i=1}^{m} x_{ij}$	27	28	28	28	28	30	23	29	28	28	28	28	28	30	391		4080,36	

Например, для первого фактора:

$$\Delta_1 = \left(16{,}5 - 55{,}86\right) = -39{,}36.$$

Квадрат отклонения по первому фактору:

$$S_1 = \Delta^2 = \left(-39{,}36\right)^2 = 1549{,}21.$$

Сумма квадратов отклонений всех оценок рангов каждого фактора от среднего арифметического рангов (таблица 4):

$$S_{\sum} = \sum_{j=1}^{m} \Delta^2 = 4080{,}36$$

Для оценки средней степени согласованности опрашиваемых экспертов коэффициент конкордации (согласования) несвязанных рангов [6, 7]:

$$W = \frac{12S}{k^2\left(m^3 - m\right)} = \frac{12 \cdot 4080{,}36}{14^2\left(7^3 - 7\right)} = 0{,}743 \qquad (5)$$

где m = 7 – количество факторов, k = 14 – количество экспертов.

При полном совпадении мнений экспертов W=1, при несовпадении W=0, при W= 0,743> 0,4 ÷ 0,5 – совпадение мнений удовлетворительно.

Так как в матрице таблице 3 имеются связанные ранги, то коэффициент конкордации уточняется по формуле [1, 2]:

$$W = \frac{S}{\frac{1}{m}k^2 \cdot \left(m^3 - m\right) - k \cdot \sum_{i=1}^{k} T_i}, \qquad (6)$$

где $\sum_{i=1}^{k} T_i = \frac{1}{14} \sum_{j=1}^{m} \left(t^3 - t\right)$, здесь t – число связанных рангов в каждом столбце рангов.

В таблице 3 связанные ранги обозначены курсивом. Величина T_i определяется для каждого столбца, а затем суммируется:

$T_1 = (1/14)\cdot[(2^3–2) +(2^3–2)] = 0{,}857$; $T_1 = T_7$;

$T_2=(1/14)\cdot[(2^3–2)] =0{,}4286$; $T_2 = T_5 = T_6$;

$T_3=(1/14)\cdot[(5^3–5)] =8{,}571$; $T_3= T_{11} = T_{12} = T_{13} = 8{,}571$;

$T_4= 0$;

$T_8=(1/14)\cdot[(2^3–2)+(2^3–2)+ (2^3–2) +(2^3–2)] =1{,}286$;

$T_9 = (1/14)\cdot(4^3–4) = 4{,}286$; $T_9 = T_{14} = 4{,}286$;

$T_{10} = (1/14)\cdot[(2^3–2) +(3^3–3)] = 2{,}143$;

$T_{14}=(1/12)\cdot[(3^3–3)+(3^3–3)+ (2^3–2) +(3^3–3)] =6{,}5$;

Значит

$$\sum_{i=1}^{k} T_i = 0{,}857 + 0{,}42864 + 8{,}571 + 0 + 0{,}4286 + 0{,}4286 + 0{,}857 + 1{,}286 + 4{,}286 + 2{,}143 + 8{,}571 + 8{,}571 + 8{,}571 + 6{,}5 = 51{,}5.$$

Коэффициент конкордации определяется по формуле (6):

$$W = \frac{4080{,}36}{\frac{1}{7} \cdot 14^2 \cdot \left(7^3 - 7\right) - 14 \cdot 51{,}5} = \frac{4080{,}36}{9408 - 721} = 0{,}47$$

Следовательно мнение экспертов согласовано, так как при W= 0,47 >0,4÷0,5 – совпадение мнений удовлетворительно.

Значимость коэффициента конкордации проверяется сравнением расчетного и критического значения χ^2 – критерия Пирсона [1, 2], который подчиняется нормальному распределению с числом степеней свободы (m -1):

- при наличии несвязанных рангов

$$\chi_p^2 = k(m-1)\cdot W = \frac{12S}{km(m+1)} \quad (7)$$

- при наличии связанных рангов

$$\chi_p^2 = \frac{S}{\frac{1}{12}km(m+1) - \frac{1}{(m-1)\sum_{i=1}^{k} T_i}} = \frac{4080,36}{\frac{1}{12}\cdot 14\cdot 7(7+1) - \frac{1}{(7-1)\cdot 51,5}} = 62,45 \quad (8)$$

Для числа степеней свободы (m -1)= 7–1= 6 определяется табличное критическое значение $\chi^2_{кр}$= 18,48 для уровня статистической значимости q = 0,01 [1, 2].

Сравнивая расчетные значения $\chi^2_р$ = 62,45 с табличными критическими значениями $\chi^2_{кр}$= 18,48 для уровня значимостиq = 0,01, можно говорить о достаточной степени согласованности мнений экспертов и результатов исследования.

По сумме рангов производится распределение места по степени значимости фактора. Минимальной сумме рангов соответствует наиболее важный фактор, получающий первое место (M_1 =16,5), далее факторы располагаются по мере возрастания суммы рангов. Результаты экспертного опроса для выявления значимых факторов позволили определить следующую их важность:

1 место – 16,5 баллов;
2 место – 29 баллов;
3 место – 46,5 баллов;
4 место – 61,5 балла;
5 место – 73,5 баллов.

Выводы

В результате экспертного анализа влияния важности факторов определенно предполагаем, что выбранные факторы играют главную роль, то есть являются ведущим звеном в определении здоровья мужчин.

Данное исследование системных связей между факторами одного заболевания позволяет выявить персонифицированные закономерности развития гипертонической болезни, повысить эффективность управления течением заболевания и совершенствованием подходов с использованием современных методов обработки информации. Далее можно приступить к аппаратной реализации программы на основе визуализации и анализа обработки экспертной информации.

Список литературы

1. Williams C. Informashion processing in electronic medical records: A survey validation / C. Williams, H. Hamadi, C. Cummings, N. Zakari // Journal of Evaluation in Clinical Practice. – Volume 25. – 2019. – PP. 97-103. DOI: 10.1111/jep.13017.

2. Levy S. Information systems that support effective clinical decision making / S. Levy, B. Heyes, // Nursing Management. – Volume 19. – 2012. – PP. 20–22. DOI: 10.7748/nm2012.11.19.7.20.c9371.

3. Vlahos G.E. The use of computer-based information systems by German managers to support decisionmaking / G.E. Vlahos, T.W. Ferratt, G. Knoepfle // Information and Management. – Volume 41. – 2004. – PP. 763-779. DOI: 10.1016/j.im.2003.06.003.

4. Говорухина А.А. Взаимосвязь параметров фи-зического и психического здоровья и компонентного состава организма женщин, работающих в нефтегазовой отрасли / А.А. Говорухина, Е.Н. Слюсарь // Вестник нижневартовского государственного университета. – 2017. – № 4. – С. 88 –95.

5. Информационная технология экспертного опроса специалистов нефтегазовой отрасли для предотвращения пожаров на объектах защиты / Н.Ю. Зуев [и др] // Пожаровзрывобезопасность. – 2018. – № 5. – Т. 18. – С. 17–25.

6. Лукиянчина Е.В. Привлекательность регионов Крайнего Севера для креативного класса / Е.В. Лукиянчина // Вестник Сургутского государ-ственного университета. – 2018. – № 1 (19). – С. 57–62.

7. Moser T. Efficient automation systems engineering process support based on semantic integration of engineering knowledge / T. Moser, R. Mordinyi, D. Winkler, M. Melik-Merkumians, S. Biffl // IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2011, 16th Conference on Emerging Technologies and Factory Automation. – 2011. DOI: 10.1109/ETFA.2011.6059098.

8. Ordonez C. A tool for statistical analysis on network big data / C. Ordonez, T. Johnson, D. Striyastava, S. Urbanek // International Workshop on Database and Expert Systems Applications. DEXA. – Volume 2017. – 2017. – PP. 32–36. DOI: 10.1109/DEXA.2017.23.

9. Бакановская Л.Н. К вопросу о разработке нейросетевых экспертных систем [Текст] / Л.Н. Бакановская // Conduct of modern scince – 2015 Materials of the XI International scientific and practical conference. – 2015. – PP. 11–14.

10. Bakanovskaya L.N. Improvement of information processing quality in preproduction engineering efficiency assessment / L.N. Bakanovskaya, I.A. Chekardovskaya // IOP conference series: Materials science and engineering. "International Scientific-Practical Conference of Students, Graduate Students and Young Scientists "Transport and Storage of Hydrocarbons"". – 2016. – P. 012018. DOI: 10.1088/1757-899X/154/1/012018.

11. Ahmadulin R.K. Object-oriented programming when developing software in geology and geophysics / R.K. Ahmadulin, L.N. Bakanovskaya // IOP conference series: Earth and environmental science. Current Problems and Solutions. "Ecology and Safety in the Technosphere:

Current Problems and Solutions". – 2017. – P. 012049. DOI: 10.1088/1755-1315/50/1/012049.

12. Bublyk M. Structuring the fuzzy knowledge base of the IT industry impact factors / M. Bublyk, O. Rybytska, A. Karpiak, Y. Matseliukh // 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 – Proceedings. – Volume 1. – 2018. – PP. 21–24. DOI: 10.1109/STC-CSIT.2018.8526760.

13. Yusupbekov N.R. Knowledge-based planning for industrial automation systems: The way to support decisionmaking / N.R. Yusupbekov, S.M. Gulyamov, S.S. Kasimov, N.B. Usmanova // Advances in Intelligent Systems and Computing, 13th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2018. – Volume 896. – 2019. – PP. 873-879. DOI: 10.1007/978-3-030-04164-9_115.

14. Sohail M.N. A comprehensive looks at data mining techniques contributing to medical data growth: A survey of researcher reviews / M.N. Sohail, R. Jiadong, M.N. Uba, M. Irshad // Advances in Intelligent Systems and Computing. International Conference on Intelligent Computing, Communication and Devices, ICCD 2017. – Volume 752. – 2019. – PP. 21–26. DOI: 10.1007/978-981-10-8944-2.

15. Melnykova N. The special ways of application of neural networks for medical information processing / N. Melnikova, P. Mukalov, D. Koziy // Conference on Computer Sciences and Information Technologies, CSIT 2018. 13th International Scientific and Technical. – Volume 1. – 2018. – PP. 428–431.

16. Мелихова Ю.М. К вопросу использования ком-пьютерных технологий для анализа информации в научном исследовании / Ю.М. Мелихова, Л.Н. Еременко // Научный альманах. – 2017. – № 2 – 2(28). – С. 175–179. DOI: 10.17117/na.2017. 02.02.175.

17. Портал «Медицинская наука» [Электронный ресурс]. – Режим доступа: http://medical-science.ru/?page_id=2963.

18. 2017ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA. Руководство по профилактике, выявлению, оценке и управлению высоким кровяным давлением у взрослых. Резюме докладов американской коллегии кардиологов, американской кардиологической ассоциации по клиническим практикам. – DOI 10.1161/ HYP0000000000000066. – Режим доступа http://hyper.ahajournals.org.

19. Наймушина А.Г. Распространенность факторов артериальной гипертензии в открытых популяциях г. Тюмени у лиц со стабильными базовыми социальными характеристиками / А.Г. Наймушина, С.В. Соловьева // Медицина труда и промышленная экология. – 2011. – № 4. – С. 15–19.

20. Наймушина А.Г. Гендерные аспекты стресс-индуцированной артериальной гипертензии / А.Г. Наймушина, С.В. Соловьева // Вестник Тюмен-ского государственного университета. – 2014. – № 6. – С.150–156.

21. Иванов В.А. Эффективность и оптимизация ор-ганизации производства при обслуживании вспомога-тельного оборудования на компрессорных станциях / В.А. Иванов, И.Г. Волынец, И.А. Чекардовская // Известия вузов. Нефть и газ. – 2009. – № 1. – С. 48–52.

22. Чекардовская И.А. Методика анализа орга-низационно-технического уровня производства / И.А. Чекардовская // Сборник материалов: Всероссийская научно-практическая конференция «Актуальные проблемы строительства, экологии и энергосбережения в условиях западной Сибири». – ТюмГАСУ, 2008. – 289 с.

УДК 004.9 303.732

К.Ф. Тагирова, В.А. Кощеев, А.М. Вульфин, К.В. Миронов

РАЗРАБОТКА ОБОБЩЕННОГО АЛГОРИТМА ОБРАБОТКИ ДАННЫХ РАЗЛИЧНОЙ ФИЗИЧЕСКОЙ ПРИРОДЫ НА ОСНОВЕ СИСТЕМНОГО АНАЛИЗА

Аннотация

Предложена методология интеллектуального когнитивного проектирования систем поддержки принятия решений по выбору эффективных методов обработки первичных данных различной физической природы в области автоматического контроля, диагностики, управления. На основе системного анализа известных задач и методов их решения в различных предметных областях сформирована база типовых задач, обобщенного представления целей, критериев качества результата для итеративной координации процедур принятия решений.

Ключевые слова: методология интеллектуального когнитивного проектирования, система поддержки принятия решений, электронный кадастр, адаптивный интерактивный анализ.

Введение

Современные системы контроля, диагностики, управления характеризуются повышенными требованиями к точности, достоверности обработки данных при все возрастающем объеме и сложности обрабатываемых данных. С другой стороны, использование современных интеллектуальных технологий предоставляет широкие возможности в обеспечении качества функционирования таких систем, в том числе за счет автоматизации основных операций. Предлагается концепция структурно-логического подхода к автоматизации выбора методов обработки данных различной физической природы, основанная на системном подходе, ориентирующем на достижение целостности разрабатываемой информационной системы с учетом многообразия связей и свойств субъектов (данных, методов обработки). Соответствующие процедуры исследования заключаются в разделении сложной задачи на отдельные этапы, в выборе на основе логического системного анализа наиболее подходящих специальных методов для их решения, которые позволяют затем синтезировать общее решение для определенного класса задач с заданным уровнем качества. Процесс синтеза итеративный, с оценкой эффективности принятых решений для достижения определенной цели в автоматическом режиме или с участием лица, принимающего решение. Основой эффективности методологии является качество системного анализа задач в рассматриваемых предметных областях.

В результате проведенных исследований [1–13] осуществлен анализ алгоритмов обработки данных различной физической природы (технологические временные ряды, оптические изображения с различных источников и носителей информации, акустические сигналы). Выявлены специфические и типовые процедуры входного контроля сигналов, предварительной обработки, целевой обработки и поддержки принятия решений в конкретных задачах (пороговая фильтрация, нормирование, статистический анализ, корреляционный и регрессионный анализ, распознавание, кластеризация, прогноз, оптимизация).

Проведенные исследования [10–12] позволили расширить эмпирическую базу для разработки методологии обоснования выбора интеллектуальных методов обработки различной физической природы и поддержки принятия решений в сложных динамических технических системах.

Автоматизация отдельных процедур, связанных с обработкой данных (в особенности – выбора метода обработки), благодаря использованию типовых решений и алгоритмов позволяет существенно сократить временные и материальные затраты при создании систем обработки данных (например, систем технического зрения), снизить требования к квалификации и специализации разработчиков.

Цель работы – автоматизация процедуры выбора и реализации методов и алгоритмов обработки данных. Для достижения этой цели работы необходимо решить задачу разработки автоматизированного выбора алгоритмов обработки данных с учетом класса задачи, её особенностей, целей, а также параметров исходных данных.

Системный анализ типовых данных и задач в системах поддержки принятия решений

Как в любой интеллектуальной системе, основная сложность заключается в создании системы формализованного представления данных, методов и алгоритмов их обработки. Предложено использовать нормализованные шкалы, которые в каждой предметной области основаны на обобщенных представлениях технических характеристик и требований. Например, в области обработки изображений это может быть диапазон оптических частот в изображении, для электрических сигналов (технологические временные ряды) – спектральный состав, диапазон амплитуд. Для пространства методов и алгоритмов формализуется диапазон заданного качества в виде нормированной шкалы, связанной с конкретной точностью и достоверностью принятия решений с использовани-

ем конкретных интеллектуальных технологий. Для нейросетевых методов и алгоритмов это может быть: архитектура (сверточная, рекуррентная, прямого распространения, гибридная) и параметры архитектуры нейронной сети - количество нейронов и слоев, функции активации нейронов, параметры алгоритма обучения и оптимизатора, параметры процедуры тестирования и верификации сети, параметры оценки обобщающей способности и требуемых вычислительных ресурсов.

На основе системного анализа известных задач в области измерения, контроля, диагностики и управления сформулированы следующие требования к кадастру методов и алгоритмов обработки данных и алгоритму поиска в нём.

1. Требования к структуре кадастра [8–12]:

1) гибкость структуры кадастра, позволяющая вносить изменения в описания задач, методов обработки, и типов данных без нарушения его работы;

2) наличие возможности дополнять кадастр после его создания, и, соответственно, специальных инструментов его администрирования;

3) наличие функции выявления методов и алгоритмов обработки, а также типов данных, являющихся в информационном смысле тождественными или содержательно противоречивыми;

4) наличие иерархической классификации имеющихся в базе типов данных, методов и алгоритмов их обработки, а также способов их применения;

5) наличие раздела «синонимы» и возможность их распознавания;

6) наличие списка литературы в текстах описания.

2. Требования к процедуре поиска:

1) простота и гибкость формализации поискового запроса, простота поисковой модели и наглядность визуального представления различных характеристик данных и методов обработки;

2) простота работы с полученными результатами, возможность повторного поиска по полученным результатам и применение различных фильтров;

3) возможность сохранять и пользоваться результатами предыдущих (удачных) поисков.

3. Требования к описаниям элементов кадастра:

1) структурированность текста описаний;

2) стандартизированные названия для типов данных, методов их обработки, а также для их характеристик;

3) поддержка графических и символьных изображений, формул, табличных данных в выходной карте описания метода обработки и типов поддерживаемых данных;

4) наличие примеров использования методов обработки данных.

Выявлены следующие наиболее актуальные и часто используемые типы данных: «плоские» таблицы (flat table); реляционные таблицы; временные ряды (включая технологические); текстовые данные; графические данные (в том числе видео); аудиоданные; транзакции, наборы; пространственные данные и данные, связанные с мобильными устройствами; XML; Web-контент [1–7].

Необходимо отметить, что не все элементы этого списка обязательно находятся на одном уровне предполагаемой системы классификации.

В начале сессии выбора процедуры обработки данных пользователь, в зависимости от обстоятельств (наличия данных, наличия средств телеметрии на предприятии) может располагать разнородной информацией о своих задачах. Необходимо предусмотреть следующие ситуации:

1. Пользователь имеет данные для обработки и знает, какую информацию ему необходимо из них извлечь.

2. Пользователь имеет определённые данные и хочет узнать, какую информацию из них можно извлечь.

3. Пользователю необходимо получить некую целевую информацию, но у него нет необходимых данных, и он хочет узнать, какие данные ему нужны.

Процесс работы системы состоит из пяти основных этапов:

1. Задание цели. Задание цели должно определять одну из описанных выше ситуаций.

2. Загрузка данных и выбор класса задачи, а также конкретизации задачи (целевой информации). После выбора цели пользователь загружает в систему имеющиеся у него данные и выбирает класс задач из списка, из которого система предварительно исключает те задачи, для решения которых загруженный пользователем тип данных не подходит.

3. Проверка качества и структуры данных. Под качеством данных подразумевается, например, качество изображений для диагностики технических объектов, которое нередко регламентируется в ГОСТ (например, требования к контрастности [14]). Требования к структуре данных подразумевают их конкретную организацию структуры. Например, наличие всех необходимых проекций объекта для обработки данных компьютерной томографии. В случае, если предоставленные данные не удовлетворяют требованиям качества или структуры, система сообщает об этом пользователю.

4. Выбор и применение процедуры обработки данных [15, 16]. На основании полученных на предыдущих этапах данных производится принятие решения о процедуре для решения поставленной пользователем задачи. Оценка качества данных на этом этапе зависит от требований (к зашумленности изображения, к количеству изображений для анализа.

5. Получение результатов. После того, как был осуществлен выбор процедуры обработки, система использует эту процедуру для получения целевой информации [17–25] в виде: двоичные изображения с выделенными на них областями интереса, массивы векторов признаков этих выделенных областей.

На этом этапе существует возможность вывода описания выполненной процедуры.

На этапах 2–5 система должна располагать данными следующего вида:

- классы задач обработки данных;
- задачи обработки внутри этих классов;
- существующие требования для типов данных (в соответствии с ГОСТ, если он имеется);
- существующие требования к структуре данных (для решения задачи обработки);
- процедуры, с помощью которых решаются конкретные задачи – их может быть несколько.

В свою очередь, «процедуры» должны представлять собой связанный набор данных, состоящий из:

- требований, предъявляемых этими процедурами к исходным данным и условий использования процедур;
- информации об операциях (методах, алгоритмах) (или, иначе говоря, «ссылок» на эти операции в виде вызова функций, например) из которых состоит процедура, а также параметрах этих операций, если они есть;
- способов представления информации, которая получается на выходе после применения соответствующих процедур.

Всё это может быть организовано в виде единой базы типовых задач обработки данных, алгоритмов их решения, методов обработки данных.

Разработка кадастра методов и алгоритмов обработки данных

Конечный вид алгоритма, который соответствует вышеперечисленным этапам взаимодействия пользователя с системой, представлен на рисунке 1.

В начале процесса происходит задание пользователем общей цели. Затем осуществляется загрузка данных. На основе загруженных данных система должна выводить перечень классов задач (на основе информации в базе типовых задач), соответствующих загруженным пользователем данным, и пользователь осуществляет выбор класса задачи. Далее выводится перечень конкретных задач в выбранном пользователем классе, которые могут быть решены. Когда задача выбрана, осуществляется проверка качества и структуры загруженных пользователем данных – они должны соответствовать предъявляемым требованиям к качеству и структуре, которые выражаются с помощью объективных параметров данных, рассчитываемых в соот-

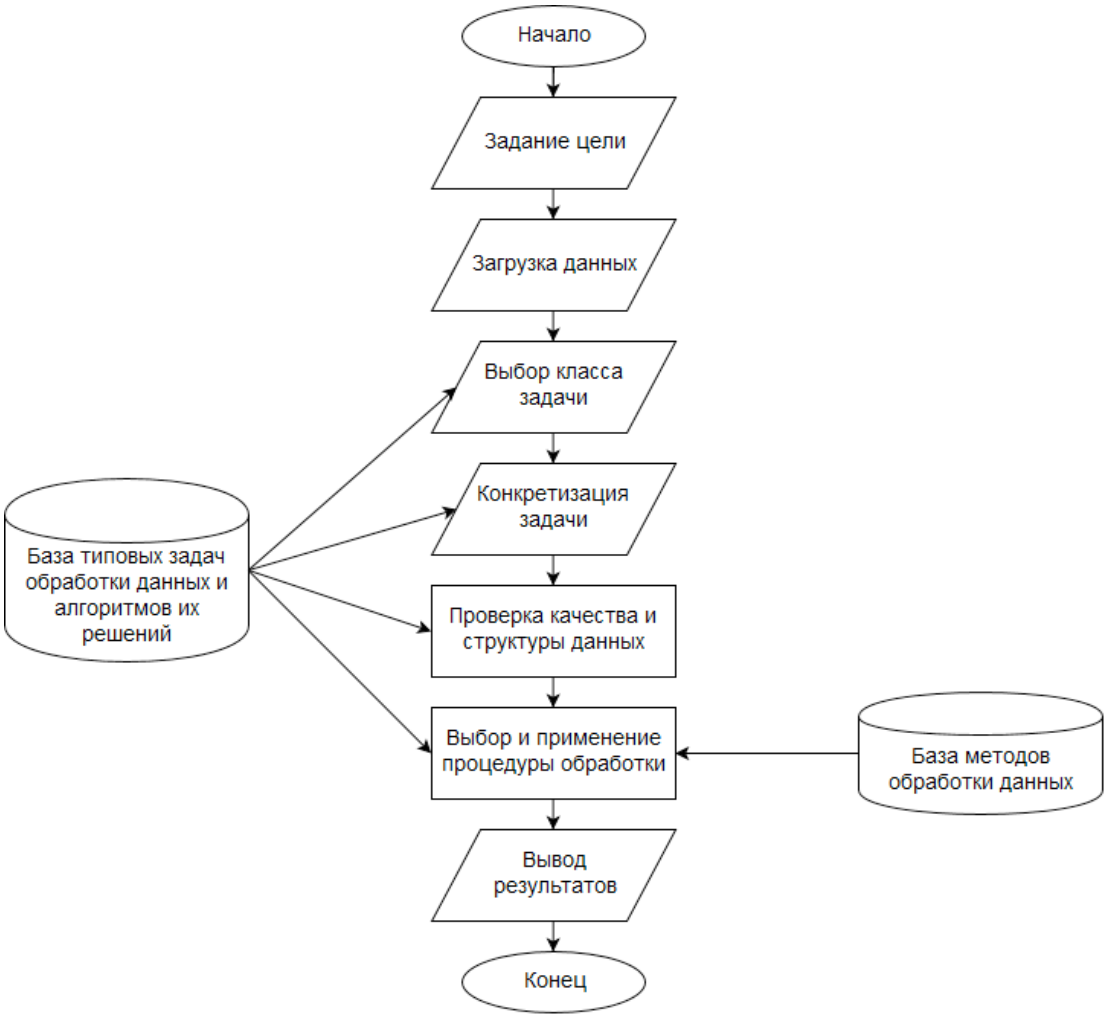


Рис. 1. Общий вид алгоритма работы системы

ветствии с информацией в базе типовых задач (каждый элемент «задача» должен сопровождаться перечнем параметров данных, которые нужно рассчитать для оценки качества и структуры). После этого система производит выбор процедуры обработки на основе информации о процедурах, имеющихся в базе типовых задач для заданной пользователем задачи – здесь учитываются условия использования процедур, которые должны быть описаны для каждой отдельной процедуры.

Применение процедуры обработки осуществляется с использованием базы методов обработки данных. Наконец, система производит вывод результатов в зависимости от запросов пользователя.

Для работы описанной системы, представляющей собой формализованный кадастр методов обработки данных, созданы: база методов обработки данных, база типовых задач обработки данных и алгоритмов их решений с возможностью расширения и соответствующим алгоритмом поиска имеющихся в базах методов, процедур и задач.

На рисунке 2 приведен алгоритм работы кадастра методов и алгоритмов обработки данных.

Разработаны алгоритмы основных подпрограмм формализованного кадастра:

- подпрограмма поиска методов, процедур, и задач обработки данных (предоставляет пользователю возможность вести поиск как в соответствии с обозначенными полями описания методов, процедур, и задач обработки, так и с помощью одиночных ключевых фраз).

- подпрограмма обработки данных (расчёт параметров данных в соответствии с имеющимися условиями и требованиями к данным, на основе которых может приниматься решение об использовании той или иной процедуры обработки, или метода в составе этих процедур; проверка качества и структуры данных: формат, размерность, наличие специальных идентификаторов или метаданных, и т.д.).

- подпрограммы пополнения базы методов и базы типичных задач (проверка введённых значений для исключения дублирования методов и процедур, а также некорректных последовательностей использования методов при несоответствии выходного типа данных предыдущего метода входному типу данных текущего). Для пополнения базы методов требуется указание входного и выходного типов данных, введение программного кода метода, описание параметров метода (количества и структуры), ввод описания к добавленному методу обработки. Аналогично, при пополнения базы типовых задач требуется указание входного и выходного типов данных, ввод условий использования процедуры, ввод последовательности

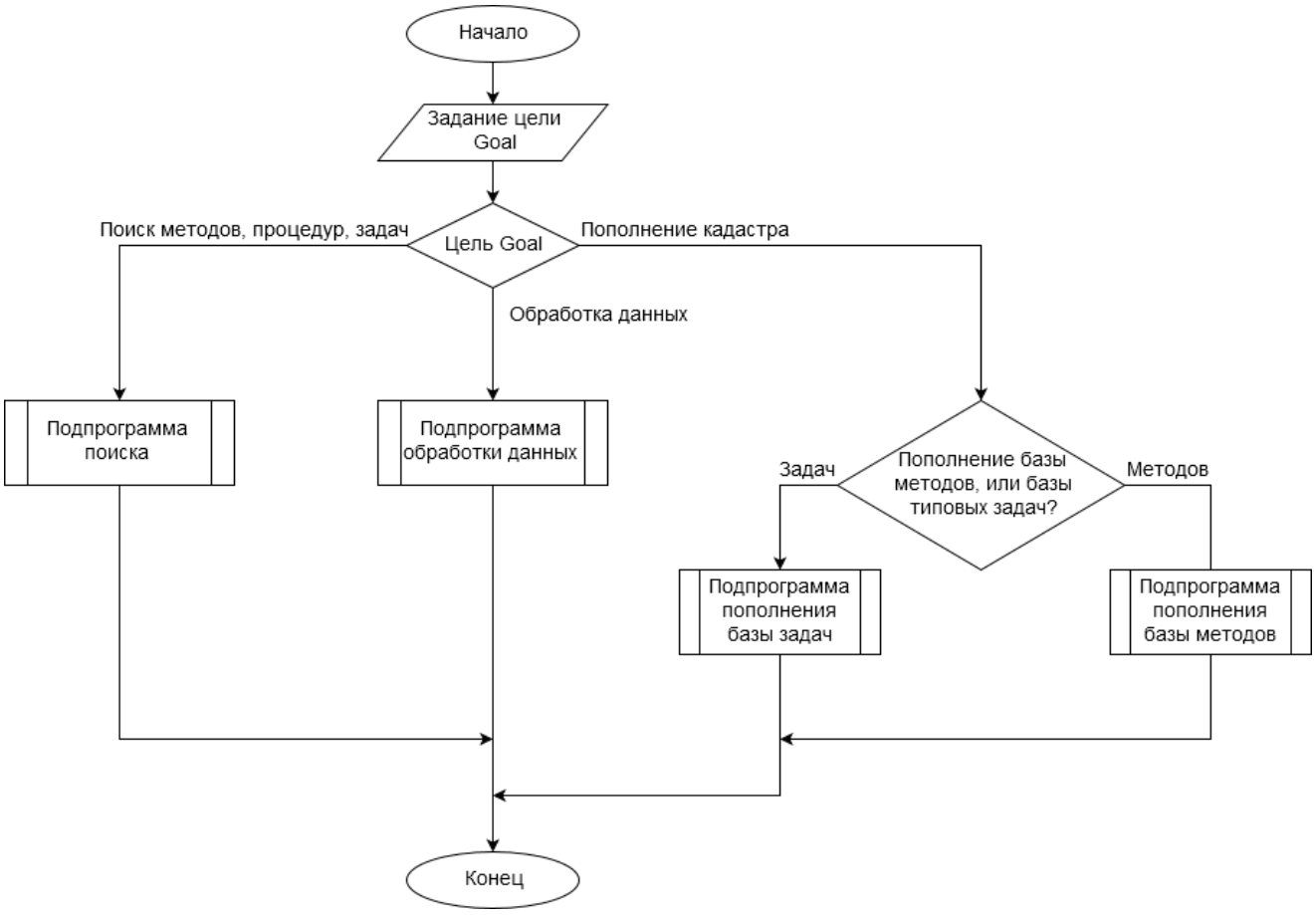


Рис. 2. Алгоритм работы кадастра методов и алгоритмов обработки данных

методов обработки, их параметров, а также условий использования методов, ввод описания к добавленной процедуре обработки.

Реализация разработанных алгоритмов обеспечит повышение эффективности систем диагностики и управления за счет использования адаптивного алгоритма анализа и обработки входных данных на основе комплексного критерия «качество данных – качество диагностики – качество управления».

Выводы

Таким образом, разработана методология автоматизированного выбора методов и алгоритмов обработки данных различной физической природы. Разработаны алгоритмы для реализации предложенного подхода, позволяющие учитывать класс задачи, её особенности, цели, а также параметры исходных данных. Разработана структура базы типовых задач обработки данных и алгоритмов их решений, а также структура базы методов обработки данных. В совокупности указанные результаты образуют электронный кадастр методов обработки данных, предназначенного для интеллектуального когнитивного проектирования систем поддержки принятия решений по выбору эффективных методов обработки первичных данных различной физической природы в области автоматического контроля, диагностики, управления.

Ценность и актуальность методологии автоматического синтеза алгоритмов интеллектуального анализа, соответствующих природе входных данных и качеству принимаемых решений, существенно возрастает в рамках реализации перспективной концепции цифрового производства в различных отраслях промышленности и цифровой экономики в целом.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта 17-08-01569 «Разработка методологии выбора методов интеллектуального анализа и обработки данных в задачах диагностики и управления сложными техническими системами на основе структурно-логического подхода».

Список литературы

1. Loskot P., 2018. Automation is Coming to Research. IEEE Signal Processing Magazine, May 2018, 35(4). Date Views 15.01.2019 https://www.researchgate.net/publication/325177106_Automation_is_Coming_to_Research.

2. Липатова С.В., А.А. Смагин, А.М. Фирулин, П.И. Смикун, В.Г. Захаров, 2016. Методика построения системы компактного хранения данных, предназначенных для аналитической обработки Журнал Автоматизация процессов управления, 4(46): 66–75.

3. Razu Ahmed, Arifa Khatun, Asraf Ali and Kenneth Sundaraj, 2018. A literature review on NoSQL database for big data processing. International Journal of Engineering and Technology, 7(2): 902–906.

4. Sagiroglu S., D. Sinanc. (2013) Big data: A review. Collaboration Technologies and Systems (CTS), 2013 International Conference on Digital Object Identifier, pp. 42–47.

5. Feldman R., G.G. Curry, 2011. Manufacturing Systems Modeling and Analysis. Berlin, Heidelberg: Springer Verlag, pp: 335.

6. Wang Jun, (2001). Computational Intelligence in Manufacturing Handbook. Boca Raton: CRC Press LLC, pp: 560.

7. Lin Gu, Deze Zeng, Peng Li and Song Guo, 2014. Cost Minimization for Big Data Processing in Geo-Distributed Data Centers. IEEE Transactions on Emerging Topics in Computing, 2(3): 314–323.

8. Стулова Е.В., С.А. Фоменков, Д.М. Коробкин, 2013. Система формирования и модификации описания физических эффектов в базе данных. Известия Волгоградского государственного технического университета, 14 (117), с. 126–131.

9. Глазунов В.Н., 2007. БАЗА ДАННЫХ «ЭФФЕКТЫ 200». В кн.: «ТРИЗ-Конференция – 2007», Москва, 7–8 июля 2007.

10. Tagirova, K.F., A.M. Vulfin, 2018. Automation of cognitive design of visual information processing systems. Journal of Advanced Research in Technical Science, 12: 90–96.

11. Tagirova K.F., A.R. Ramazanov, A.M. Vulfin and A.V. Nikonov, 2018. Storage subsystem for intelligent oil wells control system. In the Proceedings of the International Conference Information Technology and Nanotechnology. Session Data Science, Samara, Russia, pp. 255–261.

12. Koshcheev V.A., Tagirova K.F. and Vulfin A.M., 2018. Automatic formation of complex image processing. In the Proceeding of All-Russian Scientific Conference “Information Technologies for Intelligent Decision Making Support”, Ufa-Stavropol, Russia, 2018, Vol.2, pp. 153–156.

13. Chaware S.R., N.H. Khandare, 2016. Image noise removal using different median filtering techniques A review. International Journal of Scientific & Engineering Research, Vol. 7, Issue 2, February-2016, pp. 279–283.

14. ГОСТ 23479-79. Контроль неразрушающий. Методы оптического вида. Общие требования (с Изменениями № 1, 2)

15. Habib A.A., 2016. Statistics and Mean Filters Based on Spatial Noise: A Comparative Study. International Journal of Advanced Research in Computer and Communication Engineering, 5(8): 495–500.

16. Parfenov, E.V., A. Yerokhin and A. Matthews, 2009. Small signal frequency response studies for plasma electrolytic oxidation. Surface & Coatings Technology, 203(19): 2896-2904.

17. Injadat M.N., F. Salo and A.B. Nassif, 2016.

Data mining techniques in social media: A survey. Neurocomputing, Vol. 214: 654–670.

18. Roiger R.J., 2017. Data mining: a tutorial-based primer. Chapman and Hall/CRC, pp: 487.

19. Hong M. et al., 2016. A unified algorithmic framework for block-structured optimization involving big data: With applications in machine learning and signal processing. IEEE Signal Processing Magazine, 1(33): 57–77.

20. Zaharia M. et al., 2016. Apache spark: a unified engine for big data processing. Communications of the ACM, 11(59): 56–65.

21. Armbrust M. et al., 2015. Spark SQL: Relational data processing in spark. In the Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, Melbourne, Victoria, Australia, 2015, pp. 1383–1394.

22. Zheng, Z. et al., 2015. Real-time big data processing framework: challenges and solutions. Applied Mathematics & Information Sciences, 6(9): 3169–3190.

23. Ji, C. et al., 2012. Big data processing in cloud computing environments. In the Proceedings of the 2012 12th International Symposium on Pervasive Systems, Algorithms and Networks. IEEE, 2012, pp. 17–23.

24. Wu, X. et al., 2014. Data mining with big data. IEEE transactions on knowledge and data engineering, 1(26): 97–107.

25. Dittrich, J., J.A. Quiané-Ruiz, 2012. Efficient big data processing in Hadoop MapReduce. In the Proceedings of the VLDB Endowment, Vol. 5, Issue 12, pp. 2014-2015.

УДК 519.63

И.В. Бычин, А.В. Гореликов, А.В. Ряховский

АЛГОРИТМ ЧИСЛЕННОГО РЕШЕНИЯ НАЧАЛЬНО-КРАЕВОЙ ЗАДАЧИ О РАЗМАГНИЧИВАНИИ ПРОВОДЯЩЕГО ШАРА

Аннотация

Рассматривается начально-краевая задача с вакуумными граничными условиями о диффузии магнитного поля в проводящем шаре. В рамках метода контрольного объема разработан алгоритм численного решения данной задачи. Представлены результаты тестовых расчетов. Работа выполнена при поддержке РФФИ, грант: 16-29-15105 офи_м.

Ключевые слова: диффузия магнитного поля, вакуумные граничные условия, численное решение.

Введение

При математическом моделировании явления гидромагнитного динамо в задачах астрофизики и геодинамики возникает необходимость решать уравнение магнитной индукции с вакуумными граничными условиями [1–5]. При постановке задач с вакуумными граничными условиями магнитное поле считается всюду непрерывным, а в вакууме потенциальным и соленоидальным [6]. Уравнение диффузии магнитного поля, которое рассматривается в данной работе, является частным случаем квазистационарного уравнения индукции для неподвижной среды [7–13]. При численном решении, как уравнения диффузии, так и уравнения индукции, возникает проблема получения соленоидального магнитного поля. В настоящее время наиболее распространенными методами вычислительной магнитной гидродинамики являются псевдоспектральные [14, 15] и конечно-разностные [16] методы, а также метод контрольного объема [15, 17]. В рамках данных методов существуют различные подходы для обеспечения бездивергентности магнитного поля [18]: использование векторного потенциала; процедура коррекции магнитного поля путем прибавления к нему поправочного поля, т.е. метод искусственного скалярного потенциала [19]; применение смещенных сеток и запись дискретного аналога уравнения индукции в такой форме, чтобы его решение автоматически удовлетворяло сеточному уравнению неразрывности [20–22] (такой подход использован и в данной работе). Несмотря на достаточно большое число работ в этой области, разработка эффективных алгоритмов решения уравнения индукции с вакуумными граничными условиями остается актуальной задачей вычислительной магнитной гидродинамики.

Постановка задачи и математическая модель

В данной работе рассматривается задача о размагничивании находящегося в вакууме однородного и изотропного проводящего шара G радиуса a. Магнитная проницаемость полагается равной единице. Задача решается в сферических координатах r, θ, φ. Предполагается, что в вакууме (т.е. в области $R^3 \setminus \overline{G}$) магнитное поле потенциально, соленоидально и, следовательно, его скалярный потенциал является регулярной на бесконечности гармонической функцией. На границе шара G все компоненты магнитного поля полагаются непрерывными. Математическая модель состоит из уравнения диффузии магнитного поля (1) в области G ($\mathbf{B}(\mathrm{r},t)$ – индукция магнитного поля, ν_m – коэффициент магнитной диффузии), уравнения Лапласа на скалярный потенциал $u(\mathbf{r},t)$ магнитного поля в вакууме (при $r \ge a$ $\mathbf{B} = \nabla u$), граничного условия (3), начальных условий (4), (5) и условия согласования начальных и граничных условий (6):

$r < a,\ t > 0$:

$$\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t} = -\nu_m \mathrm{rot}\,\mathrm{rot}\,\mathbf{B}(\mathbf{r},t), \quad (1)$$

$r > a,\ t > 0$:

$$\nabla^2 u(\mathbf{r},t) = 0 \ (u(\mathbf{r},t) \text{ регулярна при } r \to +\infty).$$

Граничное условие ($r = a,\ t > 0$): $\mathbf{B} = \nabla u$. (3)

Начальные условия $t = 0$:

$$r \le a: \ \mathbf{B} = \mathbf{b}(\mathbf{r}), \text{ где } \operatorname{div} \mathbf{b}(\mathbf{r}) = 0, \quad (4)$$

$$r > a: \ u = u_0(\mathbf{r}), \text{ где } \nabla^2 u_0(\mathbf{r}) = 0, \ (u_0 \text{ регулярна при } r \to +\infty) \quad (5)$$

$$r = a: \ \mathbf{B} = \nabla u_0(\mathbf{r}). \quad (6)$$

Соленоидальность магнитного поля в начальный момент времени (4) обеспечивает соленоидальность поля ($\mathbf{B}(\mathrm{r},t)$ в последующие моменты времени.

Дискретные аналоги

Дискретный аналог уравнения диффузии (1) получен в рамках метода контрольного объема [23] в сферических координатах с использованием полностью неявной схемы, таким образом, чтобы численное решение уравнения (1) удовлетворяло сеточному уравнению неразрывности:

$$\nabla \cdot \mathbf{B} = 0, \quad (7)$$

где

$$\nabla \cdot \mathbf{B} = \frac{1}{\Delta \mathrm{V}} \sum_{i=1}^{6} \mathbf{B} \cdot \mathbf{n}_i \Delta S_i,$$

$\Delta \mathrm{V}$ – объем контрольного объема, ΔS_i – площадь грани контрольного объема с нормалью $\mathbf{n}_i$. Для этого при аппроксимации оператора *rot* в (1) использован подход, аналогичный предложенному в работе [20].

В результате сеточное уравнение диффузии записывалось в виде:

$$\frac{1}{\Delta t}\left(\mathbf{B}^{(m)} - \mathbf{B}^{(m-1)}\right) = \mathbf{\Lambda}(\mathbf{B}^{(m)}), \tag{8}$$

где $\mathbf{\Lambda}(\mathbf{B})$ – дискретный аналог оператора $\mathbf{B}$; $\mathbf{B}^{(m-1)}$, $\mathbf{B}^{(m)}$ – сеточные значения индукции магнитного поля в моменты времени t_{m-1} и t_m; $\Delta t = t_m - t_{m-1}$ шаг по времени.

При заданном на границе $\mathbf{B}$ потенциал $u(\mathbf{r},t)$ вне шара является регулярным на бесконечности решением внешней задачи Неймана для уравнения Лапласа с граничными условиями:

$$\left(r = a,\ t > 0\right):\ \frac{\partial u}{\partial r} = B_r \tag{9}$$

Используя преобразование обратных радиусов [24]: $r r' = a^2$, $u(r',\theta,\varphi) = r\, w(r,\theta,\varphi)$ ($(0 \le r \le a, r' \ge a$), внешнюю задачу Неймана на потенциал u можно свести к решению внутренней краевой задачи для уравнения Лапласа на w с граничными условиями третьего рода:

$$\left(r = a,\ t > 0\right):\ a\frac{\partial w}{\partial r} + w = -B_r. \tag{10}$$

Сеточное уравнение Лапласа можно записать в виде:

$$L(w) = 0, \tag{11}$$

где $L(w)$ – дискретный аналог оператора Лапласа, полученный методом контрольного объема в сферических координатах.

Алгоритм решения начально-краевой задачи с вакуумными граничными условиями о диффузии магнитного поля в проводящем шаре

Сеточные значения $\mathbf{B}$ в точках на границе шара для удобства записи алгоритма обозначим $\mathbf{f}$, (т.е. $\mathbf{f} = \mathbf{B}_{r=a}$). В данной работе предложен следующий алгоритм получения численного решения задачи (1)–(6).

Пусть в момент времени t_{m-1} известно численное решение во внутренних расчетных точках $\mathbf{B}^{(m-1)}$ и в точках на границе шара $\mathbf{f}^{(m-1)}$ и пусть нижний индекс k нумерует внутренние итерации (и соответствующие приближения), которые необходимо сделать, чтобы получить решение на следующем временном слое. В качестве нулевого приближения для $\mathbf{B}$ и $\mathbf{f}$ на следующем временном слое возьмем $\mathbf{B}^{(m-1)}$ и $\mathbf{f}^{(m-1)}$, т.е. при $k = 0$: $\mathbf{B}_0 = \mathbf{B}^{(m-1)}$, $\mathbf{f}_0 = \mathbf{f}^{(m-1)}$.

Шаг 1: Из сеточного уравнения неразрывности (7) записанного для контрольных объемов, прилегающих к границе расчетной области, $\nabla\cdot\mathbf{B}_k\big|_{\partial G} = 0$ определяется следующее приближение нормальной (радиальной) компоненты магнитного поля на ∂G, т.е. вычисляется $\mathrm{f}_{r,k+1/2}$. Тангенциальные компоненты берутся с предыдущего приближения: $\mathrm{f}_{\theta,\,k+1/2} = \mathrm{f}_{\theta,\,k}$; $\mathrm{f}_{\phi,k+1/2} = \mathrm{f}_{\phi,k}$. Таким образом, вычисляется вектор $\mathrm{f}_{k+1/2}$.

Шаг 2: Для нахождения $\mathrm{B}_{k+1/2}$ решается следующая задача:

$$G:\ \frac{1}{\Delta t}\left(\mathbf{B}_{k+1/2} - \mathbf{B}^{(m-1)}\right) = \mathbf{\Lambda}(\mathbf{B}_{k+1/2}), \qquad \partial G:\ \mathbf{B}_{k+1/2} = \mathbf{f}_{k+1/2}.$$

Шаг 3: Из уравнения неразрывности для приграничных контрольных объемов $\nabla\cdot\mathbf{B}_{k+1/2}\big|_{\partial G} = 0$ определяется следующие приближение $\mathrm{f}_{r,k+1}$ нормальной компоненты индукции на границе расчетной области.

Шаг 4: Потенциал на границе шара $u = aw$, соответствующий $\mathrm{f}_{r,k+1}$ определяется из решения третьей внутренней краевой задачи на уравнение Лапласа:

$$G:\ L(w) = 0;\ \partial G:\ a\frac{w(r+\Delta r) - w(r)}{\Delta r} + w = -\mathrm{f}_{r,k+1}.$$

Шаг 5: Вычисляются следующие приближения для тангенциальных компонент индукции магнитного поля на ∂G:

$$\mathrm{f}_{\theta,\,k+1} = \frac{1}{r}\frac{u(\theta+\Delta\theta) - u(\theta)}{\Delta\theta}\bigg|_{r=a},\quad \mathrm{f}_{\varphi,\,k+1} = \frac{1}{r\sin\theta}\frac{u(\varphi+\Delta\varphi) - u(\varphi)}{\Delta\varphi}\bigg|_{r=a}.$$

Шаг 6: Находится следующее приближение для $\mathbf{B}$ во внутренних расчетных точках шара:

$$G:\ \frac{1}{\Delta t}\left(\mathbf{B}_{k+1} - \mathbf{B}^{(m-1)}\right) = \mathbf{\Lambda}(\mathbf{B}_{k+1}),\ \partial G:\ \mathbf{B}_{k+1} = \mathbf{f}_{k+1}.$$

Шаги 1–6 повторяются до тех пор, пока не будет достигнута сходимость на данном временном слое. В качестве критериев сходимости ставились следующие условия:

$$1)\ \max_G\left|\nabla\cdot\mathbf{B}_{k+1}\right| \le \varepsilon_1, \tag{12}$$

$$2)\ \left\|\mathbf{f}_{k+1} - \mathbf{f}_k\right\|_2 \le \varepsilon_2. \tag{13}$$

При достижении сходимости $\mathbf{B}^{(m)} = \mathbf{B}_{k+1}$, $\mathbf{B}^{(m)}\big|_{r=a} \equiv \mathbf{f}^{(m)} = \mathbf{f}_{k+1}$.

Результаты тестовых расчетов

В качестве теста рассматривалась аксиально-симметричная задача с вакуумными граничными условиями о размагничивании шара единичного радиуса, которая имеет точное решение в элементарных функциях. Если задать начальное магнитное поле $\mathbf{b} = b_r(r,\theta)\mathbf{e}_r + b_\theta(r,\theta)\mathbf{e}_\theta$ и потенциал $u_0(r,\theta)$ в виде:

$$r \le 1:\ b_r = \frac{\cos\theta}{(\pi r)^2}\left(\frac{\sin\pi r}{\pi r} - \cos\pi r\right),$$

$$b_\theta = \frac{\sin\theta}{2(\pi r)^2}\left(\frac{\sin\pi r}{\pi r} - \cos\pi r - \pi r\sin\pi r\right), \tag{14}$$

$$r \ge 1:\ u_0(r,\theta) = -\frac{\cos\theta}{2\pi^2 r^2}, \tag{15}$$

то поле $\mathbf{B}(r,\theta,t) = e^{-\pi^2 t}\,\mathbf{b}(r,\theta)$ $(r \le 1)$ и потенциал $u(r,\theta,t) = e^{-\pi^2 t}\,u_0(r,\theta)$ $(r \ge 1)$ являются решением задачи (1)–(6) с начальными условиями (14)–(15).

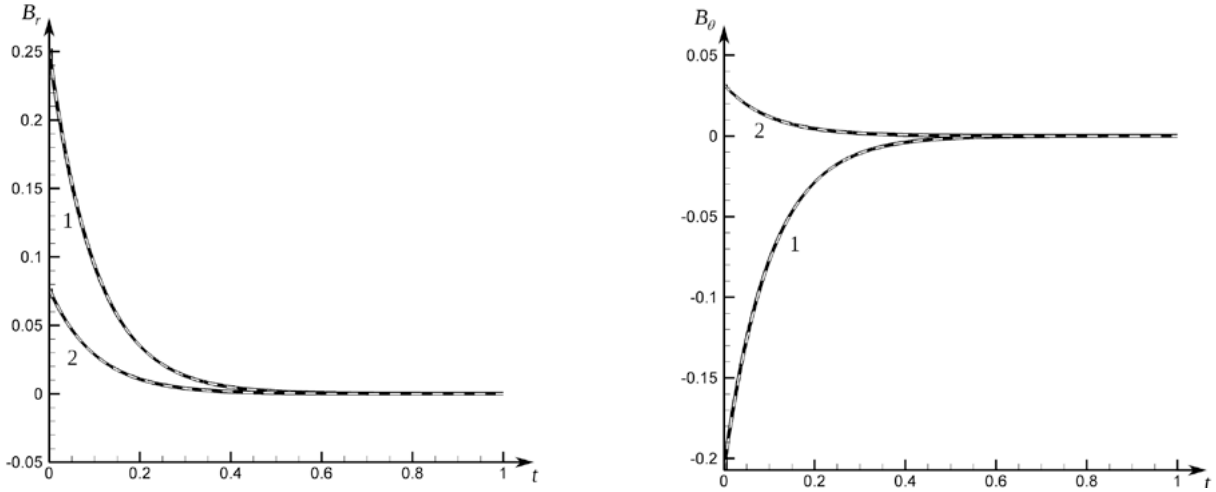
Расчеты проводились до момента времени T ($T = 1$)

на сетках ($n_r \times n_\theta$), где n_r, n_θ, – количество контрольных объемов вдоль соответствующей координатной линии. Во всех представленных расчетах шаг по времени $\Delta t = 10^{-4}$. На каждом временном слое критерии (12), (13) выполнялись с точностью $\varepsilon_1 = \varepsilon_2 = 10^{-11}$. Точность численного решения оценивалась по максимальным абсолютным ΔF и средним относительным δF погрешностям:

$$\Delta F = \max_{\bar{G}\otimes(0,T]} |F - F_{an}|, \qquad \delta F = \frac{1}{V_G}\int_G \frac{|F - F_{an}|}{\bar{F}_{an}}\, d\vartheta \cdot 100\%,$$

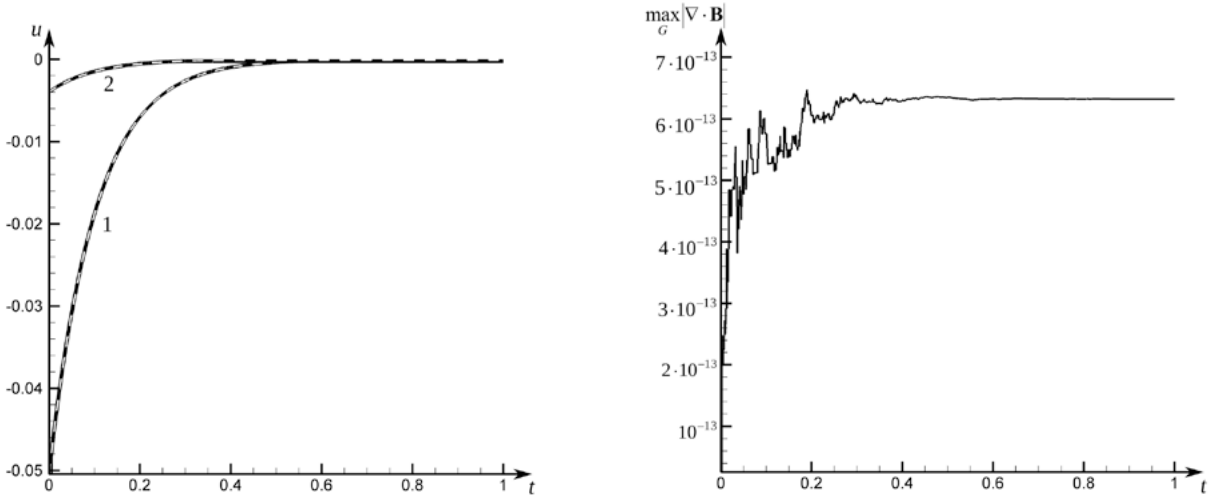
где $\bar{F}_{an} = \frac{1}{V_G}\int_G |F_{an}|\, d\vartheta$, $V_G = \int_G d\vartheta$ – объем области.

Результаты тестов на численную сходимость на последовательности вложенных расчетных сеток, приведенные в таблице 1, демонстрируют, что предложенный алгоритм решения начально-краевых задач с вакуумными граничными условиями занимает промежуточное место по порядку аппроксимации по пространственным переменным между схемами первого и второго порядков. Значения средних относительных погрешностей приведены на момент времени $t = 0.25$. На рис. 1–2 представлены некоторые результаты расчетов на сетке 22×62 в сравнении с аналитическим решением. Разработанный алгоритм может быть обобщен на решение задач магнитной гидродинамики с вакуумными граничными условиями.



а) *б)*

Рис.1. а) $B_r(r,\theta,t)$ в точке 1 с координатами (0.05, 0.71), $B_r(r,\theta,t)$ в точке 2 с координатами (1.00, 0.71); б) $B_\theta(r,\theta,t)$ в точке 1 с координатами (0.08, 0.71), $B_\theta(r,\theta,t)$ в точке 2 с координатами (1.00, 0.71).
Численное решение – черная сплошная линия. Аналитическое решение – белая пунктирная линия



а) *б)*

Рис. 2. а) $u(r,\theta,t)$ в точке 1 с координатами (1.00, 0.08), $u(r,\theta,t)$ в точке 2 с координатами (1.00, 1.49). Численное решение – черная сплошная линия. Аналитическое решение – белая пунктирная линия. б) Максимальное по области G сеточное значение дивергенции магнитного поля $\max_G |\nabla \cdot \mathbf{B}|$

Таблица 1

Максимальные абсолютные и средние относительные погрешности

Сетка $n_r \times n_\theta$	$\lvert\Delta B_r\rvert$	$\lvert\Delta B_\theta\rvert$	$\lvert\Delta u\rvert_{r=a}$	δB_r , %	δB_θ , %	$\delta u\rvert_{r=a}$, %
12×32	$2.12\cdot10^{-3}$	$3.14\cdot10^{-3}$	$8.32\cdot10^{-5}$	0.34	0.62	0.21
22×62	$7.78\cdot10^{-4}$	$1.04\cdot10^{-3}$	$2.00\cdot10^{-5}$	0.17	0.21	0.15
42×122	$2.51\cdot10^{-4}$	$3.19\cdot10^{-4}$	$8.23\cdot10^{-6}$	0.13	0.13	0.13

Список литературы

1. Olson P.L., G.A. Glatzmaier and R.S. Coe, 2011. Complex polarity reversals in a geodynamo model. Earth Planet. Sci. Lett., 304: 168–179.

2. Olson P., P. Driscoll and H. Amit, 2009. Dipole collapse and reversal precursors in a numerical dynamo. Phys. Earth. Planet. Inter., 173: 121–140.

3. Jouve L., Brun, A.S., Arlt, R., Brandenburg, A., Dikpati, M., Bonanno, A., Käpylä, P.J., Moss, D., Rempel, M., Gilman, P., Korpi, M. J. and Kosovichev, A.G., 2008. A solar mean field dynamo benchmark. Astron. Astrophys., 483: 949–960.

4. Glatzmaier G.A., 2018. Computer simulations of Jupiter's deep internal dynamics help interpret what Juno sees. Proc. Nat. Acad. Sci., 115 (27): 6896–6904.

5. Jones C. 2014. A dynamo model of Jupiter's magnetic field. Icarus, 241: 148–159.

6. Jones C.A., 2000. Convection-driven geodynamo models. Phil. Trans. R. Soc. Lond. A., 358: 873–897.

7. Ландау Л.Д., Лифшиц Е.М., 2005. Теоретическая физика: Учеб. пособ.: Для вузов. В 10 т. Т. VIII. Электродинамика сплошных сред. М.: ФИЗМАТЛИТ, 656 с.

8. Кнопфель Г., 1972. Сверхсильные импульсные магнитные поля. Методы генерации и физические эффекты, связанные с созданием импульсных полей мегаэрстедного диапазона / перевод с английского Ф.А. Николаева и Ю.П. Свириденко. М.: Мир, 392 с.

9. Смайт, Вильям Р., 1954. Электростатика и электродинамика / перевод со второго американского изд. А.В. Гапонова и М.А. Миллера. М.: Изд-во иностранной лит., 604 с.

10. Моффат Г.К., 1980. Возбуждение магнитного поля в проводящей среде / перевод с английского А.А. Рузмайкина. М.: Мир, 339 с.

11. Куликовский А.Г., Любимов Г.А., 2011. Магнитная гидродинамика : учебное пособие – 3-е изд. М.: Логос, 324 с.

12. Каулинг Т., 1987. Магнитная гидродинамика / Пер. с англ. В.Г. Петрова. М.: Атомиздат, 144 с.

13. Галанин М.П., Попов Ю.П., 1995. Квазиста-ционарные электромагнитные поля в неоднородных средах: Математическое моделирование. М.: Наука. Физматлит, 320 с.

14. Glatzmaier, G.A., 1984. Numerical simulations of stellar convective dynamos I. The model and method. J. Comp. Phys., 55: 461–484.

15. Решетняк М.Ю., 2013. Моделирование в геодинамо. Saarbrücken: Ламберт Академик Паблишинг, 180 с.

16. Kageyama A., T. Miyagoshi and T. Sato, 2008. Formation of current coils in geodynamo simulations. Nature, 454: 1106–1109.

17. Hejda P. and M. Reshetnyak, 2003. Control volume method for the dynamo problem in the sphere with the free rotating inner core. Studia geoph. et geod, 47: 147–159.

18. Куликовский А.Г., Н.В. Погорелов и А.Ю. Семенов, 2001. Математические вопросы численного решения гиперболических систем уравнений. ФИЗМАТЛИТ, 608 с.

19. Бетелин, В.Б., В.А. Галкин и А.В. Гореликов, 2015. Алгоритм типа предиктор-корректор для численного решения уравнения индукции в задачах магнитной гидродинамики вязкой несжимаемой жидкости. Доклады Академии наук, 464(5): 525–525.

20. Колмычков, В.В., О.С. Мажорова и Е.Э. Федосеев, 2009. Численный метод решения уравнений МГД. Препринты ИПМ имени М.В. Келдыша, 30: 28.

21. 21. Toth G., 2000. The constraint in shock-capturing magnetohydrodynamics codes. J. Comput. Phys., 161: 605.

22. Попов М.В. и Т.Г. Елизарова, 2013. Моделирование трехмерных МГД-течений в рамках магнитной квазигазодинамики. Препринты ИПМ им. М.В. Келдыша, 23: 32.

23. Патанкар С., 1984. Численные методы решения задач теплообмена и динамики жидкости. М.: Энергоатомиздат, 152 с.

24. Тихонов А.Н. и А.А. Самарский, 1972. Уравнения математической физики: Учебник для ун-тов / 4-е изд. испр. М.: Наука, 746 с.

УДК 591.7

М.А. Куприянов

ОЦЕНКА ПОГРЕШНОСТИ ИЗМЕРЕНИЯ ПЕРИМЕТРА ПЛОЩАДНОГО ОБЪЕКТА ПО ЕГО ЦИФРОВОМУ ИЗОБРАЖЕНИЮ

Аннотация

С использованием моделирования дана оценка погрешности измерения периметра объектов произвольной формы по их цифровым изображениям. Установлено, что погрешность измерения периметра по цифровым изображениям составляет не менее 10 %. При увеличении разрешающей способности изображения погрешность измерения периметра стремится к величине около 25–27 %.

Ключевые слова: погрешность измерения периметра, растеризация, цифровые изображения.

Введение

В последние несколько десятилетий в научных работах для исследования различных геометрических характеристик объектов используется косвенный способ измерения, который основан на измерении размера и количества пикселей цифрового изображения, представляющих измеряемый объект. Например, в [1] проводили количественную оценку формы листовых пластин у растений, в [2, 3] измерялись геометрические характеристики озёр по космическим снимкам.

На наш взгляд основным достоинством измерения геометрических характеристик объектов по цифровым изображения является возможность автоматизации процесса измерения, что немаловажно при большом количестве исследуемых объектов. В тоже время недостатком применения такого способа измерения является искажение формы объекта, связанное с дискретной природой цифровых изображений, элементарной частицей которых, является пиксель. Как показано в [4] при увеличении разрешающей способности изображения одного и того же объекта произвольной геометрической формы погрешность измерения его площади будет уменьшаться. Как при этом будет изменяться погрешность измерения периметра объекта – осталось неизученным.

В связи с этим целью настоящей работы является модельное исследование погрешности измерения периметра замкнутого контура по его представлению на цифровом изображении.

Методические вопросы исследования

Для создания наборов модельных данных была использована ранее разработанная методика формирования модельных векторных объектов (далее эталонов). Методика подробно описана в [4]. С использованием указанной методики было сгенерировано 6 выборочных последовательностей эталонов. В каждой выборке содержалось по 1 000 эталонов, площадь каждого них равнялась 10 000 м^2. Эталоны внутри одной выборки имели одинаковую величину показателя формы f, которая определялась по формуле [5; 6, с. 68]:

$$f = \frac{P^2}{4\pi S}, \quad (1)$$

где f – показатель формы эталона, P – периметр границы эталона, S – площадь эталона. Эталоны в выборках имели следующие значения показателя формы f: 1; 1,5; 2; 3; 4; 5.

Важным методическим вопросом исследования является выбор метода растеризации эталонов, т.к. от использованного метода зависит результат измерения. Для проведения исследования в настоящей работе был использован один из наиболее простых методов растеризации «maximum combined area method» подробно описанный в [7]. Основная идея метода заключается в том, что в ячейку растрового изображения записывается значение 1 в том случае, если суммарная площадь контура эталона накрывает по крайней мере 50 % площади растровой ячейки, в противном случае записывается значение 0.

На рисунке 1 приведён пример бинарной растеризации (эталон показан жирной чёрной линией). Как видно на рисунке всего 4 пикселя накрываются эталоном более чем на 50 %, поэтому им присвоены значения 1 (серый цвет), остальным пикселям присвоено значение 0 (белый цвет).

При растеризации устанавливался фиксированный линейный размер ячейки растрового изображения (L). Для растеризации каждого эталона использовались различные значения пространственного разрешения ячейки L так, чтобы расчётное количе-

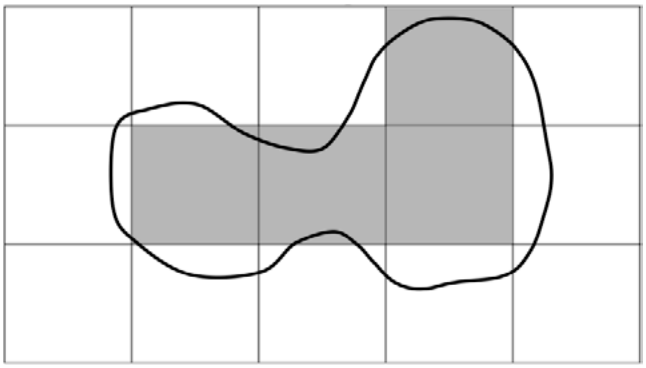


Рис. 1. Пример растеризации эталона

ство пикселей (N_p), приходящееся на эталон, варьировалось в границах от 1 до 100 000. Расчет величины N_p производился по формуле:

$$N_p = \frac{S}{L^2}, \quad (2)$$

где N_p – расчётное количество пикселей в эталоне при фиксированном пространственном разрешении изображения, шт., S – площадь эталона, м^2, L – пространственное разрешение изображения, м.

Для удобства дальнейшего анализа и представления результатов были приняты следующие значения величины N_p: 1, 2, 5, 10, 20, 50, 100, 200, 500, 1 000, 2 000, 5 000, 10 000, 20 000, 50 000, 100 000.

Для измерения периметра и площади эталонов использовалось программное обеспечение ArcGIS Desktop. Измерение площади растеризованных эталонов происходило путём умножения квадрата пространственного разрешения на количество пикселей со значением 1, а периметр рассчитывался как сумма длин всех составляющих внешнюю границу ребер растеризованного эталона.

Анализ и оценка погрешности

С использованием описанной в [4] методики было сформировано 6 выборочных наборов по 1 000 эталонов в каждой. Для каждой выборки были рассчитаны средние значения относительной погрешности измерения периметра растеризованных эталонов по формуле:

$$\delta_{ср} = \frac{1}{n}\sum_{i=1}^{n}\delta_i, \quad (3)$$

где $\delta_{ср}$ – средняя относительная погрешность измерения периметров растеризованных эталонов, n – количество эталонов в выборке, шт., δ_i – относительная погрешность измерения периметра i-го растеризованного эталона, которая рассчитывается по формуле:

$$\delta = \frac{|P - P_и|}{P}, \quad (4)$$

где δ – относительная погрешность измерения периметра растеризованного эталона, P – периметр эталона, м^2, $P_и$ – измеренный периметр растеризованного эталона, м^2.

В результате проведенного анализа были получены значения величины $\delta_{ср}$ для всех выборок с указанными выше значениями показателя формы f и величины N_p. Результаты представлены на рисунке 2. Как видно из рисунка, $\delta_{ср}$ меняется в зависимости от показателя формы и количества пикселей, содержащихся в изображении объекта. При возрастании N величина $\delta_{ср}$ показывает меньшую зависимость от формы измеряемого объекта и стремится к величине приблизительно равной 25–27 %. Практически при любых значениях N_p и f значение величины $\delta_{ср}$ больше 10 %.

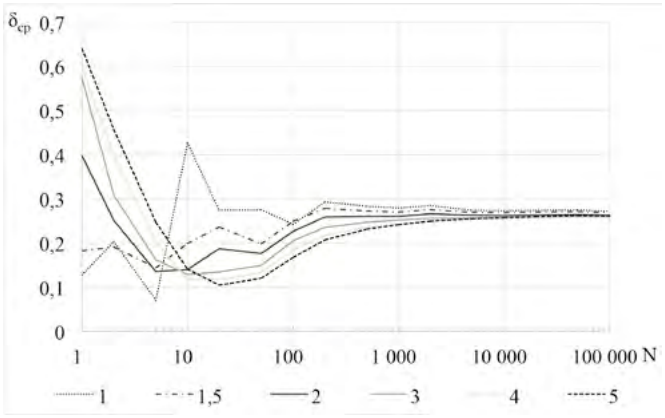


Рис. 2. Зависимость $\delta_{ср}$ от N при различных показателях формы f (обозначены различными линиями и оттенками серого)

Выводы

С использованием методов моделирования выявлена зависимость между относительной погрешностью измерения периметра по цифровому изображению и количеством пикселей из которых состоит измеряемый объект для объектов различной формы. Установлено, что при увеличении разрешающей способности изображения средняя относительная погрешность определения периметра объекта по цифровому изображению стремится к величине около 25–27 %.

Таким образом, для достижения приемлемой погрешности измерения периметра по цифровым изображениям требуется проведение дополнительных процедур, например, сглаживания. Однако сглаживание может привести к погрешностям при измерении площади. Представляет интерес найти наиболее оптимальный метод сглаживания, который бы сохранял как исходную величину периметра измеряемого объекта, так и исходную величину площади.

Список литературы

1. Аникьев А.А., Федоряка Н.И., Аникьева Э.Н. Способ количественной оценки формы листа как уникального сортового признака и его вариабельности у растений земляники //Сельскохозяйственная биология. – 2008. – №. 1. – С. 116.

2. Verpoorter C. et al. A global inventory of lakes based on high□resolution satellite imagery //Geophysical Research Letters. – 2014. – Т. 41. – №. 18. – С. 6396–6402.

3. Verpoorter C., Kutser T., Tranvik L. Automated mapping of water bodies using Landsat multispectral data //Limnology and Oceanography: Methods. – 2012. – Т. 10. – №. 12. – С. 1037–1050.

4. Куприянов М.А., Кочергин Г.А., Полищук Ю.М. Анализ погрешности дистанционного измерения площади пространственных объектов на основе моделирования // Вестник Югорского государственного университета. – 2018. – №. 3 (50).

5. Miller V.C. Quantitative geomorphic study of drainage basin characteristics in the Clinch Mountain area, Virginia and Tennessee // Technical report (Columbia University. Department of Geology); no. 3. – 1953.

6. Берлянт, А.М. Картографический метод исследования [Текст] / А.М. Берлянт. – Москва: Изд-во Моск. ун-та, 1978. – 257 с.

7. How Polygon To Raster works [сайт]: // ESRI: http://desktop.arcgis.com/en/arcmap/10.3/tools/conversion-toolbox/how-polygon-to-raster-works.htm (дата обращения: 08.02.2019).

УДК 519.635.4

Ротко В.В.

О НЕКОТОРЫХ КЛАССАХ ОБРАТНЫХ ЗАДАЧ ДЛЯ СИСТЕМ КОНВЕКЦИИ-ДИФФУЗИИ И ИХ ОБОБЩЕНИЙ

Аннотация

В работе рассматривается вопрос о корректности в пространствах Соболева коэффициентных обратных задач для квазилинейных математических моделей конвекции-диффузии. Неизвестные функции, зависящие от времени, входят в функцию источника и в сам оператор в качестве коэффициентов. В качестве условий переопределения рассматриваются значения решения в некотором наборе внутренних точек области. Приведены условия гарантирующие локальную по времени корректность задачи в классах Соболева. Условия на данные задачи минимальны. Полученные результаты являются точными.

Ключевые слова: параболическая система, обратная задача, функция источника, конвекция-диффузия, тепломассоперенос.

Введение

Мы рассматриваем вопрос о определении вместе с решением правой части специального вида и коэффициентов уравнения в квазилинейных параболических уравнениях и системах. Пусть G – область в R^n с границей Γ класса C^{2m} и $Q=(0,T)\times G$. Параболическое уравнение имеет вид

$$u_t + A(t,x,D)u = g(x,t,u,Du,\dots,D^{2m-1}u) + \sum_{i=1}^{r} b_i(t,x)q_i(t) + f, \qquad (1)$$

где (t,x)∈Q, A – матричный эллиптический оператор порядка $2m$ с матричными коэффициентами размерности $h\times h$, представимый в виде

$$A(t,x,D) = \sum_{i=r+1}^{sh} q_i(t,x')A_i(t,x,D_x) + A_{sh+1}(t,x,D_x),$$

$$A_i = \sum_{|\alpha|\le 2m} a_{i\alpha}(t,x)D^{\alpha}\ (i=r+1,\dots,sh+1), \qquad D=(\partial_{x_1},\partial_{x_2},\dots,\partial_{x_n})$$

и g – некоторая нелинейная функция, зависящая от u и производных $D^{\alpha}u$ с $|\alpha|\le 2m-1$. Под $D^k u$ понимаем вектор, состоящий из всех производных порядка k функции u. Уравнение (1) дополняется начальными и граничными условиями

$$u|_{t=0} = u_0,$$

$$B_j u|_S = \sum_{|\beta|\le m_j} b_{j\beta}(t,x)D^{\beta}u|_S = g_j(t,x), \qquad (2)$$

где $m_j<2m$, $j=1,2,\dots,m$ и $S=(0,T)\times\Gamma$. Неизвестными в (1), (2) являются решение u, функции $q_i(t)$ ($i=1,2,\dots,hs$, $sh\ge r$), входящие как в правую часть (1) так и в оператор A как коэффициенты. Условия переопределения для нахождения этих функций q_i имеют вид

$$u|_{x=x_i}=\psi_i(t),\ i=1,2,\dots,s, \qquad (3)$$

где $\{x_i\}$ – набор точек из G. Проблемы подобного вида возникают при описании процессов тепломассопереноса, диффузионных процессов, процессов фильтрации и во многих других областях [15, 16]. Прежде всего мы сошлемся на работу [4], где получена теорема существования и единственности решений задачи (1)–(3) в пространствах Гельдера в случае $m = h = r = 1$, $s = 1$ и линейной по своим аргументам функции f. В случае $n = 1$, $r = 1$ и $G = R$ аналогичный результат в получен в работах [9], [2]. Общие теоремы о разрешимости абстрактных задач такого вида в квазилинейном случае получены в монографии [17] в пространствах функций, удовлетворяющих условию Гельдера по t в случае когда главная часть оператора A не зависит от неизвестных функций и $D(A)$ не зависит от времени. Результаты применимы и к задачам вида (1)–(3) и при выполнении некоторых (довольно жестких и вообще говоря значительно завышенных) условий на данные задачи гарантируют локальную по времени разрешимость этой задачи. Линейные задачи вида (1)–(3) в случае $m = 1$, были рассмотрены в работах авторов в [6, 18] (здесь условия на данные минимальны), а квазилинейные в работе [7], где были ослаблены условия на данные по сравнению с теми, которые были использованы в [17]. Стоит отметить, что глобальных по времени результатов о разрешимости задач вида (1)–(3) для нелинейной правой части f, по видимому, не имеется за исключением работы [5].

Отметим, что численные методы решения различных модельных задач, входящих в класс (1)–(3) рассматривались, например, в книгах [16, 1] и большом количестве работ (см., например, [11, 12]), выделим работу [14], где как раз рассматривался случай квазилинейной параболической системы.

В данной работе мы получим локальную по времени теорему о существовании и единственности решений задачи (1)–(3). Схема доказательства может быть использована и для построения численных алгоритмов решения подобных задач (что собственно говоря подтверждается численными экспериментами).

Опишем содержание работы. В первом параграфе описаны условия на данные задачи и сформулированы основные результаты. Во втором параграфе приведено их доказательство. Обозначения функциональных пространств стандартные (см., например, [8]).

Определения, обозначения, и формулировка основных результатов

Через $B(X,Y)$ обозначим пространство линейных ограниченных операторов, определенных на X со значениями в Y. Если $X=Y$ пишем $B(X)$ вместо $B(X,X)$. Как обычно символы ρ(L), σ(L) обозначают спектр и

резольвентное множество оператора L. Пусть E – банахово пространство. Через $L_p(G;E)$ (G – область в R^n) обозначается пространство сильно измеримых функций, определенных на G со значениями в E и конечной нормой $\|\|u(x)\|_E\|_{L_p(G)}$ [8]. Мы также используем пространства $C^k(\bar{G};E)$, состоящие из функций, имеющих в G все производные до порядка k включительно, непрерывные в G и допускающие непрерывное продолжение на замыкание $\bar{G}$. Обозначения для пространств Соболева $W_p^s(G;E)$, $W_p^s(Q;E)$ и т.д. стандартные [8, 13, 10]. При нецелых s пространство Соболева $W_p^s(G;E)$ совпадает с пространством Бесова $B_{p,p}^s(G;E)$. Если $E = C$ или $E = C^n$, то последнее пространство обозначаем просто через $B_{p,p}^s(G)$. Аналогично вместо $W_p^s(G;E)$ или $C^k(\bar{G};E)$ используем обозначение $W_p^s(G)$ или $C^k(\bar{G})$. Таким образом, включение $u\in W_p^s(G)$ (или $u\in C^k(\bar{G})$) для данной вектор-функции $u=(u_1,u_2,\dots,u_k)$ означает, что каждая из компонент u_i принадлежит пространству $W_p^s(G)$ (или $C^k(\bar{G})$). В этом случае под нормой вектора понимаем сумму норм координат. Будем считать, что аналогичное соглашение справедливо и для матриц, т.е. включение $a\in W_p^s(G)$ для данной матрицы-функции $a=\{a_{ij}\}_{j,i=1}^k$ означает, что $a_{ij}(x)\in W_p^s(G)$ для всех ij. Для данного интервала $J=(0,T)$, положим $W_p^{s,r}(Q)=W_p^s(J;L_p(G))\cap L_p(J;W_p^r(G))$. Соответственно, $W_p^{s,r}(S)=W_p^s(J;L_p(G))\cap L_p(J;W_p^r(G))$.

Определение вложения $\Gamma\in C^{2m}$ может быть найдено в [3]. Далее, мы считаем, что параметр $p>n+2m$ зафиксирован. Пусть $B_\delta(x_i)$ – шар радиуса δ с центром в точке x_i. Дан набор точек $\{x_j\}$ из (3), параметр $\delta>0$ назовем допустимым, если $\overline{B_\delta(x_i)}\subset G$, $\overline{B_\delta(x_i)}\cap\overline{B_\delta(x_j)}=\emptyset$ для $i\neq j$, $i,j=1,2,\dots,r$. Пусть $Q^\tau=(0,\tau)\times G$, $G^\delta=\cup_i B_\delta(x_i)$, $Q_\delta=(0,T)\times G_\delta$, $Q_\delta^\tau=(0,\tau)\times G_\delta$.

Далее во всех условиях мы считаем, что допустимый параметр $\delta>0$ зафиксирован и не будем это оговаривать. Зафиксируем также некоторую функцию $\varphi\in C_0^\infty(R^n)$ такую, что $supp\,\varphi\subset G_\delta$ и $\varphi(x)=1$ для $x\in G_{\delta/2}$.

Условия согласования и гладкости данных могут быть записаны в виде:

$$u_0(x)\in B_{p,p}^{2m-2m/p}(G),\ g_j\in W_p^{k_j,2mk_j}(S),\quad k_j=(2m-m_j-1/p)/(2m), \tag{4}$$

$$\varphi(x)\nabla u0(x)\in B_{p,p}^{2m-2m/p}(G_\delta) \tag{5}$$

$$\psi_j(t)\in C^1([0,T]),\ \psi_j(0)=u_0(x_j),\ j=1,2,\dots,s. \tag{6}$$

Условия на коэффициенты операторов A,B_j более или менее стандартные. Более того, для простоты выкладок мы будем использовать не самые точные условия на коэффициенты. Мы считаем, что

$$\begin{gathered}(a_{i\alpha}(t,x)\in L_\infty(Q)\ \ (|\alpha|<2m),\\ a_{i\alpha}\in C(\bar{Q})\ (|\alpha|=2m, i=r+1,\dots,sh+1),\\ b_{j\beta}\in C^{1-m_j/2m,2m-m_j}(\bar{S})\ (j=1,\dots,m,|\beta|\le m_j),\end{gathered} \tag{7}$$

$$b_j(t,x)\in L_p(Q),\nabla_x b_j(t,x)\in L_p(Q_\delta)\ (j=1,2,\dots,r) \tag{8}$$

$$\begin{gathered}\nabla_x a_{i\alpha}(t,x)\in L_\infty(Q_\delta)\ \ (|\alpha|<2m),\\ \nabla_x a_{i\alpha}\in L_\infty(Q_\delta)\ (|\alpha|=2m, i=r+1,\dots,sh+1),\end{gathered} \tag{9}$$

Мы будем искать функции q_i в классе непрерывных функций. В связи с этим потребуем также, чтобы

$$\begin{gathered}(a_{i\alpha}(t,x_j),b_l(t,x_j)\in C([0,T]),g(x_j,t,p)\\ \in C([0,T]\times B_R)\ \forall R>0\end{gathered} \tag{10}$$

при все $l=1,\dots,r$, $j=1,2,\dots,s$ и $|\alpha|<2m$.

Рассмотрим матрицу $B(t)$ размера $sh\times sh$, строчки которой с номерами с номерами от $(j-1)h+1$ до jh занимают вектора-столбцы

$$\begin{gathered}(-b_1(t,x),-b_2(t,x),\dots,b_r(t,x),A_{r+1}\\ u_0(x),\dots,A_{sh}\,u_0(x))|_{x=x_j}.\end{gathered} \tag{11}$$

Можно показать, используя условия (4), (5), (7), (11) и теоремы вложения (см. лемму 1 ниже), что элементы этой матрицы непрерывны на $0,T]$. Потребуем, чтобы существовала постоянная $\delta_0>0$ такая, что

$$|\det B(t)|\ge\delta_0\ \ \forall t\in[0,T]. \tag{12}$$

Рассмотрим систему уравнений

$$B(t)\vec{q}^0=\vec{g},\vec{q}^0=(q_1^0,q_2^0,\dots,q_{sh}^0),$$

где $\vec{q}$ – вектор-столбец, координаты которого с номерами от $(j-1)h+1$ до jh есть вектор $f(t,x_j)-A_{(sh+1)}\,u_0(x_j)-\psi_{jt}(t)+g(t,x_j,u_0\ (x_j),\dots,D^{2m-1}\,u_0(x_j))$. При выполнении условия (12) система (13) имеет единственное решение $\vec{q}^0=(q_1^0,\dots,q_{sh}^0)=(B(t))^{-1}\vec{q}(t)$. Приведенные выше условия на данные задачи гарантируют, что $\vec{q}^0\in C([0,T])$.

Рассмотрим операторы

$$A_0(t,x,D)=\sum_{i=r+1}^{sh}q_i^0(t,x')A_i(t,x,D_x)+A_{hs+1}(t,x,D_x),$$

$$A_0^0(t,x,D)=\sum_{i=r+1}^{sh}q_i^0(t,x')A_{i0}(t,x,D_x)+A_{sh+10}(t,x,D_x),$$

где $A_{i0}(t,x,\xi)=\sum_{|\alpha|=2m}a_{i\alpha}(t,x)\xi^\alpha$, и предположим что оператор ∂_t+A_0 параболичен, т.е. найдется постоянная $\delta_1>0$ такая, что любой корень p многочлена

$$\det(A_0^0(t,x,i\xi)+pE)=0$$

(E – единичная матрица) удовлетворяет неравенству

$$\mathrm{Re}\,p\le-\delta_1|\xi|^{2m},\forall\xi\in R^n,\forall(x,t)\in Q. \tag{13}$$

Условие Лопатинского запишется в виде: для любой точки $(t_0,x_0)\in S$ запишем операторы A_0^0,B_{j0} ($B_{j0}=\sum_{|\beta|=mj}b_{j\beta}\,D^\beta$) в локальной системе координат y (ось y_n направлена по нормали к S, а остальные оси лежат в касательной плоскости) и предположим, что система

$$(\lambda E+A_0^0(i\xi',\partial_{y_n}))v(z)=0\ \ B_{j0}(i\xi',\partial_{y_n}))v(0)=h_j,) \tag{14}$$

($\xi'=(\xi_1,\dots,\xi_{n-1},y_n\in R^+,j=1,2,\dots,m$) имеет единственное решение из $C(\bar{R}^+;E)$ убывающее на бесконечности для всех $\xi'\in R^{n-1}$, $|\arg\lambda|\le\pi/2$ и $h_j\in E$ таких что $|\xi'|+|\lambda|\neq0$.

Алгебраические условия, гарантирующие выполнение (15) могут быть найдены, например, в [3]. Справедлива следующая теорема (первая половина утверждения вытекает из теоремы 10.4 в [3], а доказательство второй половины по существу содержится в доказательстве теоремы 4 в [6]).

Теорема 1. *Пусть G – ограниченная область с границей класса C^{2m}, условия* (4), (7), (14), (15) *выполнены и $k_j\neq1/p$ для всех $j=1,2,\dots,m$. Тогда если $g\in L_p(Q)$, то существует единственное решение $u\in W_p^{1,2m}(Q)$ задачи*

$$\begin{gathered}u_t+A_0(t,x,D_x)u=g,u|_{t=0}=u_0(x),\\ B_j\,u|_S=g_j(j=1,\dots,m),\end{gathered} \tag{15}$$

удовлетворяющее оценке

$$\|u\|_{W_p^{1,2m}(Q)} \le c\left[\|g\|_{L_p(Q)} + \sum_{j=1}^{m} \|g_j\|_{W_p^{k_j,2mk_j}(S)} + \|u_0\|_{B_{pp}^{2m-2m/p}(G)}\right],$$

где c – постоянная, не зависящая от данных задачи g,g_j,u_0 и решения u. Пусть дополнительно выполнены условия (5), (8), (9) *и* $\varphi g \in L_p(0,T;W_p^1(G))$. *Тогда решение задачи* (15) *обладает свойством* $\varphi u \in L_p(0,T;W_p^{2m+1}(G))$, $\varphi u_i \in L_p(0,T;W_p^1(G))$, $\varphi\nabla u \in W_p^{1,2m}(Q)$. *Если* $u_0=0, g_j=0$ *для всех j, то найдется постоянная c>0 такая, что решение удовлетворяет оценке*

$$\|u\|_{W_p^{1,2m}(Q^\tau)} \le c\|g\|_{L_p(Q^\tau)},$$

где постоянная c не зависит от $\tau\in(0,T]$. *Соответственно, при дополнительном условии найдется постоянная c>0, не зависящая от* $\tau\in(0,T]$ *и такая, что*

$$\|\varphi\nabla u\|_{W_p^{1,2m}(Q^\tau)} \le c(\|g\|_{L_p(Q^\tau)} + \|\varphi g\|_{L_p(0,\tau;W_p^1(G))}).$$

Опишем свойства нелинейной вектор-функции $g(t,x,u,Du,\dots,D^{2m-1}u)$. Обозначим через p вектор $p=(p_1,p_2,\dots,p_{r_0})$, где r_0 – число h умноженное на количество различных производных вида $D^\alpha u$ с $0 \le |\alpha| \le 2m-1$. Вектор p отвечает аргументам $u,Du,\dots,D^{2m-1}u$ функции g. Пусть B_R – шар радиуса R с центром в нуле в R^{r_0}.

Условие (А). Функция $g(t,x,p)$ непрерывна по параметру $p\in R^{r_0}$ при п.в. $(t,x)\in Q$ и измерима по $(t,x)\in Q$ для каждого p; для каждого $R>0$ найдется функция такая, что $|g(t,x,p^1)-g(t,x,p^2)| \le |\Phi_0(t,x)||p_1-p_2|$,
для всех $p^1,p^2 \in B_R$; найдется допустимое $\delta>0$ такое, что для каждого $R>0$ $g\in L_p(0,T;W_p^1(G;C(\overline{B}_R)))$; при почти всех $(t,x)\in Q$ g непрерывно дифференцируема по p и производные g_{x_j}, g_{p_i} ($j=1,2,\dots,n$, $i=1,2,\dots,r_0$) удовлетворяют условию Липшица в следующем смысле: для каждого $R>0$ найдутся функции $\Phi_i(t,x) \in L_p(Q)$ такие, что

$$|g_{p_i}(t,x,p^1) - g_{p_i}(t,x,p^2)| \le |\Phi_1(t,x)||p^1 - p^2|, \ j=1,2,\dots,r_0,$$
$$|g_{x_j}(t,x,p^1) - g_{x_j}(t,x,p^2)| \le |\Phi_2(t,x)||p^1 - p^2|, \ i=1,2,\dots,r_0,$$

для всех $p^1,p^2 \in B_R$ и п.в. $(t,x)\in Q_\delta$.

Теорема 2. *Пусть условия* (А), (4)–(14) *выполнены,* $k_j \ne 1/p$ *для всех j,* $f \in L_p(Q)$, $\varphi f \in L_p(0,T;W_p^1(G))$. *Тогда найдется число* $\tau_0\in(0,T]$ *такое, что существует единственное решение* $(u,q_1,\dots,q_{sh})$ *задачи* (1)–(3) *из класса*

$u\in W_p^{1,2m}(Q^{\tau_0})$: $\varphi\nabla u\in W_p^{1,2m}(Q^{\tau_0})$, $q_j\in C([0,T])$, $j=1,2,\dots,sh$.

Доказательство основных результатов

Нам понадобится одно вспомогательное утверждение, вытекающее из теорем вложения, доказательство которого имеется, например, в работе [6].

Лемма 1. *Если* $u\in W_p^{1,2m}(Q^\tau)$ ($\tau>0, p>n+2m$), *то производная вида* $D_x^\alpha u$ *при* $|\alpha|\le 2m-1$ *быть может после изменения на множестве меры нуль принадлежит* $C(\overline{Q}^\tau)$ *и если* $u(0,x)=0$, *то для всех* α *с* $|\alpha|<2m$ *справедлива оценка*

$$\|D^\alpha u\|_{C(Q^\tau)} \le c\|u\|_{W_p^{1,2m}(Q^\tau)}\tau^\beta,$$

где β, *c – некоторые положительные постоянные, не зависящие от u.*

Доказательство теоремы 2. Пусть u – решение задачи (1)–(3). Рассмотрим вспомогательную задачу

$$\Phi_t + A_0(t,x,D)\Phi = g_0,$$
$$\Phi(0,x) = u_0(x), B_i\,\Phi|_S = g_i (j=1,2,\dots,m), \qquad (16)$$

где $g_0 = f + g(x,t,u_0,Du_0,\dots,D^{2m-1}u_0) + \sum_{i=1}^{r} b_i(t,x)q_i^0(t)$.

(Сделав замену переменных $u=v+\Phi$, мы придем к задаче

$$v_t + A_0 v + A(\vec{\mu})v = \tilde{g} - A(\vec{\mu})\Phi + \sum_{i=1}^{r} b_i(t,x)\mu_i(t), \qquad (17)$$

$$v|_{t=0}=0,\ B_j v|_S = 0 \ (j=1,2,\dots,m). \qquad (18)$$

$$v(t,x_j) = \psi_j - \Phi(t,x_j) = \psi_j,\ i=1,2,\dots,s, \qquad (19)$$

где $\tilde{g}=g(t,x,v+\Phi,D(v+\Phi),\dots,D^{2m-1}(v+\Phi)) - g(x,t,u_0,Du_0,\dots,D^{2m-1}u_0)$, $\mu_i(t)=q_i(t)-q_i^0(t)$, $A(\vec{\mu})=\sum_{i=r+1}^{sh}\mu_i(t)A_i$, $\vec{\mu}=(\mu_1,\dots,\mu_{sh})$. Таким образом, свели задачу (1)–(3) к эквивалентной и более простой задаче (17)–(19), которую мы и будем исследовать. Фиксируя функции $\mu_j\in C([0,\tau])$ и находя решение v задачи (18)–(19) на интервале $(0,\tau)$, мы получим отображение $v=v(\vec{\mu})$. Изучим его свойства. Положим $\|\vec{\mu}\|_{C([0,\tau])}=\sum_{i=1}^{sh}\|\mu_i\|_{C([0,\tau])}$. Используя теорему 1, сведем задачу о нахождении функции $v=v(\vec{\mu})$ к исследованию уравнения

$$v = (\partial_t + A_0)^{-1}\tilde{g} - (\partial_t - A_0)^{-1}A(\vec{\mu})(v+\Phi) + (\partial_t + A_0)^{-1}\sum_{i=1}^{r} b_i(t,x)\mu_i(t). \qquad (20)$$

Из оценки теоремы 1 для решения вытекает и оценка для нормы оператора $(\partial_t+A_0)^{-1}$ на произвольном интервале $(0,\tau)$ ($\tau \le T$). Определим пространство H_τ функций $v\in W_p^{1,2m}(Q^\tau)$ таких, что $\varphi\nabla v\in W_p^{1,2m}(Q^\tau)$ и удовлетворяет однородным начальным и граничным условиям. Введем норму в нем норму

$$\|v\|_\tau = \|v\|_{W_p^{1,2m}(Q^\tau)} + \|\varphi\nabla v\|_{W_p^{1,2m}(Q^\tau)}.$$

Далее, из теоремы 1 имеем

$$\|(\partial_t+A_0)^{-1}A(\vec{\mu})v\|_\tau \le c\|A(\vec{\mu})v\|_{L_p(Q^\tau)} + \|\varphi A(\vec{\mu})v\|_{L_p(0,\tau;W_p^1(G))} \le cc_1\|v\|_\tau\|\vec{\mu}\|_{C([0,\tau])},$$

где постоянная c_1 зависит от норм коэффициентов операторов A_i в $L_\infty(Q)$ и не зависит от τ. Выберем шар, в котором мы будем искать решение $\vec{\mu}$. Далее считаем, что

$$\|\vec{\mu}\|_{C([0,\tau])} \le \frac{1}{4cc_1} = r_1, \qquad (21)$$

т.е. $\mu \in B_{r_1}(\tau)$ – шар в пространстве $C([0,\tau])$ радиуса r_1. При выполнении этого условия норма оператора $(\partial_t+A)^{-1}A(\vec{\mu})v$ в $B(H_\tau)$ оценивается через 1/4. Оценим оставшиеся слагаемые в (20). Обозначим правую часть (20) через $R(v)$. Аналогично имеем, что

$$\|(\partial_t - A_0)^{-1}A(\vec{\mu})\Phi + (\partial_t + A_0)^{-1}\sum_{i=1}^{r} b_i(t,x)\mu_i(t)\|_\tau \le c_2 r_1\|\Phi\|_T + c_3 r_1, \qquad (22)$$

где постоянные c_i не зависят от τ. Положим $R_0=4(c_2r_1\|\Phi\|_T + c_3r_1)$. Ищем решение уравнение (20) в шаре $U_{R_0}=\{v\in H_\tau : \|v\|_\tau \le R_0\}$. Наконец, используя условие (А), легко получить оценку

$$\|(\partial_t+A_0)^{-1}\tilde{g}\|_\tau \le c(R_0)(\|v\|_{C([0,\tau];C^{2m-1}(\overline{G}))} + \|\Phi - u_0\|_{C([0,\tau];C^{2m-1}(\overline{G}))}). \qquad (23)$$

В силу леммы 1 правая часть оценивается через

$$R_2 = c_1(R_0)(\tau^\beta\|v\|_{W_p^{1,2m}(Q^\tau)} + \|\Phi - u_0\|_{C([0,\tau];C^{2m-1}(\overline{G}))}).$$

Выберем τ_0 такое, что $R_2 \le R_0/2$ при $\tau \le \tau_0$. При таком выборе, в силу (22), (23) имеем, что $\|R(v)\|_\tau \le R_0$ при всех $v \in U_{R_0}$, т.е. оператор R переводит шар U_{R_0} в себя. Покажем сжимаемость оператора. Совершенно аналогично имеем оценку (при $\tau \le \tau_0$)

Выбирая $\tau_1 \le \tau_0$ так чтобы $c_4(R_0)\tau^\beta \le 1/4$, мы получим, что оператор R, сжимающий и значит уравнение (20) имеет единственное решение в шаре U_{R_0} из H_τ при $\tau \le \tau_1$. Получим некоторые дополнительные оценки. Пусть v_1, v_2 два решения уравнения (20), отвечающие двум различным векторам $\vec{\mu}^1, \vec{\mu}^2 \in B_{r_1}(\tau)$. Вычитая равенства (20) и используя условие (А) и лемму 1, получим оценку

$$\|v_1 - v_2\|_\tau = \|R_1(v_1) - R_2(v_2)\|_\tau$$
$$\le (c\tau^\beta + 1/4)\|v_1 - v_2\| + c(R_0)\|\vec{\mu}^1 - \vec{\mu}^2\|_{C([0,\tau])},$$

откуда, взяв $\tau_2 \le \tau_1$ достаточно малым, получим оценку

$$\|v_1 - v_2\|_\tau \le 2c(R_0)\|\vec{\mu}^1 - \vec{\mu}^2\|_{C([0,\tau])}, \quad (24)$$

при $\tau \le \tau_2$. Перейдем к построению решения обратной задачи. Пусть v – решение задачи (17)–(19) из класса H_τ. Для сокращения записи, мы ниже пишем в качестве аргумента у функции $\tilde{g}$ только три величины (t,x,v). Из определения этого класса вытекает в частности, что определены следы в точке $x = x_j$ всех функций входящих в уравнение (17) и мы имеем

$$\tilde{\psi}_{jt} + A_0 v(t,x_j) + A(\vec{\mu})v(t,x_j) = \tilde{g}(t,x_j,v(t,x_j)) - A(\vec{\mu})\Phi(t,x_j) + \sum_{i=1}^{r} b_i(t,x_j)\mu_i(t), j = 1,2,\dots,s. \quad (25)$$

Эти равенства можно записать в виде операторного уравнения

$$\tilde{B}\vec{\mu} = F + S(\vec{\mu}), \quad (26)$$

где матрица $\tilde{B}(t)$ имеет размер $sh \times sh$ и ее строчки с номерами с номерами от $(j-1)h+1$ до jh занимают вектора-столбцы

$$(-b_1(t,x), -b_2(t,x), \dots, -b_r(t,x),$$
$$A_{r+1}\Phi(t,x), \dots, A_{sh}\Phi(t,x))|_{x=x_j}.$$

Имеем $\tilde{B}(0) = \tilde{B}(0)$ и значит найдется $\tau_3 \le \tau_2$ такое, что $|\det \tilde{B}(t)| > \delta_1 > 0$ на $[0,\tau_3]$, где δ_1 – некоторая положительная постоянная. В правую часть (26) входит вектор-функция $S(\vec{\mu})$, координаты которой с номерами от $(j-1)h+1$ до jh занимают вектора-столбцы – $A_0 v(t,x_j) - A(\vec{\mu})v(t,x_j) + \tilde{g}(t,x_j,v(t,x_j)) - \tilde{g}(t,x_j,0)$, где функция $v = v(\vec{\mu})$) – решение задачи (17), (18). По доказанному, этот оператор $\vec{\mu} \to v(\vec{\mu})$ определен для всех векторов $\vec{\mu} \in B_{r_1}(\tau)$ при $\tau \le \tau_2$. Свойства этого отображения $\vec{\mu} \to v(\vec{\mu})$ мы уже исследовали. У вектор-функция F координаты с номерами от $(j-1)h+1$ до jh занимают вектора-столбцы $F_j = -\tilde{\psi}_{jt}$. Система (26) и есть искомая система для нахождения величин μ_i. Покажем, что можно найти такое $\tau_4 \le \tau_3$, что оператор $B^{-1}F + B^{-1}S(\vec{\mu}) : C([0,\tau_4]) \to C([0,\tau_4])$, определен, переводит шар $B_{r_1}(\tau_4)$ в пространстве $C([0,\tau_4])$ в себя и является в нем сжимающим. По построению величин q_j^0 $F_j(0) = 0$. Тем же свойством обладает функция $S(0)$. Следовательно, найдется $\tau^1 \le \tau_3$ такое, что $\|F - S(0)\|_{C([0,\tau])} \le r_1/2$ при всех $\tau \le \tau^1$. Используя лемму 1, получим $(\tau \le \tau^1)$

$$\|B^{-1}S(\vec{\mu}_1)(t) - B^{-1}S(\vec{\mu}_2)(t)\|_{C([0,\tau])} \le c_1 \sum_{i=1}^{s} (\|A_0(v_1 - v_2)(t,x_j)\|_{C([0,\tau])} + \|A(\mu_1 - \mu_2)v_1(t,x_j)\|_{C([0,\tau])} + \|A(\mu_2)(v_1 - v_2)(t,x_j)\|_{C([0,\tau])} + \|A_0(v_1 - v_2)(t,x_j)\|_{C([0,\tau])} + \|\tilde{g}(t,x_j,v_1(t,x_j)) - \tilde{g}(t,x_j,v_2(t,x_j))\|_{C([0,\tau])}.$$

Тогда используя лемму 1 и оценку (24) получим неравенство

$$\|B^{-1}S(\vec{\mu}_1)(t) - B^{-1}S(\vec{\mu}_2)(t)\|_{C([0,\tau_1])} \le c_6(R_0)\tau^\beta \|\mu^q - \mu^2\|_{C([0,\tau])}, \quad (27)$$

где β – некоторая положительная постоянная. Выберем $\tau^2 \le \tau^1$ такое, что $c_6(R_0)(\tau^2)^\beta \le 1/2$. Положим $\tau_4 = \tau^2$. Тогда из (27) вытекает, что оператор $B^{-1}(F + S(\vec{\mu}))$, определен, переводит шар $B_{r_1}(\tau_4)$ в себя и является в нем сжимающим. Применяя теорему о неподвижной точке, получим, что в шаре $B_{r_1}(\tau_4)$ существует единственное решение системы (26).

Положим $v = v(\vec{\mu})$. Покажем, что построенная функция удовлетворяет условиям (19). По построению v – решение задачи (17)–(18). Взяв $x = x_j$ в (17), получим

$$v_t(t,x_j) + A_0 v(t,x_j) + A(\vec{\mu})v(t,x_j) = \tilde{g}(t,x_j,v(t,x_j)) - A(\vec{\mu})\Phi(t,x_j) + \sum_{i=1}^{r} b_i(t,x_j)\mu_i(t), j = 1,2,\dots,s.$$

Вычитая это равенство из (25) получим $(v(t,x_j) - \tilde{\psi}_j)_t = 0$. для всех j. Откуда получим требуемое.

Выводы

Способ доказательства достаточно конструктивен и позволяет на его основе строить соответствующий численный алгоритм решения. Численные эксперименты в случае систем второго порядка были проведены и показали достаточно хорошую сходимость алгоритма.

Благодарности

Работа поддержана грантом РФФИ 18-01-00620а.

Список литературы

1. Алифанов О.М., Артюхов Е.А., Ненароком А.В. Обратные задачи сложного теплообмена. М: Янус-К, 2009.

2. Белов, Ю.Я., Коршун К.В. О задаче идентификации функции источника для уравнения типа Бюргерса // J. of Siberian Federal University. Mathematics & Physics. 2012. Т. 5(4), P. 497–506.

3. Ладыженская О.А. Солонников В.А., Уральцева Н.Н. Линейные и квазилинейные уравнения параболического типа. М.: Наука, 1967.

4. Прилепко А.И., Соловьев В.В. Теоремы разрешимости и метод Ротэ в обратных задачах для уравнения параболического типа // Дифференц. уравнения. 1987. Т. 23, 10. С. 1791–1799.

5. Соловьев В.В. Существование решения в «целом» обратной задачи определения источника в квазилинейном уравнении параболического типа, Дифференц. уравнения, 1996, том 32, номер 4, 536-544.

6. Пятков С.Г., Самков М.Л. О некоторых классах коэффициентных обратных задач для параболических систем уравнений // Матем. тр. 2012. Т. 15, 1. С. 155–177.

7. Пятков С.Г., Ротко В.В. Об определении функции источника в квазилинейных параболических задачах с точечными условиями переопределения // Вестник ЮУрГУ. Серия «Математика. Механика. Физика». 2017. Т. 9, 4. С. 19-26.

8. Трибель Х. Теория интерполяции. Функциональные пространства. Дифференциальные операторы. М.: Мир, 1980.

9. Afinogenova O.A., Belov Yu. Ya., Frolenkov I.V. Stabilization of the solution to the identification problem of the source function for a one-dimensional parabolic equation // Doklady Mathematics. 2009. V. 79, N 1. P. 70–72.

10. Amann H. Operator-valued Fourier multipliers, vector-valued Besov spaces, and applications // Math. Nachr. 1997. V. 186, N 1, P. 5–56.

11. Badia A. El, Hamdi A. Inverse source problem in an advection-dispersion- reaction system: application to water pollution // Inverse Problems. 2007. V. 23. P. 2103–2120.

12. Badia A. El, Ha-Duong T. Inverse source problem for the heat equation. Application to a pollution detection problem. // J. Inv. Ill-Posed Problems. 2002. V. 10, N 6. P. 585–599.

13. Denk R., Hieber M., Prüss J. Optimal -estimates for parabolic boundary value problems with inhomogeneous data // Math. Z. 2007. V. 257, N 1. P. 193-224.

14. Mamonov A.V., Tsai Y-H. R. Point source identification in nonlinear advection-diffusion-reaction systems // Inverse Problems. 2013. V. 29, N 3. 26 p.

15. Marchuk G.I. Mathematical Models in Environmental Problems. Studies in Mathematics and its Applications. V. 16. Amsterdam: Elsevier Science Publishers, 1986.

16. Ozisik M.N., Orlande H. R. B. Inverse Heat Transfer. New York: Taylor & Francis, 2000.

17. Prilepko A.I., Orlovsky D.G., Vasin I. A. Methods for solving inverse problems in Mathematical Physics. New-York: Marcel Dekker, Inc., 1999.

18. Pyatkov S.G., Rotko V.V. On some parabolic inverse problems with the pointwise overdetermination // AIP Conference Proceedings. 2017. V. 1907, 020008.

УДК 51-73

Е.И. Сафонов, С.Н. Шергин

ЧИСЛЕННОЕ РЕШЕНИЕ ОБРАТНЫХ ЗАДАЧ ДЛЯ УРАВНЕНИЯ КВАЗИСТАЦИОНАРНЫХ ЭЛЕКТРОМАГНИТНЫХ ВОЛН В АНИЗОТРОПНЫХ НЕМЕТАЛЛИЧЕСКИХ СРЕДАХ

Аннотация

В работе рассматриваются обратные задачи эволюционного типа для математических моделей квазистационарных электромагнитных волн. В модели предполагается, что длина волны мала по сравнению с пространственными неоднородностями. Вводя электрический и магнитный потенциал получаем эллиптическое уравнение второго порядка по пространственным переменным, содержащее интегральные слагаемые типа свертки по времени. После дифференцирования по времени задача сводится к уравнению составного типа с интегральным слагаемым. Определению вместе с решением подлежат неизвестные коэффициенты в интегральном операторе. Дополнительно к краевым условиям задаются условия переопределения в виде заданного набора функционалов от решения, которые могут иметь произвольны вид (интегралы от решения с весом, значения решения в отдельных точках, и пр.). Мы строим численный алгоритм решения этой обратной задачи и приводим результаты численных экспериментов. Алгоритм основан на полученных ранее теоретических результатах, где задача сводилась к некоторому уравнению Вольтерра и доказывалась сходимость метода последовательных приближений. Соответствующий оператор являлся сжимающим на малом промежутке времени. В данной работе в качестве метода используется метод последовательных приближений и метод конечных элементов.

Ключевые слова: Уравнение Соболевского типа, эллиптическое уравнение, уравнения с памятью, обратная задача, краевая задача, численное решение, метод конечных элементов.

Введение

Мы будем рассматривать математические модели возникающие при описании процессов распространения электромагнитных волн в анизотропных средах [1] и при рассмотрении нестационарных внутренних волн в несжимаемой стратифицированной вращающейся жидкости [2]. Если длина волны мала по сравнению с пространственными неоднородностями, то эти уравнения можно записать в виде, учитывающем только временную дисперсию и после введения потенциалов электрического и магнитного поля $E = -\nabla\varphi(x,t)$, $H = -\nabla\psi(x,t)$ и некоторых преобразований, мы придем к уравнениям (см. [1], стр. 28)

$$\sum_{i=1}^{3}(1+4\pi\kappa_i *)\varphi_{x_i x_i} = -4\pi\rho + F_d, \sum_{i=1}^{3}(1+4\pi\kappa_i *)\psi_{x_i x_i} = F_b \quad (1)$$

где κ_i – диагональные элементы тензоров электрической и магнитной восприимчивости и $\kappa_i * \varphi(x,t) = \int_0^t \kappa_i(t-\tau)\varphi(x,\tau)\,d\tau$. В данной работе мы рассмотрим обратные задачи об определении коэффициентов k_i для общих уравнений вида

$$L_0 u + \sum_{i=1}^{m} \kappa_i * L_i(x,t)u = f, \quad (2)$$

где $L_k u = \sum_{i,j=1}^{n} a_{ij}^k(x,t)u_{x_i x_j} + \sum_{i=1}^{n} a_i^k(x,t) + a_0^k(x,t)u$, $(x,t) \in Q = G\times(0,T)$, $G \subset \mathbb{R}^n$. Уравнение (2) дополняется условиями переопределения

$$\Psi_j(u)(t) = \psi_j(t), \quad (3)$$

где Ψ_i – некоторые функционалы, и краевыми условиями

$$Bu|_S = g(x,t), \qquad S = \partial G\times(0,T), \quad (4)$$

где $Bu = u$ или $Bu = \sum_{i=1}^{n}\gamma_i(x,t)u_{x_i} + \sigma(x,t)u$. Близкие к (2) уравнения и системы возникают в теории упругости (материалы с памятью) [3, 4], физике (модели фазового поля, теплоперенос) [10]–[11], и во многих других областях. Наиболее полно рассмотрен случай параболического (см. [3]–[9]) или гиперболического (см. [11, 12] оператора L_0, причем рассмотрен и самый общий случай, когда $L_0 = \partial_t - A$ или $L_0 = \partial_t^2 - A$, где оператор A уже является не обязательно дифференциальным, а является генератором аналитической полугруппы (см., например, [6]-[9]). Случай псевдопараболического оператора L_0 рассмотрен в [13]. В случае $L_0 = \partial_t$, мы приходим к моделям типа Гуртина-Пипкина (см. [14,15]). Эллиптический случай, по-видимому, не рассматривался, за исключение одной модельной ситуации (см. [16]), где n = 1.Обратные задачи для псевдопараболических уравнений рассматривались, например, в [18]–[20]. В данной работе, основываясь на результатах работы [21], мы описываем численный алгоритм решения задачи (2)–(4) и приводим результаты численных экспериментов

1. Определения и вспомогательные результаты

Мы будем использовать результаты из работы [21]. Считаем, что оператор L_0 эллиптичен,

$$a_{ij}^0 \in C(\overline{Q}),\quad a_i^0, a_0^0 \in C\left([0,T];L_p(G)\right) a_{ij}^k \in L_\infty(Q),$$
$$a_{ijt}^0, a_{ijt}^k \in L_p(0,T;L_\infty(G)),\qquad a_i^k, a_0^k \in L_\infty\left(0,T;L_p(G)\right), a_{0t}^k, a_{it}^k \in L_p(Q)$$
$$(i,j=1,2,..,n,\qquad k=1,2,\dots,m),\qquad \gamma_i,\gamma_{it},\sigma,\sigma_t \in C^1(\overline{S})\ (i=1,\dots,n), \tag{5}$$
$$\left|\sum_{i=1}^n \gamma_i n_i\right| \geq \delta_1 > 0\ \ \forall (x,t) \in S,\qquad p > n,$$

и выполнены условия согласования

$$Bu_0(x) = g(x,0), \tag{6}$$
$$\Psi_j(u_0(x)) = \psi_j(0)\quad j = 1,2,\dots,m. \tag{7}$$

Пусть B – матрица с элементами $b_{ij} = \Psi_i(L_0^{-1}L_j(x,0)u_0(x))$. Мы предположим, что существует $\delta_2 > 0$ такая, что

$$|\det B| \geq \delta_2 \quad \forall t \in [0,T], \tag{8}$$

где $L_0^{-1}f$ есть решение v_0 задачи: $L_0 v_0 = f$, $Bv_0|_\Gamma = 0$. Мы рассматриваем задачу (2)–(4) считая, что функционалы Ψ_j удовлетворяют условиям

$$\Psi_j \in L(W_p^2(G),\mathbb{R}),\quad \Psi_j(u_0) = \psi_j(0),\quad j = 1,2,\dots,m, \tag{9}$$

и, таким образом, $\Psi_j \in (W_p^2(G))^*$. Кроме того будем считать, что $a_0^0 \leq 0$ п.в. в случае граничных данных Дирихле и $a_0^0 < -\delta_1 < 0$ $(\delta_1 > 0)$ п.в. в случае условий с косой производной.

Теорема 1 [21]. *Пусть* $f, f_t \in L_p(Q)$, *выполнены условия (5)-(9),* $g \in C([0,T];W_p^{S_0}(\Gamma))$, $g_t \in L_p(0,T;W_p^{S_0}(\Gamma))$ $(p > n)$, $\psi_j \in W_p^1(0,T)$ $(j = 1,2,\dots,m)$. *Тогда существует единственное решение задачи* (2)-(4) *такое, что* $u_t \in L_p(0,T;W_p^2(G))$, $u \in C([0,T];W_p^2(G))$, $k_i \in L_p(0,T)$ $(i = 1,2,\dots,m)$.

Доказательство этой теоремы достаточно конструктивно и мы будем использовать рассуждения их него для построения численного алгоритма. Для простоты мы рассмотрим случай условия граничных Дирихле, т.е. $Bu = u$ в (4). Пусть Φ есть решение задачи

$$L_0(t)\Phi = f(x,t),\qquad \Phi|_S = g(x,t), \tag{10}$$

такое, что $\Phi(x,t) \in C([0,T];W_p^2(G))$, $\Phi_t \in L_p(0,T;W_p^2(G))$, $\Phi|_S = g(x,t)$ и $\Phi(x,0) = u_0(x)$, где u_0 есть решение задачи (10) при $t = 0$. Сделаем замену $u = v + \Phi$. Дифференцируя уравнение (2) по t, придем к равенству

$$(L_0 v)_t + \int_0^t \sum_{j=1}^m k_i(\tau)((L_j v)_t(x,t-\tau) + \left(L_j\Phi\right)_t(x,t-\tau) \tag{11}$$
$$\sum_{j=1}^m k_j(t)L_j(x,0)u_0(x) = 0,$$
$$Bv|_\Gamma = 0,\quad v(x,0) = 0, \tag{12}$$
$$\Psi_j(v) = \tilde{\psi}_j = \psi_j - \Psi_j(\Phi),\qquad j = 1,2,\dots,m. \tag{13}$$

Преобразуя это уравнение получим

$$L_0 v(x,t) = -\int_0^t \int_0^\xi \sum_{j=1}^m k_j(\tau)(\tau)((L_j v)_t(x,\xi-\tau) + (L_j\Phi)_t(x,\xi-\tau))\,d\tau d\xi \tag{14}$$
$$-\sum_{j=1}^m \int_0^t k_j(\xi)\,d\xi L_j(x,0)u_0(x) = G(v)$$

или

$$v(x,t) = L_0^{-1}G(v). \tag{15}$$

С другой стороны будем иметь

$$v_t(x,t) + L_0^{-1}L_{0t}v + L_0^{-1}\int_0^t \sum_{j=1}^m k_j(\tau)((L_j v)_t(x,t-\tau) + \tag{16}$$
$$\left(L_j\Phi\right)_t(x,t-\tau)\big)d\tau = -\sum_{j=1}^m k_j(t)L_0^{-1}L_j(x,0)u_0(x),$$

Применяя функционал Ψ_i к (16), получим, что

$$\tilde{\psi}_{it} + \Psi_i(L_0^{-1}L_{0t}v) + \Psi_i(L_0^{-1}\int_0^t \sum_{j=1}^m k_j(\tau)(L_j v)_t(x,t-\tau)\,d\tau) = \tag{17}$$
$$-\sum_{j=1}^m k_j(t)\Psi_i(L_0^{-1}L_j(x,0)u_0(x) - L_0^{-1}\int_0^t \sum_{j=1}^m k_j(\tau)(L_j\Phi)_t(x,t-\tau)\,d\tau).$$

В силу условия (8) из этого равенства вытекает, что

$$\vec{k} = B^{-1}\vec{F} - B^{-1}A(\vec{k}), \tag{18}$$

где $\vec{F}$ имеет координаты $F_i = -\tilde{\psi}_{it} \in W_p^1(0,T)$ $(i = 1,2,\dots,m)$ и оператор $A(\vec{k})$ координаты

$$A_i(\vec{k}) = \Psi_i(L_0^{-1}L_{0t}v + L_0^{-1}\int_0^t \sum_{j=1}^m k_j(\tau)((L_j v)_t(x,t-\tau) + (L_j\Phi)_t(x,t-\tau))d\tau),$$

где $v = v(\vec{k})$ решение задачи (11), (12). Получили систему для нахождения вектор-функции $\vec{k} = (k_1,k_2,..,k_m)$.

2. Описание алгоритма

Приведем описание численного алгоритма в случае краевых условий Дирихле при $m = 2$ и уравнения вида (1)

$$\Delta u + \int_0^t \sum_{m=1}^2 p_m(s)u_{x_m x_m}(t-s)ds = f(x,t), \tag{19}$$

рассматриваемого в области $Q = G\times(0,T)$, $x = (x_1,x_2) \in G \subset \mathbb{R}^2$, $t \in [0,T]$. Уравнение (19) дополняется граничными условиями

$$u|_S = g(x,t),\qquad u(x,0)|_\Gamma = g(x,0). \tag{20}$$

Функция $p_m(s)$ определяется на основе дополнительных измерений вида

$$u(x_j,t) = \psi(t),\ \text{j=1,2} \tag{21}$$

Численный алгоритм можно разбить на три части.

1. Поиск приближения начальных данных u_0. Начальные условия при $t = 0$, $u(x,0) = u_0(x)$ определяются как решение задачи

$$\Delta u_0 = f(x,0),\quad u_0|_\Gamma = g(x,0).$$

Функцию u_0 ищем в виде $u_0(x) = \tilde{g} + v_0$, где $\tilde{g}(x)$ – некоторая функция такая, что $\tilde{g}(x)|_\Gamma = g(x,0)$, т.е. $\tilde{g}(x)$ есть продолжение функции $g(x,0)$ внутрь области. Функция v_0 есть решение задачи

$$\Delta v_0 = f(x,0) - \Delta\tilde{g}(x),\quad v_0|_\Gamma = 0. \tag{22}$$

Строим приближение v_0 в виде

$$v_0 = \sum_{i=1}^N C_i\varphi_i(x), \tag{23}$$

где $\varphi_i(x)$ – кусочно-линейные базисные функции в методе конечных элементов, равные 0 на Γ, построенные для некоторой фиксированной триангуляции области.

Числа C_i находятся из системы

$$-\sum_{i=1}^{N} C_i(\nabla\varphi_j, \nabla\varphi_i) = (f(x,0), \varphi_i) + (\nabla\tilde{g}(x), \nabla\varphi_i), \quad (24)$$

где $i = 1,2,\dots,N$, N - количество узлов в области.

2. Поиск решения Φ вспомогательной задачи.

Продифференцируем уравнение (19) по t и сделаем замену $u_t = \Phi + v$, где Φ есть решение вспомогательной задачи

$$\Delta\Phi = f_t, \qquad \Phi|_S = g_t(x,t).$$

Приближение функции Φ можно найти как сумму

$$\Phi = \tilde{g} + \sum_{i=1}^{N} C_i\varphi_i,$$

где $\tilde{g}$ - продолжение функции g_t внутрь области.

3. Поиск решения задачи v. Рассмотрим уравнение

$$\Delta v + \sum_{m=1}^{2}\int_0^t p_m(s) v_{x_m x_m}(t-s)ds + \sum_{m=1}^{2}\int_0^t p_m(s)\Phi_{x_m x_m}(t-s)ds + \sum_{m=1}^{2} p_m(s) u_{0 x_m x_m}(t-s)ds = 0 \quad (25)$$

Скалярно умножая уравнение (25) на φ_i, получим $(i = 1,2,\dots,N)$

$$(\nabla v, \nabla\varphi) + \int_0^t \sum_{m=1}^{2} p_m(s)(v_{x_m}, \varphi_{x_m})ds + \int_0^t \sum_{m=1}^{2} p_m(s)(\Phi_{x_m}, \varphi_{x_m})ds + \sum_{m=1}^{2} p_m(s)(u0_{x_m}, \varphi_{x_m})ds = 0. \quad (26)$$

Ищем приближение функции v в виде $v = \sum_{i=1}^{N} C_i\varphi_i$. Тогда уравнение (26) переписывается в виде

$$\sum_{j=1}^{N} C_j(\nabla\varphi_j, \nabla\varphi_i) + \sum_{m=1}^{2}\int_0^t p_m(s)\sum_{j=1}^{N} C_j(\varphi_{jx_m}, \varphi_{ix_m})(t-s)ds + \sum_{m=1}^{2}\int_0^t p_m(s)(\Phi_{x_m}, \varphi_{ix_m})(t-s)ds + \sum_{m=1}^{2} p_m(t)(u_{0x_m}\varphi_{ix_m}) = 0, \quad i,j = 1,2,\dots N. \quad (27)$$

Введем величины

$$K:\ k_{ij} = (\varphi_j, \varphi_i),$$

$$B_m:\ \beta_{ij}^m = (\varphi_{ix_m}, \varphi_{jx_m}),$$

$$\vec{F}0_m:\ F0_{mi} = (u0_{x_m}, \varphi_{ix_m}),\ \vec{F}_m:\ F_{mi} = (\Phi_{x_m}, \varphi_{ix_m})$$

Тогда уравнение (30) перепишется в виде

$$K\vec{C} + \int_0^t \sum_{m=1}^{2} p_m(s)B_m\vec{C}(t-s)ds + \int_0^t \sum_{m=1}^{2} p_m(s)\vec{F}_m(t-s)ds + \sum_{m=1}^{2} p_m(s)\vec{F}0_m = 0. \quad (28)$$

3. Численная реализация алгоритма

Задаём триангуляцию области G, узлы триангуляционной сетки $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_m$ и соответствующие базисные кусочно-линейные функции $\{\varphi_i(x)\}$. Без ограничения общности считаем, что точки замера x_j являются узлами сетки. Тогда найдутся номера j_1, j_2 такие, что $x_k = \tilde{x}_{j_k},\ k = 1,2$.

1. Поиск приближения u_0. Определим матрицу жесткости K состоящую из элементов $K_{ji} = \int_G (\nabla\varphi_i\nabla\varphi_j)dx$, а также вектора $\vec{F}0, \vec{\tilde{g}}$ с координатами $\int_G f(0)\varphi_i dx$ и $\int_G \Delta\tilde{g}\varphi_i dx$. Приближённое решение (22) ищем в виде $v_m = \sum_{i=1}^{m} C_i\varphi_i(x)$. Вектор $\vec{C} = (C_1, C_2, \dots, C_m)$ записывается в виде $\vec{C} = -K^{-1}(\vec{F}0 - \vec{\tilde{g}})$.

2. Поиск решения вспомогательной задачи Φ.

Ищем приближение функции Φ в виде $\Phi^m = \sum_{i=1}^{m} c_i(t)\varphi_i(x) + \tilde{g}$. Вектор функция $\vec{c}(t)$ определяем следующим образом. Найдем вектора $\vec{F}_t$, $\vec{\tilde{g}}$ с координатами $\int_G f_t\varphi_i dx$ и $\int_G \Delta\tilde{g}\varphi_i dx$. Здесь $\tilde{g}$ – продолжение функции g_t внутрь области. Тогда $\vec{c}(t) = -K^{-1}(\vec{F}_t - \vec{\tilde{g}})$.

3. Поиск решения вспомогательной задачи v. Заменим интегралы в (28) на интегральные суммы

$$KC_l + \tau\sum_{m=1}^{2}\sum_{j=1}^{l} p_{m,j}B_mC_{l-j} + \tau\sum_{m=1}^{2}\sum_{j=1}^{l} p_{m,j}F_m((l-j)\tau) + \sum_{m=1}^{2} p_m\vec{F}0_m = 0, \quad l = 1,2,\dots,n, \quad (29)$$

где τ – шаг по времени и таким образом промежуток $[0,T]$ разбивается на промежутки вида $[(k-1)\tau, k\tau]$, $k = 1,2,\dots,n$. Перепишем условия переопределения (21): $v(y_m,t) = \psi_{mt}(t) - \Phi(y_m,t)$. Пусть $y_1 = \tilde{x}_{j_1}$, $y_2 = \tilde{x}_{j_2}$, тогда

$$\psi_{1t} - \Phi(\tilde{x}_{j_1},t) = v(\tilde{x}_{j_1},t) = C_{j_1}(t), \psi_{2t} - \Phi(\tilde{x}_{j_2},t) = v(\tilde{x}_{j_2},t) = C_{j_2}(t),,$$
$$(C_i)_{j_1} = \psi_{1t} - \Phi(\tilde{x}_{j_1},t)|_{t=i\tau}, (C_i)_{j_2} = \psi_{2t} - \Phi(\tilde{x}_{j_2},t)|_{t=i\tau}, \quad (30)$$

где $i = 1,2,\dots,n$. Имеем $(k = 1,2)$

$$(C_i)_{j_k} + \tau\sum_{m=1}^{2}\sum_{j=1}^{l} p_{m,j}(K^{-1}B_mC_{l-j})_{j_k} + \tau\sum_{m=1}^{2}\sum_{j=1}^{l} p_{m,j}(K^{-1}F_m((l-j)\tau))_{j_k} + \sum_{m=1}^{2} p_{m,l}(K^{-1}\vec{F}0_m)_{j_k} = 0 \quad (31)$$

Определим $p_{m,0}$, подставив в (31) $l = 0$:

$$(C_0)_{j_k} = (\psi_{kt}(0) - \Phi(x_{j_k},0)) = -\sum_{m=1}^{2} p_{m,0}(K^{-1}\vec{F}0_m)_{j_k}$$

Отсюда найдем $p_{m,0} = p_m(0)$ $(m = 1,2)$ обращая соответствующую матрицу с элементами $f_{km} = (K^{-1}\vec{F}0_m)_{j_k}$. Далее вектор C_0 будет найден как

$$C_0 = -\sum_{m=1}^{2} p_{m,0}K^{-1}\vec{F}0_m.$$

Перепишем формулу(29) как

$$C_l + K^{-1}(\tau(\sum_{m=1}^{2}\sum_{j=1}^{l-1} p_{m,j}B_mC_{l-j} + \sum_{m=1}^{2} p_{m,l}B_mC_0 + \sum_{m=1}^{2}\sum_{j=1}^{l-1} p_{m,j}F_m((l-j)\tau) + \sum_{m=1}^{2} p_{m,l}F_m(0)) + \sum_{m=1}^{2} p_{m,l}\vec{F}0_m) = 0. \quad (32)$$

Для того, чтобы найти вектор $p_{m,l}$ (при известных векторах $p_{m,l-1}\dots$) используем равенства (30), (31). Вектор $p_{m,l}$ выразим из (31) обращая матрицу с элементами f_{kj} определенную выше. Тогда вектор C_l определится из равенства

$$C_l = -K^{-1}(\tau(\sum_{m=1}^{2}\sum_{j=1}^{l-1} p_{m,j}B_mC_{l-j} + \sum_{m=1}^{2} p_{m,l}B_mC_0 + \sum_{m=1}^{2}\sum_{j=1}^{l-1} p_{m,j}F_m((l-j)\tau) + \sum_{m=1}^{2} p_{m,l}F_m(0)) + \sum_{m=1}^{2} p_{m,l}F0_m) \quad (33)$$

Повторяя рассуждения на каждом временном слое, мы определим вектора C_l и величины $p_{m,l}$.

Результаты численных экспериментов

Характеристики испытуемой ЭВМ: Процессор: Intel(R) Core(TM) i7-3517U CPU @ 1.90GHz 2.40GHz; ОЗУ: 10,00 ГБ; Тип системы: 64-разрядная операционная система Windows 7 Ultimate.

В результате расчётов определялось приближённое значение решения (u, p_1, p_2) задачи (19)–(21) в узлах сетки в моменты времени $t_1,t_2,\ldots,t_N$. В качестве области возьмём круг единичного радиуса с координатами центра (0,0).

Для численного решения задачи будут рассматриваться несколько сеток для описанной области с количеством узлов зависящим от расположения точек наблюдения и количества узлов сетки.

Решение задачи (19)–(21), а так же граничные условия запишем в виде:

$$u(x,t)=(x^2+1)(y^2+1)(t^2+1),$$

$$p_1(t)=-\sqrt{((t+1))},\ p_2(t)=\ln(t^2+1),\ g(x,0)|_\Gamma=x^2y^2+2.$$

Полученный вектор правой части: $f=1.06t+(2(x^2+1)(12atan(t)-12t+18t^2\cdot atan(t)+3t^3\cdot\ln(t^2+1)-11t^3))/9-2y^2\cdot(0.66(t+1)^{3/2}-0.66)-1.33\cdot(t+1)^{3/2}+2(t\hat{}2+1)(x^2+1)+2(t^2+1)(y^2+1)+2y^2(0.53t+0.66t^2-0.019(t+1)^{3/2}(16t+8t^2+8)+0.15)+1.33t^2-0.038\cdot(t+1)^{3/2}(16t+8t^2+8)+1.638$.

Введем несколько величин: $\varepsilon_p=\max_i|p^m(i\tau)-p_i^m|$ – ошибка вычислений, использующая среднее арифметическое $\varepsilon_u=\max_{i,j}|u_{i,j}-u(i,j\tau)|$, где $i=1,2,\ldots,n$, $j=1,2,\ldots,l$. Возьмём за τ_s время вычислений в секундах.

Результаты численных экспериментов для шага по времени $t_s=0.02$ представлены в таблице 1. В них изменяются положения точек наблюдения, в связи с чем меняется и сетка области.

Как можно заметить увеличение точности вычислений, зависит в первую очередь от однообразия треугольников в триангуляционной сетке и не жёстко связано с их количеством. В работе не будут представлены графики исходной и восстановленной функций p_1 и p_2, как как полученные графики, в виду высокой точности вычислений, практически идентичны. Для следующих нескольких экспериментов увеличим количество временных промежутков в два раза, рассмотрим только результаты для сеток 1–5 и 11–15.

Таблица 1

Результаты численных экспериментов

№	N_x	ε_p	ε_u	τ_s
1	229	0,037002	0,046466	4,147049
2	236	0,011253	0,044536	4,184662
3	237	0,015125	0,047002	4,242981
4	236	0,021869	0,04542	4,004972
5	239	0,027241	0,046797	4,036889
6	669	0,048357	0,04537	49,12766
7	693	0,037822	0,044257	52,20402
8	694	0,014884	0,046501	51,6243
9	698	0,015127	0,045502	51,32633
10	689	0,037445	0,049313	48,90397
11	897	0,017754	0,041931	81,42314
12	907	0,012555	0,04202	86,20059
13	911	0,012972	0,042302	90,10167
14	905	0,015084	0,042173	83,36037
15	915	0,008126	0,042166	89,42623
16	2639	0,022252	0,042182	1755,792
17	2735	0,019808	0,041656	1846,957
18	2736	0,014147	0,042299	1801,659
19	2749	0,013294	0,042023	1790,621
20	2710	0,009836	0,043133	1811,33

Таблица 2

Результаты численных экспериментов

№	N_x	ε_p	ε_u	τ_s
1	229	0,03007	0,027882	15,77482
2	236	0,01078	0,026787	16,59276
3	237	0,00867	0,028076	16,87762
4	236	0,01492	0,027352	16,27913
5	239	0,02724	0,027731	16,263
6	897	0,01076	0,021724	321,624
7	907	0,00555	0,021355	334,0045
8	911	0,00602	0,022035	339,4294
9	905	0,00807	0,021586	324,4042
10	915	0,00713	0,021937	329,5462

Выводы

Представлен метод определения неизвестных параметров в уравнении (19). Численный алгоритм основан на методе конечных элементов. Результаты численных экспериментов показывают хорошую точность вычислений. Точность определения параметров p_1, p_2 возрастает с уменьшением шага по времени, а так же с увеличением количества узлов, при условии, однородности пространственной сетки.

Благодарности

Работа авторов поддержана РФФИ и правительством Ханты-Мансийского автономного округа – Югры (Грант 18-41-860003, р_урал_а).

Список литературы

1. Свешников А.Г., Альшин А.Б., Корпусов М.О., Плетнер Ю.Д. Линейные и нелинейные уравнения Соболевского типа. ФМЛ, 2007.

2. Габов С.А., Свешников А.Г. Линейные задачи теории нестационарных внутренних волн. М.: Наука, 1990.

3. Colombo F., Guidetti D. An Inverse Problem for a Phase-field Model in Sobolev Spaces. Nonlinear Elliptic and Parabolic Problems. Progress in Nonlinear Differential Equations and Their Applications. Basel: Birkhäuser Verlag. 2005. Vol. 64, P. 189–210.

4. Guidetti D., and Lorenzi A. A mixed type identification problem related to a phase-field model with memory // Osaka J. Math. 2007. Vol. 44. p. 579–613.

5. Colombo F., Guidetti D. A global in time existence and uniqueness result for a semilinear integrodifferential parabolic inverse problem in Sobolev spaces // Mathematical Models and Methods in Applied Sciences Vol. 2007.Vol. 17, No. 4.p.537–565

6. Colombo F. On some methods to solve integro-differential inverse problems of parabolic type // ВестникЮУрГу. Серия «Математическое модели- рование и программирование». 2015. Т. 8, 3. С. 95–115.

7. Favini A., Lorenzi A. Identication problems for singular integro-differential equations of parabolic type II // Nonlinear Analysis. 2004. Vol. 56, No.6 p. 879–904.

8. Lorenzi A., Tanabe H. Inverse and direct problems for nonautonomous degenerate integro-differential equations of parabolic type with Dirichlet boundary conditions. Differential Equations: inverse and Direct Problems. Lecture notes in pure and applied mathematics. Vol. 251. Boca Raton, London, New York: Chapman & Hall/CRC Taylor & Francis Group. 2006. P. 197–244.

9. Abaseeva N., Lorenzi A. Identification problems for nonclassical integro-differential parabolic equations // J. Inv. Ill-Posed Problems. 2005. Vol. 13, No. 6, p. 513–535.

10. Lorenzi A., Paparone I. Direct and inverse problems in the theory of materials with memory. Rend. Sem. Mat. Univ. Padova. 1992. Vol. 87. p. 105–138.

11. Janno J., Von Wolfersdorf L. Inverse Problems for Identification of Memory Kernels in Viscoelasticity // Mathematical Methods in the Applied Sciences. 1997. Vol. 20, P. 291–314.

12. Durdiev D.K., Safarov Zh.Sh. Inverse problem of determining the one-dimensional kernel of the viscoelasticity equation in a bounded domain // Mathematical Notes. 2015. Vol. 97, No. 6, p. 867–877.

13. Асанов А., Атаманов Э. Р., Обратная задача для операторного интегро-дифференциального псевдопа-раболического уравнения // Сиб. матем. журн. 1995. Т. 36, 4, С. 752–762.

14. Avdonin S.A., Ivanov S.A., Wang j. Inverse Problems for the Heat Equation with Memory. 2016. https://arxiv.org/abs/1612.02129

15. Pandolfi L. Identification of the relaxation kernel in diffusion processes and viscoelasticity with memory via deconvolution. 2016. https://arxiv.org/abs/1603.04321.

16. Денисов А.М. Обратная задача для квази-линейного интегро-дифференциального уравнения // Дифференц. уравнения 2001. Т. 37, 10, с. 1350–1356.

17. Pyatkov S.G., Shergin S.N. On some mathematical models of filtration theory. // Вестник ЮУрГУ. Серия «Математическое моделирование и программирование». 2015. Т. 8, 2. С. 105–116.

18. Lyubanova A.Sh. Identification of a coefficient in the leading term of a pseudoparabolic equation of filtration, Sibirskiimatematicheskiizhurnal 2013 № 6, p. 1315–1330.

19. Lyubanova A.Sh., Tani A. On inverse problems for pseudoparabolic and parabolic equations of filtration // Inverse problems in science and engineering. 2011. V. 19, No. 7, p. 1023–1042.

20. Кожанов А.И. О разрешимости обратных задач восстановления коэффициентов в уравнениях составного типа // Вестн. НГУ. Сер. матем., мех., информ. – 2008. – Т. 8, Вып. 3. – С. 81–99.

21. S.G. Pyatkov, S.N. Shergin. Inverse problems for mathematical models of quasistationary electromagnetic waves in anisotropic nonmetallic media with dispersion, Вестник ЮУрГУ. Серия «Математическое моделирование и программирование» . 2018. Т. 11, 1. С. 44–59.

УДК 004.032.6

К.Ф. Тагирова, Д.А. Талипов, А.М. Вульфин, К.В. Миронов

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ РЕЖИМАМИ РАБОТЫ КАМЕРЫ СГОРАНИЯ ГАЗОТУРБИННОЙ УСТАНОВКИ

Аннотация

Рассмотрены вопросы создания интеллектуальной системы нейросетевого управления пульсацией акустического давления в камере сгорания газотурбинных установок (ГТУ), позволяющей выявлять начало процесса вибрационного горения в автоматическом режиме на основе диагностических данных с существующего агрегата.

Ключевые слова: интеллектуальные системы управления, вибрационное горение, нейросетевой анализ данных.

Введение

Возникновение вибрационного горения является существенной проблемой камер сгорания индустриальных газотурбинных установок (ГТУ), спроектированных с учётом жёстких экологических требований [1, 2]. Возникновение вибрационного горения может стать причиной разрушения конструкции ГТУ. Следовательно, актуальным является создание системы управления пульсацией акустического давления в камере сгорания ГТУ, позволяющей выявлять начало процесса вибрационного горения в автоматическом режиме. Способ обнаружения вибрационного горения основан на анализе процесса затухания колебаний и уменьшении интенсивности их генерации с помощью цифровой обработки виброакустических сигналов и нейросетевого анализа параметров наблюдаемого процесса.

Целью является автоматизация процесса обнаружения пульсаций давления в камере сгорания газотурбинной установки для своевременного предотвращения аномальных режимов.

САУ ГТУ должна выполнять непрерывный мониторинг сигналов, поступающих от датчиков, сигнальных устройств и исполнительных механизмов ГТУ, преобразование полученной информации в цифровой код и использовать эту информацию для управления режимом горения. По результатам анализа САУ ГТУ генерирует предупреждающие или аварийные сообщения о достижении предельных значений контролируемых параметров, формирует управляющие воздействия на исполнительные механизмы для остановки ГТУ.

1. Анализ подходов к диагностике вибрационного горения

Задача надежного определения и предотвращения режима вибрационного горения в работе газотурбинных двигателей при их испытаниях и в эксплуатации остается актуальной, несмотря на существование различных способов решения этой задачи.

В большинстве известных решений используется анализ спектральных характеристик контролируемых акустических сигналов. Недостатками приведенных способов являются факторы, снижающие достоверность выявления режима вибрационного горения (например, влияние вибрации корпуса камеры, внешних электростатических полей). Кроме того, в известных подходах существует необходимость установки датчиков в высокотемпературной зоне форсажной камеры сгорания, необходимость специального окна в камере сгорания для обеспечения передачи излучения от газа на фотоумножитель, юстировка оптических элементов. Указанные недостатки требуют дополнительных мер по учету (устранению) сопутствующих помех или конструктивных доработок.

Современные методы диагностики состояния ГТУ [3–10] в составе САУ достигли хороших показателей в безопасности и бесперебойной работе агрегатов. Выявление пульсаций давления в камере сгорания – вибрационного горения – с последующей возможностью управления режимами работы за счет регулирования подачи топлива в ГТУ позволяет предотвратить некоторые из проблем ещё на раннем этапе их возникновения, увеличивая продолжительность работы агрегата за счёт сохранности составляющих его элементов. Однако влияние человеческого фактора по-прежнему уменьшает время эксплуатации установки в целом и создает опасность возникновения аварийной ситуации в связи с задержкой действий оператора установки при оценке текущего режима работы установки. Для автоматизации обнаружения процессов вибрационного горения и, как следствие, повышения безопасности системы при внештатных ситуациях предлагается использование предобученного нейросетевого блока диагностики.

Использование интеллектуальных технологий диагностики на основе нейронных сетей позволит повысить качество распознавания признаков режимов работы камеры сгорания, как штатных, так и аномальных, а также нивелировать влияние неинформативных составляющих контролируемых сигналов.

2. Разработка структуры интеллектуальной системы управления режимами работы камеры сгорания газотурбинной установки

В процессе запуска и при работе в различных режимах система автоматического управления ГТУ ограничивает предельные параметры ГТУ, контролирует

параметры газотурбинного двигателя и работоспособность самой САУ ГТУ, генерирует предупреждения и сигналы тревоги в соответствии с заранее определенными алгоритмами (при выходе контролируемых параметров за пределы допустимого диапазона).

Параметры вибрации (вибросмещение, виброскорость, виброускорение), выраженные в абсолютных или относительных величинах, применяются для описания отдельных гармонических составляющих процесса вибрации. В случае сложной вибрации, которая не описывается простым гармоническим процессом, для характеристики вибрации применяют пиковое значение $x_{ПИК}$, эффективное значение $x_{ЭФ}$ и среднее значение $x_{СР}$.

Для сложного процесса вибрации пиковые значения дают ограниченную информацию, т.к. они описывают процесс в отдельные моменты времени без учета взаимосвязи между ними. Средние и эффективные значения применяются для анализа вибрации, которая не является гармонической и может быть, например, случайной величиной. Применение анализа пиковых значений вибрации и усредненных значений не позволяют полностью решить проблему своевременного выявления начала процесса вибрационного горения ввиду вариабельности параметров сигнала не только для различных агрегатов, но и для одной и той же ГТУ. Следовательно, необходимы адаптивные и самонастраиваемые инструменты анализа виброакустических сигналов, которые позволили бы повысить эффективность управления ГТУ за счет своевременного выявления нежелательных процессов и перевода установки в соответствующий режим работы. Подобные инструменты анализа предлагается реализовать на основе нейросетевых технологий обработки виброакустических сигналов.

Для анализа и разработки нейросетевого блока диагностики вибрационного горения использованы реальные данные с газотурбинного комплекса. Измерение параметров акустического поля и фиксация параметров процесса горения в камере сгорания ГТУ производилось с помощью измерительного зонда [11, 12].

В существующей системе управления установкой индикатор наличия вибрации на пульте управления оператора оповещал о начале газоколебательных процессов в камере сгорания. Наличие нейросетевого модуля диагностики ГТУ в составе САУ позволит не только автоматически выявлять начало процесса вибрационного горения, своевременно информировать оператора, но и формировать сигнал для блока управления краном подачи топлива для предотвращения развития процесса вибрационного горения.

На рисунке 1 представлена структура интеллектуальной системы управления ГТУ.

Фиксируемые виброакустические данные, полученные с зонда, подвергаются предобработке с помощью стандартных цифровых фильтров для устранения шумовой компоненты. Затем формируются гистрограммы распределения амплитуд сигнала по нормированному диапазону частот, которые используются для выделения признаков вибрационного горения с помощью нейросетевого блока.

Таким образом, предлагаемая САУ ГТУ с блоком нейросетевого анализа параметров процесса горения [13, 14] позволяет контролировать уровень вибрационных колебаний на основе предобработанных данных с зонда, установленного в камере сгорания ГТУ.

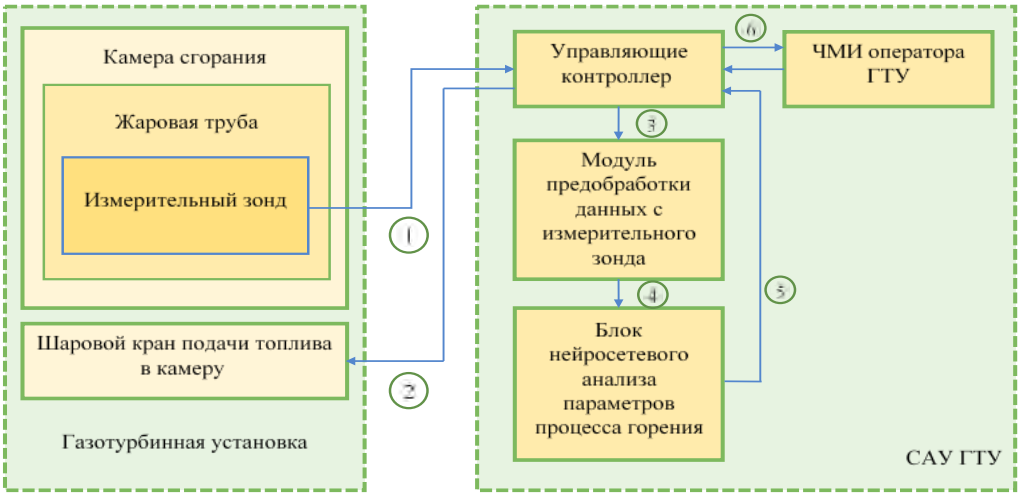
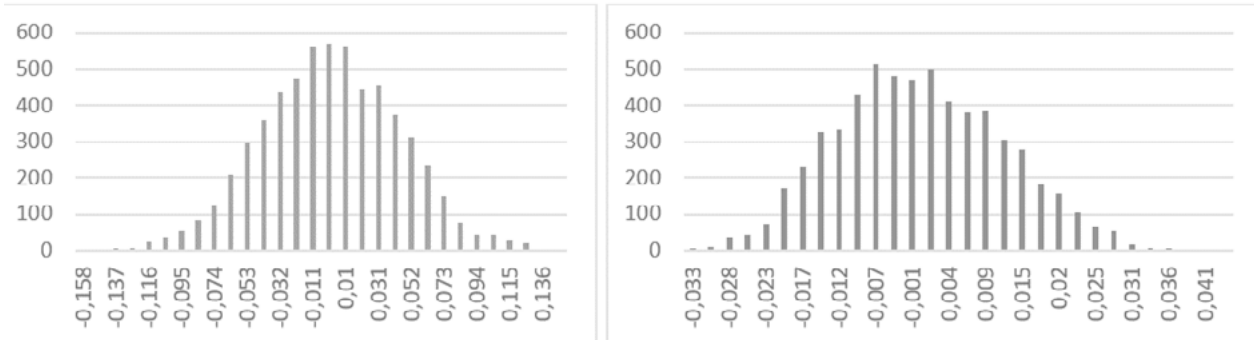


Рис. 1. Структура интеллектуальной системы управления с использованием нейросетевого блока анализа виброакустических сигналов. 1 – данные, фиксируемые виброакустическим зондом и текущие параметры ГТУ; 2 – управление положением крана подачи топлива в камеру сгорания ГТУ; 3 – виброакустические данные в виде дискретной последовательности параметров уровня акустического сигнала w(t); 4 – предобработанные (фильтрация, нормализация) параметры w(t) в виде гистрограммы распределения амплитуд сигнала (нормированные по частоте); 5 – сигнал наличия выявленного процесса вибрационного горения и степень уверенности классификатора; 6 – текущие параметры установки для человеко-машинного интерфейса (ЧМИ) оператора и управляющие воздействия оператора*



а) *б)*

Рис. 2. Гистограмма виброакустического сигнала нормального (а) процесса горения и горения с небольшими отклонениями от нормального (б) с распределением амплитуды пульсаций, где по горизонтальной оси расположены нормированные диапазоны частот колебаний, по вертикальной – количество точек гистограммы в диапазоне

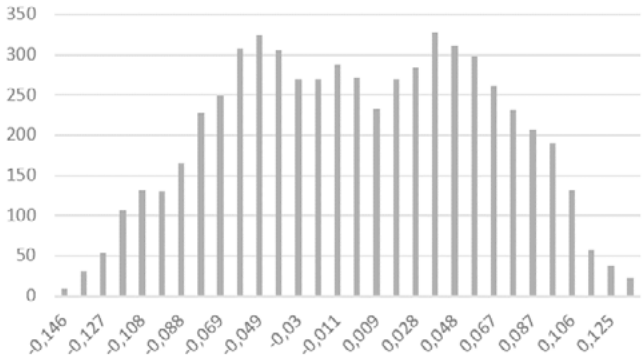


Рис. 3. Гистограмма виброакустического сигнала вибрационного горения с распределением амплитуды пульсаций, где по горизонтальной оси расположены нормированные диапазоны частот колебаний, по вертикальной – количество точек гистограммы в диапазоне

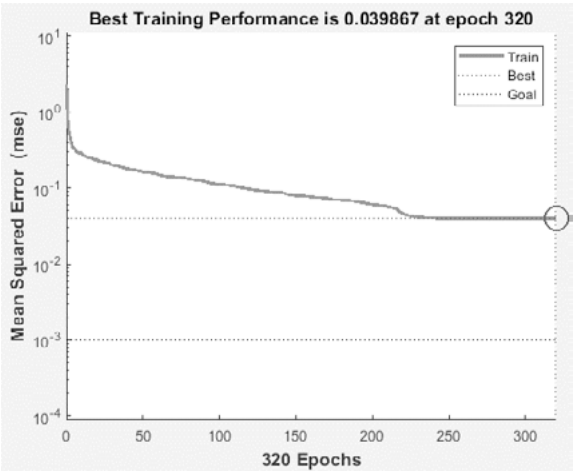


Рис. 4. График изменения ошибки на обучающей выборке для однократного запуска

3. Натурный эксперимента по диагностике возникновения вибрационного горения в камере сгорания газотурбинной установки

Результирующие гистограммы предобработанного виброакустического сигнала в камере сгорания ГТУ представлены на рисунках 2–3 и демонстрируют разницу между нормальным процессом горения и вибрационным горением.

Для анализа использована нейронная сеть (НС) прямого распространения на основе многослойного персептрона (MLP) [15–17], параметры которой приведены в таблице 1. Параметры обучения НС представлены в таблице 2.

Изменение ошибки НС на обучающей выборке при одиночном запуске обучения НС в процессе перекрестной проверки представлен на рисунке 4.

Обученная НС в виде набора весовых коэффициентов перенесена в контроллер [18–21] для малых систем автоматизации ОВЕН ПЛК150-220. В среде *CODESYS* 2.3.9.41 разработана программа, реализующая прием и обработку данных результирующих гистограмм виброакустических сигналов с помощью обученной НС. Анализ тестовых данных показал идентичность результатов работы нейросетевого модуля на ПК и ПЛК.

Выводы

Предложена структура интеллектуальной САУ ГТУ, включающая блок автоматического выявления процесса пульсаций давления (вибрационного горения) в камере сгорания газотурбинной установки на основе предобработки и нейросетевого анализе данных акустического давления. Предложенный нейросетевой блок обнаружения процесса вибрационного горения позволяет добиться выявления до 83 % случаев нежелательного развития процесса.

Разработанный нейросетевой блок диагностики режимов работы камеры сгорания ГТУ реализован на контроллере ОВЕН ПЛК150-220.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта 17-08-01569 «Разработка методологии выбора методов интеллектуального анализа и обработки данных в задачах диагностики и управления сложными техниче-

Таблица 1

Параметры нейросетевого классификатора

Параметры	Значения
Размерность входного вектора	30
Тип входного вектора	Нормированные значения с std = 1, mean = 0
Размерность выходного вектора	1
Тип выходного вектора	Значение в диапазоне от [-1; 1], отражающее степень уверенности сети в наличии или отсутствии вибрационного горения
Количество скрытых слоев	1
Количество нейронов по слоям	30, 24, 1
Функции активации нейронов по слоям	tansig, tansig
Постобработка выхода НС	нет / hardlim
Оценка ошибки сети	Среднеквадратическая ошибка

Таблица 2

Параметры обучения НС

Параметры обучения	Значение
Алгоритм обучения	Алгоритм сопряженных градиентов
Тип верификации	Перекрестная проверка с делением на k = 5 групп
Усредненное значение точности на тестовой выборке	0,833
Усредненное значение чувствительности на тестовой выборке	0,909
Усредненное значение специфичности на тестовой выборке	0,808

скими системами на основе структурно-логического подхода».

Список литературы

1. Третьякова Д.В., 2016. Контроль вибрации пламени в камере сгорания газотурбинных ГПА. Газотурбинные технологии, 1(136): 10–14.

2. Tretyakov D., 2016. A Self-Learning Diagnosis Algorithm Based on Data Clustering. Intelligent Control and Automation, 3(7): 84–92.

3. Neumeier J., B.T. Zinn, 1996. Active Control of Combustion Instabilities with Real Time observation on unstable combustor modes. AIAA 96-0758, 34th Aerospace Sciences Meting, Exhibit, Reno. NV. January 1518. 1996.

4. Tretyakov D., S. Stryapunin and S. Srinivasan, 2014-10-23. Combustor Cap for Damping Low Frequency Dynamics, Patent US 2014 311 156 A1.

5. Slobodyanskiy I. et al., 2014. Acoustically stiffened gas turbine combustor supply. Patent US 8661822 B2.

6. Hobson D.E., J.E. Fackrell and Hewitt G. Combustion instabilities in industrial gas turbines-measurements on operating plant and thermoacoustic modelling. ASME 1999 International Gas Turbine and Aeroengine Congress and Exhibition. American Society of Mechanical Engineers, pp. V002T02A015-V002T02A015.

7. Smith K.O., J. Blust, 2005. Combustion instabilities in industrial gas turbines: Solar turbines' experience. Combustion Instabilities in Gas Turbine Engines: Operational Experience, Fundamental Mechanisms, and Modeling, Vol.210: 29-42.

8. Hubbard S., A.P. Dowling, 2001. Acoustic resonances of an industrial gas turbine combustion system. Journal of engineering for gas turbines and power, 4(123): 766–773.

9. Goldmeer J. et al., 2005. Passive control of dynamic pressures in a dry, low NOx combustion system using fuel gas circuit impedance optimization. ASME Turbo Expo 2005: Power for Land, Sea, and Air, Reno, Nevada, USA, 2005, pp. 453–460.

10. Krüger U. et al., 2000. Prediction and measurement of thermoacoustic improvements in gas turbines with annular combustion systems. ASME Turbo Expo 2000: Power for Land, Sea, and Air. American Society of Mechanical Engineers, 2000. – C. V002T02A015-V002T02A015.

11. Тимофеев, В.В., 2015. «Технекон» – передовые подходы к диагностике турбомашин. Газотурбинные технологии, № 8(135): 6–9.

12. Сведения о компании ООО «Газпром трансгаз Югорск» [Электронный ресурс] Режим доступа: http://yugorsk-tr.gazprom.ru/, свободный. – (Дата обращения: 15.01.2019).

13. Sarkar, S., K.G. Lore and S. Sarkar, 2015. Early Detection of Combustion Instability by Neural-Symbolic Analysis on Hi-Speed Video. In the COCO'15 Proceedings of the 2015th International Conference on Cognitive Computation: Integrating Neural and Symbolic Approaches, pp. 93–101.

14. Cammarata, L., A. Fichera and A. Pagano, 2002. Neural prediction of combustion instability. Applied Energy, 2(72): 513–528.

15. Wang, Q., J.S. Hesthaven and D. Ray, 2018. Non-intrusive reduced order modeling of unsteady flows using artificial neural networks with application to a combustion problem. Journal of Computational Physics.

16. Lv, C. et al., 2018. Levenberg–Marquardt Backpropagation Training of Multilayer Neural Networks for State Estimation of a Safety-Critical Cyber-Physical System. IEEE Transactions on Industrial Informatics, 8(14): 3436–3446.

17. Галушкин, А.И., Цыпкин, Я.З., 2001. Распознавание образов и адаптивное управление. Нейронные сети: история развития теории, с. 211–214.

18. Tisan, A., J. Chin, 2016. An end-user platform for FPGA-based design and rapid prototyping of feedforward artificial neural networks with on-chip backpropagation learning. IEEE Transactions on Industrial Informatics. 3(12): 1124–1133.

19. Wang, Y., H. Li and X. Li, 2018. A Case of On-Chip Memory Subsystem Design for Low-Power CNN Accelerators. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 10(37): 1971–1984.

20. Программируемые логические контроллеры (ПЛК, PLC). [Электронный ресурс] Режим доступа: http://www.ovk-kom.ru/plc.html, свободный. (Дата обращения: 15.01.2019).

21. Корнеев В.В., Киселев А.В., 2003. Современные микропроцессоры. 3-е изд. СПб.: БХВ-Петербург, с. 448.

УДК 517.938.5

А.Ф. Шориков, В.И. Калёв

АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ МИНИМАКСНОГО ПРОГРАММНОГО ТЕРМИНАЛЬНОГО УПРАВЛЕНИЯ РАСХОДОМ ТОПЛИВА СРЕДСТВ ВЫВЕДЕНИЯ ПРИ НАЛИЧИИ ФАЗОВЫХ ОГРАНИЧЕНИЙ

Аннотация

Представлена формализация и алгоритм решения задачи минимаксного программного терминального управления расходом топлива жидкостной двигательной установки средств выведения при наличии фазовых ограничений. Для исходной нелинейной модели объекта управления формируется линеаризованная вдоль опорной траектории аппроксимирующая линейная дискретная динамическая система. Для аппроксимирующей системы формулируется задача минимаксного программного терминального управления с учетом заданных геометрических ограничений на реализации фазового вектора, управления и возмущения. Предлагается метод и численный алгоритм решения данной задачи. Для построения обобщенных областей достижимости аппроксимирующей системы используется модификация общего рекуррентного алгебраического метода. Предлагаемый численный алгоритм решения задачи учитывает все заданные ограничения и не зависит от размерности динамической системы. Эффективность предлагаемого подхода демонстрируется на содержательном модельном примере и подтверждается численными расчетами.

Ключевые слова: минимаксная оптимизация; нелинейный объект управления; фазовые ограничения; управление расходом топлива; средства выведения.

Введение

Задача терминального управления расходом топлива двигательных установок (ДУ) является одной из основных задач управления, решаемых для жидкостных средств выведения (СВ). Проблема заключается в синхронном и полном израсходовании компонентов топлива к заданному моменту времени. Тогда качество процесса терминального управления может характеризоваться отклонением основных координат состояния процесса от их желаемых значений в финальный момент времени. Кроме этого, важным аспектом в рассматриваемой задаче является гарантия безаварийных условий работы ДУ, т.е. нахождение фазового вектора объекта управления в каждый момент времени внутри некоторой допустимой (разрешенной) области фазового пространства.

В данной работе для исходной нелинейной модели объекта управления формируется ее линейная дискретная аппроксимация относительно заданного опорного режима функционирования, в которой учитывается возмущение, описывающее погрешность процесса аппроксимации и неконтролируемые внешние воздействия. Предполагается, что множества, ограничивающие реализации фазового вектора и вектора возмущения являются выпуклыми, замкнутыми и ограниченными многогранниками (с конечным числом вершин) в соответствующих конечномерных векторных пространствах, а множество значений допустимых управляющих воздействий в фиксированный момент времени является конечным набором векторов. Задача оптимизации режимов расхода топлива ДУ СВ формализуется с использованием минимаксного подхода [1,2], предполагающего нахождение такого управляющего воздействия, которое минимизирует наихудшие (максимальные) значения выбранного критерия качества, соответствующие возможным наихудшим реализациям допустимых значений возмущения и обеспечивает нахождение фазового вектора внутри заданных ограничений.

Предлагаемое решение поставленной задачи базируется на общем рекуррентном алгебраическом методе построения областей достижимости (ОД) линейных дискретных управляемых динамических систем [2], компьютерная реализация которого приводится и подробно описывается в работах [2–5].

В заключительной части данной работы представлены результаты компьютерного моделирования применения предлагаемого метода и численного алгоритма для решения задачи минимаксного программного терминального управления расходом топлива ДУ СВ на конкретном модельном примере.

1. Математическая модель объекта управления

На промежутке времени $[t_0, t_1]$ рассматривается нелинейная модель [6], описывающая режим работы жидкостной ДУ СВ (здесь t_0 и t_1 – моменты времени включения и выключения ДУ соответственно). Управляющее воздействие (управление) $u(\theta)$ реализуется с помощью смены положения дросселя, позволяющего изменять массовые расходы компонентов топлива из баков.

Текущие значения массовых расходов компонентов топлива из баков вычисляются посредством следующих алгебраических соотношений:

$$m_o(\theta) = \frac{(P + c_1 u(\theta)^2 + c_2 u(\theta))(K + c_5 u(\theta))}{(I + c_3 u(\theta)^2 + c_4 u(\theta))(1 + K + c_5 u(\theta))}, \quad (1)$$

$$m_f(\theta)=\frac{P+c_1u(\theta)^2+c_2u(\theta)}{(I+c_3u(\theta)^2+c_4u(\theta))(1+K+c_5u(\theta))},$$

где $\theta\in[t_0,t_1]$; P, I, K– расчетные значения силы тяги, удельного импульса тяги и коэффициента соотношения расходов компонентов топлива, соответственно; $c_1,c_2,\ldots,c_5$ – коэффициенты динамики работы ДУ.

Значения масс окислителя и горючего в баках вычисляются по формулам:

$$\begin{aligned} M_o(\theta)&=M_o^{init}-\int_{t_0}^{\theta}m_o(\theta)\,d\theta,\\ M_f(\theta)&=M_f^{init}-\int_{t_0}^{\theta}m_f(\theta)\,d\theta, \end{aligned} \qquad (2)$$

где M_o^{init}, M_f^{init} – начальные массы окислителя и горючего в топливных баках.

Исходная нелинейная модель объекта управления, описываемая системой соотношений (1), (2), линеаризуется вдоль опорной траектории вида:

$$\begin{cases} m_o^{ref}(\theta)=\dfrac{P\,K}{I+I\,K}, \\ m_f^{ref}(\theta)=\dfrac{P}{I+I\,K}, \end{cases} \quad \begin{cases} M_o^{ref}(\theta)=M_o^{init}-\dfrac{P\,K}{I+I\,K}\theta, \\ M_f^{ref}(\theta)=M_f^{init}-\dfrac{P}{I+I\,K}\theta, \end{cases} \qquad (3)$$

и далее дискретизируется с шагом ΔT. Более подробно с процедурой линеаризации и дискретизации уравнений (1), (2) можно ознакомиться в работе [6].

Рассмотрим на целочисленном промежутке времени $\overline{0,T}=\{0,1,\ldots,T\}$, $(T>0)$, поставленную в соответствие исходной нелинейной непрерывной модели линейную дискретную динамическую систему (здесь T – количество шагов, на которых будет реализовываться управление). Тогда значения массовых расходов компонентов топлива из баков вычисляются согласно следующим рекуррентным соотношениям:

$$\begin{aligned} &m_o(t+1)=m_o(t)+\alpha u(t)+\gamma_1 v_1(t),\ \ m_o(0)=m_o^{init},\\ &m_f(t+1)=m_f(t)+\beta u(t)+\gamma_2 v_2(t),\ \ m_f(0)=m_f^{init}, \end{aligned} \qquad (4)$$

где $t\in\overline{0,T-1}$; α,β – коэффициенты, полученные при линеаризации исходной модели; $u(t)$ – скалярное управление; m_o^{init}, m_f^{init} – начальные значения массовых расходов окислителя и горючего; $v_1(t), v_2(t)$ – погрешности формирования модели; γ_1,γ_2 – известные скалярные коэффициенты.

Дискретные рекуррентные уравнения, оценивающие значения масс окислителя и горючего в топливных баках, имеют следующий вид:

$$\begin{aligned} &M_o(t+1)=M_o(t)-\Delta T\,m_o(t),\ \ M_o(0)=M_o^{init},\\ &M_f(t+1)=M_f(t)-\Delta T\,m_f(t),\ \ M_f(0)=M_f^{init}, \end{aligned} \qquad (5)$$

где ΔT – шаг дискретизации.

Учитывая предыдущее, уравнения, описывающие динамику объекта управления (4), (5) могут быть записаны в следующем рекуррентном векторно-матричном виде:

$$\mathbf{x}(t+1)=A(t)\mathbf{x}(t)+B(t)\mathbf{u}(t)+D(t)\mathbf{v}(t),\ \ \mathbf{x}(0)=\mathbf{x}_0, \qquad (6)$$

где $\mathbf{x}(t)\in\mathbf{R}^4$ – фазовый вектор системы, стесненный заданным ограничением

$$\mathbf{x}(t)\in\mathbf{X}_1(t)\subset\mathbf{R}^4,\ t\in\overline{0,T}, \qquad (7)$$

формализующим условие безотказности ДУ (здесь и далее, $\mathbf{R}^n$ – n-мерное векторное пространство векторов-столбцов; $n\in\mathbf{N}$, $\mathbf{N}$ – множество всех натуральных чисел); $\mathbf{u}(t)\in\mathbf{R}$ – управление, стесненное заданным ограничением:

$$\mathbf{u}(t)\in\mathbf{U}_1(t)\subset\mathbf{R},\ t\in\overline{0,T-1}, \qquad (8)$$

где $\mathbf{U}_1(t)$ – конечный набор векторов в $\mathbf{R}$, определяющий все возможные значения управления в момент времени t; $(v_1(t),v_2(t))=\mathbf{v}(t)\in\mathbf{R}^2$ – неконтролируемое возмущение (погрешность аппроксимирующей модели), удовлетворяющее известному ограничению вида:

$$\mathbf{v}(t)\in\mathbf{V}_1(t)\subset\mathbf{R}^2,\ t\in\overline{0,T-1}; \qquad (9)$$

$A(t)\in\mathbf{R}^{4\times4}$ – матрица состояния системы;

$B(t)\in\mathbf{R}^{4\times1}$ – матрица управления;

$D(t)\in\mathbf{R}^{4\times2}$ – матрица возмущения.

Для ограничений (7), (9) выполняется следующее условие.

Условие 1. Множества $\mathbf{X}_1(t)\subset\mathbf{R}^4, t\in\overline{0,T}$, ограничивающие реализации фазового вектора системы, и множества $\mathbf{V}_1(t)\subset\mathbf{R}^2$, $\forall t\in\overline{0,T-1}$, ограничивающие реализации вектора возмущения, являются выпуклыми, замкнутыми и ограниченными многогранниками (с конечным числом вершин) в векторных пространствах $\mathbf{R}^4$ и $\mathbf{R}^2$ соответственно.

Для промежутка времени $\overline{0,T}$ и ограничения (7) введем в рассмотрение множество всех допустимых реализаций программных управлений

$$\mathbf{u}(\cdot)=\{\mathbf{u}(t)\}_{t\in\overline{0,T-1}},$$

которое описывается соотношением

$$\mathbf{U}(\overline{0,T})=\{\mathbf{u}(\cdot)\,|\,\mathbf{u}(\cdot)=\{\mathbf{u}(t)\}_{t\in\overline{0,T-1}}\in\mathbf{R}^{1\times T},\forall t\in\overline{0,T-1},\ \mathbf{u}(t)\in\mathbf{U}_1(t)\} \qquad (10)$$

и является конечным множеством. Аналогично для промежутка времени $\overline{0,T}$ и ограничения (9) введем в рассмотрение множество всех допустимых программных реализаций вектора возмущений (помех) $\mathbf{v}(\cdot)=\{v(t)\}_{t\in\overline{0,T-1}}$:

$$\mathbf{V}(\overline{0,T})=\{\mathbf{v}(\cdot)\,|\,\mathbf{v}(\cdot)=\{\mathbf{v}(t)\}_{t\in\overline{0,T-1}}\in\mathbf{R}^{2\times T},\forall t\in\overline{0,T-1},\ \mathbf{v}(t)\in\mathbf{V}_1(t)\}. \qquad (11)$$

Рассматриваемый процесс управления будем оценивать выпуклым терминальным функционалом

$\Phi: \mathbf{R}^4 \times \mathbf{U}(\overline{0,T}) \times \mathbf{V}(\overline{0,T}) \to \mathbf{R}^1$, определенным на возможных реализациях наборов $(\mathbf{x}_0, \mathbf{u}(\cdot), \mathbf{v}(\cdot)) \in \mathbf{R}^4 \times \mathbf{U}(\overline{0,T}) \times \mathbf{V}(\overline{0,T})$, значения которого определяются по формуле:

$$\Phi(\mathbf{x}_0, \mathbf{u}(\cdot), \mathbf{v}(\cdot)) = \left\| \mathbf{x}(T) - \mathbf{x}_d \right\|_4 = F(\mathbf{x}(T)),$$

(12)**Error! Bookmark not defined.**

где $\mathbf{x}(T) = \bar{\mathbf{x}}(T; \overline{0,T}, \mathbf{x}_0, \mathbf{u}(\cdot), \mathbf{v}(\cdot)) \in \mathbf{R}^4$ – финальное фазовое состояние фазовой траектории системы (6) на промежутке времени $\overline{0,T}$, соответствующее набору $(\mathbf{x}_0, \mathbf{u}(\cdot), \mathbf{v}(\cdot))$; $\mathbf{x}_d \in \mathbf{R}^4$ – вектор желаемого финального фазового состояния системы (6); функционал $F: \mathbf{R}^4 \to \mathbf{R}^1$; $\left\| \cdot \right\|_4$ – евклидова норма в пространстве $\mathbf{R}^4$.

2. Постановка задачи

Рассмотрим содержательную постановку задачу минимаксного программного терминального управления расходом топлива ДУ СВ.

Для дискретной динамической системы (6) с ограничениями (7)–(11) требуется найти допустимое программное управление $\mathbf{u}^{(e)}(\cdot) = \{\mathbf{u}^{(e)}(t)\}_{t \in \overline{0,T-1}} \in \mathbf{U}(\overline{0,T})$ такое, что гарантированный (наибольший) результат процесса управления, оцениваемый функционалом (12), будет минимальным среди результатов, возможных при любых допустимых управлениях $\mathbf{u}(\cdot) = \{\mathbf{u}(t)\}_{t \in \overline{0,T-1}} \in \mathbf{U}(\overline{0,T})$ и любых реализациях возмущений $\mathbf{v}(\cdot) = \{\mathbf{v}(t)\}_{t \in \overline{0,T-1}} \in \mathbf{V}(\overline{0,T})$, т.е. такое управление, при котором выполняется *условие минимакса*

$$\Phi(\mathbf{x}_0, \mathbf{u}^{(e)}(\cdot), \mathbf{v}(\cdot)) \le \max_{\mathbf{v}(\cdot) \in \mathbf{V}(\overline{0,T})} \Phi(\mathbf{x}_0, \mathbf{u}(\cdot), \mathbf{v}(\cdot)).$$

(13)**Error! Bookmark not defined.**

Тогда для линейной дискретной динамической системы (6) с ограничениями (7) – (11) можно сформулировать следующую задачу.

Задача 1. Для заданных целочисленного промежутка времени $\overline{0,T}$ и начального фазового вектора $\mathbf{x}(0) = \mathbf{x}_0$ требуется найти множество $\mathbf{U}^{(e)}(\overline{0,T}, \mathbf{x}(0)) \subseteq \mathbf{U}(\overline{0,T})$ допустимых программных управлений $\mathbf{u}^{(e)}(\cdot) = \{\mathbf{u}^{(e)}(t)\}_{t \in \overline{0,T-1}} \in \mathbf{U}(\overline{0,T})$, удовлетворяющих условию минимакса:

$$\mathbf{U}^{(e)}(\overline{0,T}, \mathbf{x}_0) = \{\mathbf{u}^{(e)}(\cdot) \mid \mathbf{u}^{(e)}(\cdot) = \{\mathbf{u}^{(e)}(t)\}_{t \in \overline{0,T-1}} \in \mathbf{U}(\overline{0,T}),$$
$$\Phi^{(e)} = \max_{\mathbf{v}(\cdot) \in \mathbf{V}(\overline{0,T})} \Phi(\mathbf{x}_0, \mathbf{u}^{(e)}(\cdot), \mathbf{v}(\cdot)) = \min_{\mathbf{u}(\cdot) \in \mathbf{U}(\overline{0,T})} \max_{\mathbf{v}(\cdot) \in \mathbf{V}(\overline{0,T})} \Phi(\mathbf{x}_0, \mathbf{u}(\cdot), \mathbf{v}(\cdot))\} \quad (14)$$

которое назовем множеством *минимаксных программных управлений* для данной задачи, а число $\Phi^{(e)}$ назовем ее *минимаксным результатом*.

В работах [2,4] было показано, что, так как множество $\mathbf{U}(\overline{0,T})$ конечно, решение многошаговой задачи 1 существует и сводится к реализации конечной последовательности нахождения решений только одношаговых задач линейного и выпуклого математического программирования, и операций над многогранными множествами [2,8].

3. Алгоритм решения задачи

В первую очередь рассмотрим понятие обобщенной ОД [2,3].

Определение 1. Обобщенной ОД фазового вектора линейной дискретной управляемой системы (6) с ограничениями (7)–(11) при фиксированном допустимом программном управлении $\mathbf{u}_*(\cdot) = \{\mathbf{u}_*(t)\}_{t \in \overline{0,T-1}} \in \mathbf{U}(\overline{0,T})$ на момент времени T, соответствующей набору $(\mathbf{x}(0), \mathbf{u}_*(\cdot))$, называется множество

$$\mathbf{G}(0, \mathbf{x}(0), \mathbf{u}_*(\cdot); T) = \{\mathbf{x}(T) \mid \mathbf{x}(T) \in \mathbf{R}^4, \quad (15)$$
$$\mathbf{x}(t+1) = A(t)\mathbf{x}(t) + B(t)\mathbf{u}_*(t) + D(t)\mathbf{v}(t) \in \mathbf{X}_1(t), t \in \overline{0,T-1}, \forall \mathbf{v}(t)$$

В работе [3] было показано, что множество, описывающее обобщенную ОД динамической системы, для которой выполняется условие 1, будет принадлежать к классу выпуклых замкнутых и ограниченных многогранников (с конечным числом вершин).

Алгоритм *общего рекуррентного алгебраического метода построения ОД* [2] заключается в решении последовательности одношаговых задач построения ОД $\mathbf{G}(t, \mathbf{X}(t), \mathbf{u}_*(\cdot); t+1)$, $\forall t \in \overline{1,T-1}$ (см. рисунок 1).

Инициализация:

$$\mathbf{X}(0) = \{\mathbf{x}_0\}, \quad \mathbf{u}_*(\cdot) = \{\mathbf{u}_*(t)\}_{t \in \overline{0,T-1}} \in \mathbf{U}(\overline{0,T}).$$

1: для всех $t \in \overline{0,T-1}$ выполнить

2: $\bar{\mathbf{X}}(t+1) = \left(A(t)\mathbf{X}(t) + B(t)\mathbf{u}_*(t)\right) \oplus D(t)\mathbf{V}_1(t);$

3: $\mathbf{X}(t+1) = \mathrm{ch}\left(\bar{\mathbf{X}}(t+1)\right);$

4: конец цикла.

Выходная информация:

$$\mathbf{G}\left(0, \mathbf{x}_0, \mathbf{u}_*(\cdot); t\right) = \mathbf{X}(t), \forall t \in \overline{1,T}.$$

Рис. 1. Алгоритм построения обобщенных ОД

Операция $\mathrm{ch}(\cdot)$ в алгоритме, приведенном на рисунке 1, обозначает решение задачи линейного математического программирования по нахождению крайних точек выпуклой оболочки множества. Операция суммирования множеств (по Минковскому) обозначается с помощью символа $\oplus$.

Далее, следуя [2–7], приведем предлагаемый метод решения задачи 1, заключающийся в решении следующей последовательности подзадач.

1. Упорядочение по возрастанию натурального индекса j конечного множества $\mathbf{U}(\overline{0,T})$, состоящего из N допустимых программных управлений $\mathbf{u}^{(j)}(\cdot)=\{\mathbf{u}^{(j)}(t)\}_{t\in\overline{0,T-1}}\in\mathbf{U}(\overline{0,T})$, т.е. формирование $\mathbf{U}(\overline{0,T})=\{\mathbf{u}^{(j)}(t)\}_{j\in\overline{1,N}}$.

2. Для всех $t\in\overline{1,T}$ и $j\in\overline{1,N}$ пошаговое построение ОД $\mathbf{G}(0,\mathbf{X}(0),\mathbf{u}^{(j)}(\cdot);t)$ при фиксированных управлениях $\mathbf{u}^{(j)}(\cdot)\in\mathbf{U}(\overline{0,T})$, где $\mathbf{X}(0)=\{\mathbf{x}_0\}$.

3. Формирование подмножества $\tilde{\mathbf{U}}(\overline{0,T})\subseteq\mathbf{U}(\overline{0,T})$ допустимых программных управлений $\tilde{\mathbf{u}}^{(j)}(\cdot)=\{\tilde{\mathbf{u}}^{(j)}(t)\}_{t\in\overline{0,T-1}}\in\tilde{\mathbf{U}}(\overline{0,T})$, не нарушающих фазовых ограничений, посредством удаления из множества $\mathbf{U}(\overline{0,T})$ тех допустимых программных управлений $\mathbf{u}^{(j)}(\cdot)\in\mathbf{U}(\overline{0,T})$, ОД которых нарушают фазовые ограничения (8) хотя бы для одного момента времени $t\in\overline{1,T}$ и упорядочения по возрастанию натурального индекса $i\in\overline{1,M}$ этого множества $\tilde{\mathbf{U}}(\overline{0,T})$.

4. Оптимизация для всех $i\in\overline{1,M}$ выпуклого функционала $\Phi(\mathbf{x}_0,\tilde{\mathbf{u}}^{(i)}(\cdot),\mathbf{v}(\cdot))=\left\|\mathbf{x}(T)-\mathbf{x}_{\mathrm{d}}\right\|_4$ на множестве $\mathbf{G}(0,\mathbf{X}(0),\tilde{\mathbf{u}}^{(i)}(\cdot);T)$ посредством решения следующей задачи выпуклого программирования:

$$\tilde{\Phi}^{(i)}=\Phi(\mathbf{x}_0,\tilde{\mathbf{u}}^{(i)}(\cdot),\mathbf{v}^{(e)}(\cdot))=\max_{\mathbf{v}(\cdot)\in\mathbf{V}(0,T)}\Phi(\mathbf{x}_0,\tilde{\mathbf{u}}^{(i)}(\cdot),\mathbf{v}(\cdot))=$$

$$=\max_{\mathbf{x}(T)\in\mathbf{G}(0,\mathbf{X}(0),\tilde{\mathbf{u}}^{(i)}(\cdot);T)}\left\|\mathbf{x}(T)-\mathbf{x}_{\mathrm{d}}\right\|_4=\left\|\mathbf{x}^{(e)}(T)-\mathbf{x}_{\mathrm{d}}\right\|_4,$$

где $\mathbf{x}(T)=\bar{\mathbf{x}}(T;\overline{0,T},\mathbf{x}_0,\tilde{\mathbf{u}}^{(i)}(\cdot),\mathbf{v}(\cdot))$;

$\mathbf{x}^{(e)}(T)=\bar{\mathbf{x}}(T;\overline{0,T},\mathbf{x}_0,\tilde{\mathbf{u}}^{(i)}(\cdot),\mathbf{v}^{(e)}(\cdot))$.

5. Нахождение множества $\mathbf{U}^{(e)}(\overline{0,T},\mathbf{x}(0))$ минимаксных программных управлений и числа $\Phi^{(e)}$ – оптимального гарантированного (минимаксного) результата решения задачи путем решения задачи дискретной оптимизации

$$\mathbf{U}^{(e)}(\overline{0,T},x_0)=\{\tilde{\mathbf{u}}^{(e)}(\cdot)\,|\,\tilde{\mathbf{u}}^{(e)}(\cdot)=\{\tilde{\mathbf{u}}^{(e)}(t)\}_{t\in\overline{0,T-1}}\in\tilde{\mathbf{U}}(\overline{0,T}),$$

$$\max_{\mathbf{v}(\cdot)\in\mathbf{V}(0,T)}\Phi(\mathbf{x}_0,\tilde{\mathbf{u}}^{(e)}(\cdot),\mathbf{v}(\cdot))=\min_{\tilde{\mathbf{u}}(\cdot)\in\tilde{\mathbf{U}}(0,T)}\max_{\mathbf{v}(\cdot)\in\mathbf{V}(0,T)}\Phi(\mathbf{x}_0,\tilde{\mathbf{u}}(\cdot),\mathbf{v}(\cdot))=\min_{i\in\overline{1,M}}\tilde{\Phi}^{(i)}=\Phi^{(e)}\}. \quad (17)$$

Из результатов работ [2–4] следует, что сформированное на основе (16), (17) множество допустимых программных управлений $\mathbf{U}^{(\omega)}(\overline{0,T},x_0)$ есть множество всех минимаксных программных управлений, являющихся решением задачи 1.

4. Численный пример

Продемонстрируем эффективность предлагаемого подхода на численном примере моделирования решения задачи минимаксного программного терминального управления расходом топлива ДУ РН [5]. Исходная нелинейная система описывается на промежутке времени $[0,100]$ уравнениями (1), (2), значения параметров которых указаны в таблице 1.

Таблица 1

Значения параметров нелинейной системы

I, с	P, кг	K	M_o^{init}, кг	M_o^{init}, кг	c_1	c_2	c_3	c_4	c_5
120	6000	3/2	300	200	3/50	–4500	7000	–200	125

В соответствие исходной нелинейной модели поставлена система векторно-матричных линейных рекуррентных соотношений, описывающая динамику системы, соответствующая уравнениям (4), (5) и имеющая вид:

$$\mathbf{x}(t+1)=A(t)\mathbf{x}(t)+B(t)\mathbf{u}(t)+D(t)\mathbf{v}(t),\ t\in\overline{0,4},$$

где $\mathbf{x}(t)=\{x_1(t),x_2(t),x_3(t),x_4(t)\}\in\mathbf{R}^4$; $x_1(t)$ – массовый расход окислителя; $x_2(t)$ – масса окислителя; $x_3(t)$ – массовый расход горючего; $x_4(t)$ – масса горючего; $\mathbf{u}(t)\in\mathbf{R}$ – скалярное управляющее воздействие, принимающее свои значения на множестве $\mathbf{U}_1(t)=\{1;0.3;0;-0.3;-1\}$, $\forall t\in\overline{0,4}$; $\mathbf{v}(t)\in\mathbf{R}^2$ – погрешность модели; $\mathbf{x}(0)=\mathbf{x}_0=(29.31,302,20.354,207)$.

Условия безаварийной работы ДУ диктуют фазовые ограничения

$$29\le x_1(t)\le 31,\qquad 19.5\le x_3(t)\le 20.5,\qquad \forall t\in\overline{1,5}.$$

Матрицы $A(t)$, $B(t)$ и $D(t)$ принимают значения

$$A(t)=\begin{pmatrix}1&0&0&0\\-10/T&1&0&0\\0&0&1&0\\0&0&-10/T&1\end{pmatrix},\ B(t)=\begin{pmatrix}0.7\\0\\-1/3\\0\end{pmatrix},\ D(t)=\begin{pmatrix}0.05&0\\0&0\\0&-0.04\\0&0\end{pmatrix},\ \forall t\in\overline{0,4}.$$

Качество процесса управления в данной задаче оценивается значением выпуклого терминального функционала в финальный момент времени $T = 5$

$$F(x(T)) = \sqrt{(30 - x_1(T))^2 + x_2(T)^2 + 2.25(20 - x_3(T))^2 + 2.25x_4(T)^2}\,.$$

В соответствии с предлагаемым методом решения задачи 1 для конечных множеств $\mathbf{U}_1(t)$, описывающих множества всех допустимых значений управления в моменты времени t, на основе (10) было сформировано упорядоченное множество $\overline{\mathbf{U}(0,T)} = \{\mathbf{u}^{(j)}(\cdot)\}_{j\in\overline{1,N}}$, состоящее из $N = 5^T = 3125$ допустимых программных управлений. Используя модификацию общего рекуррентного алгебраического метода построения ОД [7] для каждого допустимого программного управления $\mathbf{u}^{(j)}(\cdot) = \{\mathbf{u}^{(j)}(t)\}_{t\in\overline{0,4}}$ были вычислены соответствующие обобщенные ОД (15). Анализ областей достижимости вида (15) позволил сформировать упорядоченное множество программных управлений, удовлетворяющих ограничениям безаварийной работы ДУ, $\tilde{\mathbf{U}}(0,T) = \{\tilde{\mathbf{u}}^{(i)}(\cdot)\}_{i\in\overline{1,365}}$. Каждая из ОД, соответствующих программным управлениям из множества $\tilde{\mathbf{U}}(0,T)$, использовалась в качестве линейных ограничений (путем соответствующего преобразования) при решении задачи выпуклого математического программирования вида (16). В результате решения этой задачи максимизации был сформирован конечный набор числовых значений $\tilde{\Phi}^{(i)} = \Phi(x_0, \tilde{\mathbf{u}}^{(i)}(\cdot), \mathbf{v}^{(e)}(\cdot))$, $i \in \overline{1,365}$ – наихудших значений функционала качества, соответствующих набору допустимых программных управлений $\{\tilde{\mathbf{u}}^{(i)}(t)\}_{i\in\overline{1,365}}$. Вычисленное множество программных минимаксных управлений $\mathbf{U}^{(e)}(\overline{0,T}, x_0)$ состоит из управления $\mathbf{u}^{(e)}(\cdot) = \{\mathbf{u}^{(e)}(t)\}_{t\in\overline{0,4}} = \{1,0,0.3,-0.3,0\}$, гарантирующего результат решения задачи 1 не хуже $\Phi^{(e)} = 9.77$, т.е. обеспечивающего выполнение условия (17).

При использовании найденного программного минимаксного управления $\mathbf{u}^{(e)}(\cdot)$ в исходной нелинейной системе (1), (2) финальное фазовое состояние приняло значение $\mathbf{x}(T) = (30, 5.81, 20, 4.39)^{\mathrm{T}}$, в котором функционал (12), оценивающий качество процесса управления в исходной нелинейной системе, принял значение $F(\mathbf{x}(T))=8.82$. На рисунке 2 представлены проекции фазовых траекторий исходной системы (1), (2) и аппроксимирующей линейной дискретной системы (6), порожденные минимаксным программным управлением $\mathbf{u}^{(e)}(\cdot)$. Результаты решения задачи 1 сведены в таблицу 2.

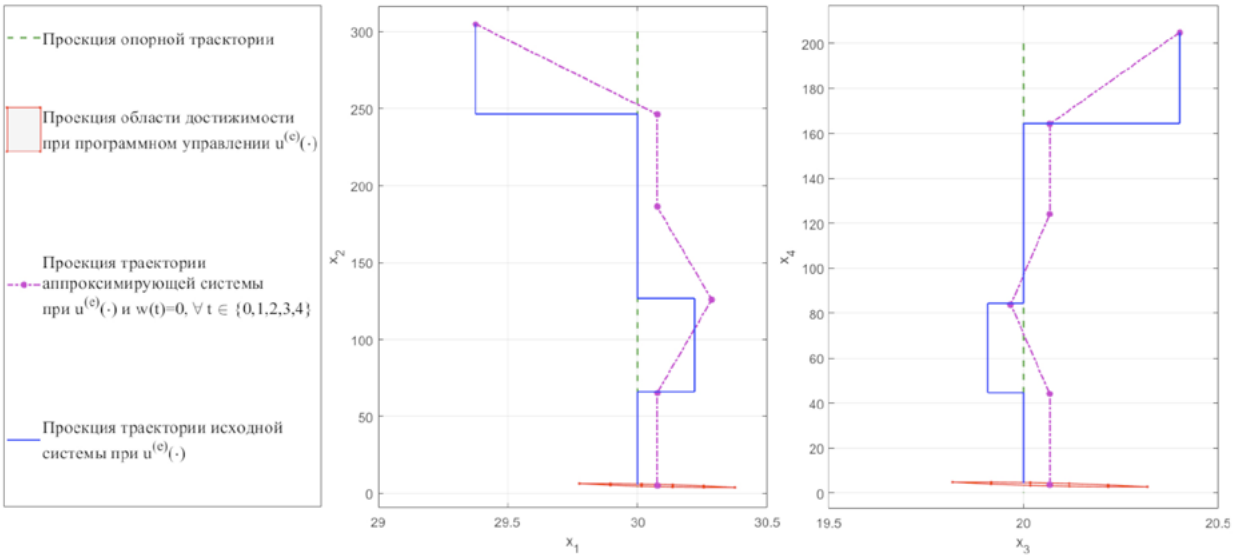


Рис. 2. Проекции фазовых траекторий для решения задачи 1

Таблица 2

Результаты решения задачи 1

	Управление	$x_1(T)$, кг/с	$x_2(T)$, кг	$x_3(T)$, кг/с	$x_4(T)$, кг	$\Phi(x(T))$
Опорная траектория	$u(t) \equiv 0$	30	0	20	0	**0**
Нелинейная система	$\mathbf{u}^{(e)}(\cdot)$	30	5.81	20	4.39	**8.82**
Линейная система	$\mathbf{u}^{(e)}(\cdot)$	29.78	6.43	19.82	4.87	**9.77**

Численное моделирование проводилось с использованием созданного авторами специализированного компьютерного программного комплекса в программной среде MATLABR2014a с процессором Intel© Core™ i7-3770 CPU @ 3.4 GHz, оперативной памятью 8 Gb. Время решения задачи составило 66 с.

Выводы

В работе приведено описание нелинейной динамической системы, описывающей динамику расхода топлива ДУ СВ. Задача оптимизации управления расходом топлива ДУ СВ формализована как задача минимаксного программного терминального управления линейной дискретной динамической системой с выпуклым функционалом качества и подробно описан предлагаемый метод решения этой задачи.

Сформулированная многошаговая задача минимаксного программного терминального управления расходом топлива ДУ СВ решается путем реализации конечной последовательности только одношаговых задач выполнения алгебраических операций, линейного и выпуклого математического программирования, а также операций над многогранными множествами, на основе общего рекуррентного алгебраического метода построения областей достижимости линейных дискретных управляемых динамических систем.

На основе описанного метода решения задачи минимаксного программного терминального управления были разработаны численный алгоритм и компьютерная программная система для моделирования ее решения и получены результаты численного моделирования. Компьютерная реализациметода ограничена только ресурсами памяти и быстродействием используемой компьютерной платформы, в то время как сам метод применим для линейных дискретных управляемых динамических систем любой конечной размерности.

Результаты численного моделирования показали эффективность предлагаемого метода для решения рассматриваемой задачи минимаксного программного терминального управления расходом топлива ДУ СВ и могут быть использованы в дальнейшем для решения аналогичной задачи при наличии обратной связи (формирования процедуры адаптивного управления).

Список литературы

1. Н.Н. Красовский, А.И. Субботин. Позиционные дифференциальные игры. Москва: Наука, 1974.

2. А.Ф. Шориков. Минимаксное оценивание и управление в дискретных динамических системах. Екатеринбург: Изд-во Урал.ун-та, 1997. 242 с.

3. В.А. Тюлюкин, А.Ф. Шориков. Об одном алгоритме построения области достижимости линейной управляемой системы // Негладкие задачи оптимизации и управление. Свердловск: УрО АН СССР, 1988. С. 55–61.

4. В.А. Тюлюкин, А.Ф. Шориков. Алгоритм решения задачи терминального управления для линейной дискретной системы // Автоматика и телемеханика, 1993. № 4. С. 115–127.

5. А.Ф. Шориков, В.А. Тюлюкин. Описание библиотеки компьютерных программ для моделирования решения задачи апостериорного минимаксного оценивания // Известия Уральского гос. экон. ун-та.1999, № 2. С. 36–49.

6. А.Ф. Шориков, В.И. Калёв. Формирование линейной дискретной динамической модели для решения задачи оптимального терминального управления расходом топлива ракеты-носителя // Информационные технологии и системы: Труды Пятой Междунар. науч. конф. Научное электронное издание – Челябинск: Изд-во Челяб. гос. ун-та, 2016. С. 61–65.

7. А.Ф. Шориков, В.В. Булаев, А.Ю. Горанов, В.И. Калев. Аппроксимация областей достижимости нелинейных дискретных управляемых динамических систем // Вестник Бурятского государственного университета. Математика, информатика, 2018. № 1. С. 52–65.

8. K. Fukuda, A. Prodon Double description method revisited // Lecture Notes in Computer Science, 1996. Vol. 1120. pp. 91–111.

A.F. Shorikov, V.I. Kalev

SOLUTION ALGORITHM FOR PROPELLANT CONSUMPTION MINIMAX OPEN-LOOP TERMINAL CONTROL PROBLEM OF LIQUID-PROPELLANT LAUNCH VEHICLE

This paper presents the formalization and solution algorithm of the propellant consumption minimax open-loop terminal control problem for launch vehicle's propulsion system in the presence of state constraint. Initial nonlinear model of plant is linearized along the reference trajectory and approximated by linear discrete-time dynamical system. Using the approximating system, we formalize the minimax open-loop terminal control problem of process under consideration subject to given geometrical constraints of state, control and disturbance vectors. To solve concerned problem new approach and numerical algorithm are proposed, which use the modification of general recurrence algebraic method to compute generalized reachable sets of linear discrete-time controlled system. The proposed solution algorithm considers all given constraints and has no dependency of dynamical system dimension. The efficiency of proposed approach is shown using informative numerical simulation example and validated by numerical calculations.

УДК 624.072.223

А.В. Юлдашева

О РАЗРЕШИМОСТИ ОДНОЙ НЕКОРРЕКТНОЙ ЗАДАЧИ ДЛЯ УРАВНЕНИЯ КОЛЕБАНИЯ БАЛКИ

Аннотация

В работе рассматривается вопрос о существовании L_2-решения краевой задачи для уравнения колебания балки в прямоугольной области. Задача вытекает из физической задачи о восстановлении колебания балки по изображению этой балки в какой-то момент времени. То есть, если мы обозначим через $u(x, t)$ сдвиг балки в момент времени t в точке x от положения равновесия и предположим, что у мы знаем положение балки в моменты времени $t = 0$ и $t = T$. То возможно ли найти $u(x, t)$ для всех $t > 0$? Предполагается, что если функция внешнего воздействия принадлежит некоторому классу Никольского, то L_2-решение рассматриваемой задачи существует и единственно при определенных условиях на размеры области, в которой рассматривается задача. О необходимости дополнительных условий для существования классического решения волнового уравнения впервые отметили Д. Боржин и Р. Даффин, для существования же L_2-решения аналогичное условие было найдено Ш.А.Алимовым. В работе также найдено условие на размеры области, при которых L_2- решение существует.

Ключевые слова: уравнение колебание балки, L_2-решение, класс Никольского, метод Фурье.

Ведение

В данной работе рассматривается вопрос о существовании L_2-решения краевой задачи для уравнения колебания балки в прямоугольной области. При этом используется определение L_2-решения краевой задачи, введенное в работе [1].

Обозначим через $u(x, t)$ сдвиг балки в момент времени t в точке x от положения равновесия. Хорошо известно уравнение свободного колебания балки

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^4 u}{\partial x^4} = f(x,t).$$

Если предположить, что у нас есть изображение балки в моменты времени $t = 0$ и $t = T$. То возможно ли найти $u(x, t)$ для всех $t > 0$?

О необходимости дополнительных условий для существования классического решения волнового уравнения впервые отметили Боржин Д. и Даффин Р. [2], Березанский Ю.М. [3, 4]. Аналогичные условия для уравнений четного высокого порядка были найдены в работах [5–8]. Для существования же L_2-решения аналогичное условие было найдено Ш.А. Алимовым [9, 10].

Краевые задачи для уравнения колебания балки хорошо изучены. Работы [11–18] лишь немногие из них, в которых доказывалась регулярная разрешимость и изучался спектр задач. Существование же L_2-решения краевой задачи для уравнения колебания балки, на сколько нам известно, изучается впервые.

Итак, в области

$D = \{(x,t) \in \mathbb{R}^2 : 0 < x < a, 0 < t < T\}$ рассмотрим следующую задачу

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^4 u}{\partial x^4} = f(x,t),\ (x,t) \in D \tag{1}$$

$$u(0,t) = u(a,t) = u_{xx}(0,t) = u_{xx}(a,t) = 0,\ 0 \le t \le T, \tag{2}$$

$$u(x,0) = u(x,T) = 0,\ 0 \le x \le a. \tag{3}$$

Введем для любого $\delta > 0$

$$\Gamma_\delta = \{(x,t) : dist((x,t), \partial D) \le \delta\}. \tag{4}$$

Будем говорить, что $u(x,t) \in L_p^0(D), p \ge 1$, если $u(x,t) \in L_p(D)$ и выполняется соотношение

$$\int\limits_{\Gamma_\delta} |u(x,t)|^p \, dx\, dt = o(\delta), \delta \to 0. \tag{5}$$

Ниже доказывается, что если $\theta = \frac{\pi T}{a^2}$ является алгебраическим (иррациональным) числом и правая часть уравнения (1) принадлежит некоторому классу Никольского, то L_p-решение задачи (1)–(3) существует.

Краевые задачи для уравнений

Прежде чем сформулировать соответствующий результат, напомним определение анизотропных классов Никольского [19].

Для функции двух переменных $f(x,t), (x,t) \in \mathbb{R}^2$, введем разность

$$\Delta_{hs}^3 f(x,t) = f(x+2h, t+s) - 2f(x+h, t+s) - f(x+2h, t) + 2f(x+h, t) + f(x, t+s) - f(x,t) \tag{6}$$

Пусть даны числа $p \ge 1, 0 \le \alpha < 1, 0 \le \beta < 1$. Функция $f \in L_p(\mathbb{R}^2)$ принадлежит классу Никольского $H_p^{\alpha,\beta}(\mathbb{R}^2)$, если выполняется условие

$$\int\limits_{\mathbb{R}^2} \left|\Delta_{hs}^3 f(x,t)\right|^p dx\, dt \le C^p |h|^{\alpha p} |s|^{\beta p}. \tag{7}$$

Для области $D \subset \mathbb{R}^2$ символом $H_p^{0,\alpha,\beta}\left(\mathbb{R}^2\right)$ будем обозначать класс функций, получаемый замыканием $C_0^\infty\left(D\right)$ по норме пространства $H_p^{\alpha,\beta}\left(\mathbb{R}^2\right)$.

Справедливо следующее утверждение.

Теорема. Пусть $\theta = \dfrac{\pi T}{a^2}$ является алгебраическим числом степени 2 и пусть

$$0 \le \alpha < 1,\ 0 \le \beta < 1,\ 2\alpha + \beta \ge \frac{1}{2}. \qquad (8)$$

Тогда для любой функции $f \in H_p^{0,\alpha,\beta}\left(\mathbb{R}^2\right)$ задача (1)–(3) имеет и при том единственное L_p- решение.

Доказательство теоремы 1 основано на поиске решения рассматриваемой
задачи методом Фурье.

Разложим правую часть уравнения (1) в ряд Фурье:

$$f(x,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} f_{nm} \sin\frac{\pi n x}{a}\sin\frac{\pi m t}{T},\ (x,t) \in D. \qquad (9)$$

Введем обозначение

$$\gamma_{nm} = \left[\left(\frac{\pi n}{a}\right)^4 - \left(\frac{\pi m}{T}\right)^2\right]^{-1} \qquad (10)$$

и докажем, что функция

$$u(x,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \gamma_{nm}^{-1} \sin\frac{\pi n x}{a}\sin\frac{\pi m t}{T},\ (x,t) \in D. \qquad (11)$$

является искомым решением. Предварительно докажем несколько лемм.

Оценка коэффициентов Фурье

Заметим, что в равенстве (9) левая часть равна нулю вне *D*, а правая часть периодична по *x* с периодом *2a* и по *t* с периодом *2T*. Следовательно, равенство (9) имеет место только в D.

Введем *2a*-периодическую по *x* и *2T*-периодическую по *t* функцию $\tilde{f} = \tilde{f}(x,t)$, ряд Фурье которой совпадает с суммой (9)

$$\tilde{f}(x,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} f_{nm} \sin\frac{\pi n x}{a}\sin\frac{\pi m t}{T},\ (x,t) \in \mathbb{R}^2. \qquad (12)$$

Тогда

$$f(x,t) = \begin{cases} \tilde{f}\left(x,t\right), \left(x,t\right) \in D, \\ 0, \left(x,t\right) \notin D. \end{cases}. \qquad (13)$$

Для функции (12) введем следующую форму:

$$\left\|\tilde{f}\right\|_{H_p^{*r}} = \left\|\tilde{f}\right\|_{L_p(D)} +$$

$$+ \sup_{h>0,s>0}\left(h^4+s^2\right)^{-r/2}\left[\int_D \left|\tilde{f}(x+2h,t+s) - \tilde{f}(x,t)\right|^p dx\,dt\right]^{1/p} \qquad (14)$$

Лемма 1. Пусть $0 < r < 1, p \ge 1$. Тогда для любого $\delta > 0$ выполняется неравенство

$$\left[\int_{\Gamma_\delta} \left|\tilde{f}(x,t)\right|^p dx\,dt\right]^{1/p} \le \mathrm{const}\,\delta^r \left\|\tilde{f}\right\|_{H_p^{*r}}, \qquad (15)$$

где множество Γ_δ определено равенством (4).

Доказательство. Заметим, что множество Γ_δ разбивается на четыре прямоугольника, а также в силу нечетности функции $\tilde{f}$, достаточно доказать верность неравенства

$$\int_{\Gamma_\delta} \left|\tilde{f}(x,t) - \tilde{f}(-x,t)\right|^p dx\,dt \le \mathrm{const}\,\delta^{rp} \left\|\tilde{f}\right\|^p_{H_p^{*r}}. \qquad (16)$$

Далее согласно [4] функция $\tilde{f}(x,t) \in \mathrm{H}_p^{*r}$ имеет представление:

$$\tilde{f}(x,t) = \sum_{k=1}^{\infty} T_k(x,t), \qquad (17)$$

где $T_k(x,t)$ – тригонометрические полиномы степени 2^k, которые удовлетворяют неравенству

$$\left\|T_k(x,t)\right\|_{L_p} \le \mathrm{const}\,\delta^{rp} \left\|\tilde{f}\right\|^p_{H_p^{*r}}. \qquad (18)$$

Так как

$$\left|T_k(\mathrm{x,t}) - T_k(-\mathrm{x,t})\right|^p \le (2x)^{p-1}\int_{-x}^{x}\left|T_k'(y,\mathrm{t})\right|^p dy \le (2\delta)^{p-1}\int_{-a}^{a}\left|T_k'(y,\mathrm{t})\right|^p dy,$$

то при помощи неравенства Бернштейна и неравенства (18) получаем

$$\left[\int_{-T}^{T} dt \int_{-x}^{x}\left|T_k(x,t) - T_k(-x,t)\right|^p dx\right]^{1/p} \le 2\min\{1,\delta 2^k\}\left\|T_k\right\|_{L_p} \le const\, 2^{-rk}\min\{1,\delta 2^k\}\left\|\tilde{f}\right\|_{H_p^{*r}}$$

Отсюда следует справедливость неравенства (16), а значит и утверждения леммы.

Лемма 2. Пусть $0 < r < 1, p \ge 1$. Если периодическая функция $\tilde{f}$ имеет конечную норму (14), то функция f, принадлежит классу $H_p^r(\mathbb{R}^2)$ и удовлетворяет оценке

$$\left\|f\right\|_{H_p^r\left(\mathbb{R}^2\right)} \le \mathrm{const}\,\delta^{rp} \left\|\tilde{f}\right\|^p_{H_p^{*r}}. \qquad (19)$$

Доказательство леммы следует из леммы 1.

Для любой функции $f \in L_2(D)$, ряд Фурье которой имеет вид (9), и любого $r > 0$ введем величину

$$A_r[f] = \sup_{R>0} R^r \left(\sum_{R \le \mu_{nm} < 2R} |f_{nm}|^2 \right)^{1/2}, \quad (20)$$

где

$$\mu_{nm} = \left[\left(\frac{\pi n}{a}\right)^4 + \left(\frac{\pi m}{T}\right)^2 \right]^{1/2}. \quad (21)$$

Лемма 3. Пусть $0 < r < 1$. Если для функции $f \in L_2(D)$ величина (20) конечна, тогда эта же функция, продолжаемая нулем вне D, принадлежит пространству $H_p^r(\mathbb{R}^2)$ и удовлетворяет неравенству

$$\|f\|_{H_p^r(\mathbb{R}^2)} \le \text{const} A_r[f]. \quad (22)$$

Доказательство леммы следует из определения нормы и элементарных тригонометрических преобразований.

Лемма 4. Пусть $0 < r < 1$. Если коэффициенты Фурье функции $f \in L_2(D)$

удовлетворяют условию

$$\sum_{R \le \mu_{nm} < 2R} |f_{nm}|^2 = o(1) R^{-2r}, \ R \to \infty, \quad (23)$$

то функция f продолженная 0 вне D, принадлежит классу $H_2^{0,r}(\mathbb{R}^2)$.

Доказательство. Заметим, что если в разложении (10) функции f отлично от нуля только конечное число слагаемых, то функция f принадлежит классу $H_2^{0,r}(D)$.

Введем последовательность линейных операторов B_k:

$$B_k f(x,t) = \sum_{\mu_{nm} < k} f_{nm} \sin\frac{\pi n x}{a} \sin\frac{\pi m t}{T}, \ (x,t) \in \mathrm{D},$$

$$B_k f(x,t) = 0, \ (x,t) \notin \mathrm{D}.$$

Следовательно, $B_k f(x,t) \in H_2^{0,r}(D)$ и осталось убедиться, что их последовательность сходится по норме $H_2^{0,r}(\mathbb{R}^2)$ к $f(x,t)$. А из леммы 3, мы имеем

$$\|f - B_k f\|_{H_2^r(\mathbb{R}^2)} \le A_r[f - B_k f],$$

откуда в силу (23) вытекает, что правая часть стремится к 0 при $k \to \infty$.

Лемма 4. При $|h| < a, |s| < T$ выполняется неравенство

$$\int_D \left|\Delta_{hs}^3 f(x,t)\right|^2 dx\,dt \le const \int_{\mathbb{R}^2} \left|\Delta_{hs}^3 f(x,t)\right|^2 dx\,dt$$

Верность этой оценки следует из нечетной периодичности функции $\tilde{f}$

и того, что $f(x,t) = 0$ почти всюду вне D.

Лемма 6. Пусть $0 \le \alpha < 1, 0 \le \beta < 1, \ p > 1$. Если $f \in H_p^{0,\alpha,\beta}(D)$, то при $h \to 0$ и $s \to 0$

выполняется соотношение

$$\int_{\mathbb{R}^2} \left|\Delta_{hs}^3 f(x,t)\right|^2 dx\,dt \le C^p |h|^{2\alpha p} |s|^{\beta p}. \quad (24)$$

Доказательство проводится в полной аналоги с доказательством предложения 3 работы [1].

Лемма 7. Пусть f разлагается в ряд Фурье (9) и пусть $0 \le \alpha < 1, 0 \le \beta < 1$.

Если $f \in H_p^{0,\alpha,\beta}(D)$, то при $\lambda \to \infty$ и $\mu \to \infty$ выполняется соотношение

$$\sum_{n<\lambda} \sum_{m<\mu} |f_{nm}|^2 n^4 m^2 = o(1) \lambda^{4(1-\alpha)} \mu^{2(1-\beta)}. \quad (25)$$

Доказательство. Из (12) следует

$$\int_{-T}^{T} \int_{-a}^{a} \left|\Delta_{hs}^3 f(x,t)\right|^2 dx\,dt = const \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |f_{nm}|^2 \sin^4 \frac{\pi n h}{2a} \sin^2 \frac{\pi m s}{2T}.$$

Поскольку $\sin x > \frac{2x}{\pi}$ при $0 < x < \frac{\pi}{2}$, то выбирая $h = \pm a/\lambda$, $s = \pm T/\mu$ из

предыдущего равенства получаем

$$\lambda^{-4}\mu^{-2} \sum_{n<\lambda} \sum_{m<\mu} |f_{nm}|^2 n^4 m^2 \le const \int_D \left|\Delta_{hs}^3 \tilde{f}(x,t)\right|^2 dx\,dt = o(1)|h|^{4\alpha}|s|^{2\beta} = o(1)\lambda^{-4\alpha}\mu^{-2\beta}$$

Лемма 7 доказана.

Следующая лемма касается оценки множителя γ_{nm}, определенного равенством (10).

Лемма 8. Пусть $\theta = \frac{\pi T}{a^2}$ является алгебраическим числом степени 2, тогда справедлива оценка

$$|\gamma_{nm}| \le const \frac{n^2 m}{\mu_{nm}^2}. \quad (26)$$

Верность оценки следует из теоремы Лиувилля для алгебраических чисел степени 2 ([20]) и элементарных преобразований.

Доказательство теоремы

Введем оператор $\Re$, который действует следующим образом: если разложение функции $f(x,t)$ в ряд Фурье имеет вид (9), то

$$\Re f(x,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \gamma_{nm}^{-1} \sin\frac{\pi n x}{a}\sin\frac{\pi m t}{T}, (x,t)\in D, \qquad (27)$$

где множители γ_{nm} определяются равенством (10).

Пусть $f(x,t)$ функция из класса Никольского $H_p^{0,\alpha,\beta}\left(\mathbb{R}^2\right)$ с показателями гладкости α и β, удовлетворяющими условиям (8). Докажем, что $\Re f(x,t)$ является L_2-решением рассматриваемой краевой задачи (1)–(3).

Убедимся вначале в том, что $\Re f(x,t)$ принадлежит классу $L_2^0(D)$. Согласно предложению 5 работы [1], для этого достаточно доказать, что $\Re f(x,t)$ принадлежит пространству $H_2^{0,r}(D)$ при $r = 1/2$, а для этого, согласно лемме 4, достаточно доказать, что при $R \to \infty$ выполняется оценка

$$\sum_{R\le\mu_{nm}<2R} \left|f_{nm}\right|^2 \gamma_{nm}^2 = o(1)R^{-1} \qquad (28)$$

Воспользуемся леммой 8, а также леммой 7 положив $\lambda = R\frac{2a}{\pi}$, $\mu = R\frac{2T}{\pi}$:

$$\sum_{R\le\mu_{nm}<2R} \left|f_{nm}\right|^2 \gamma_{nm}^2 \le R^{-4}\sum_{n<\lambda}\sum_{m<\mu}\left|f_{nm}\right|^2 n^4 m^2 = o(1)R^{-2(2\alpha+\beta)}$$

Докажем теперь, что $\Re f(x,t)$ является обобщенным решением уравнения (1).Разложим для этого произвольную функцию $\varphi(x,t)\in C_0^{\infty}(D)$ в ряд Фурье

$$\varphi(x,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \varphi_{nm}\sin\frac{\pi n x}{a}\sin\frac{\pi m t}{T}, (x,t)\in D$$

Действуя на эту функцию оператором $L = \frac{\partial^2}{\partial t^2} + \frac{\partial^4}{\partial x^4}$, получаем

$$L\varphi(x,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \gamma_{nm}^{-1}\varphi_{nm}\sin\frac{\pi n x}{a}\sin\frac{\pi m t}{T},$$

Далее применим равенство Парсеваля

$$\int_D \Re f(x,t) L\varphi(x,t)\,dx\,dt = \frac{aT}{4}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} f_{nm}\varphi_{nm} = \int_d f(x,t)\varphi(x,t)\,dx\,dt.$$

Ввиду произвольности $\varphi(x,t)\in C_0^{\infty}(D)$, это равенство и означает, что функция $\Re f(x,t)$ является обобщенным решением уравнения (1). Теорема доказана.

Выводы

В работе рассмотрена проблема о существовании решения краевой задачи для уравнения колебания балки в прямоугольной области, доказана теорема о единственном решении задачи и найдено условие на размеры области, при которых решение существует.

Список литературы

1. Алимов Ш.А. 1999. Об L_p-решениях одной краевой задачи. Узбекский математический журнал. 1: 3–9.

2. Bourgin D.G., Duffin R. 1939. The Dirichlet problem for vibrating string equation. Bull. Amer. Math. Soc. 45(12): 851–858.

3. Березанский Ю.М. 1960. О задаче типа Дирихле для уравнения колебания струны. Украин. мат. журн. Киев. 4 (12): 363-372.

4. Березанский Ю.М. Разложение по собственным функциям уравнений в частных разностях второго порядка. Тр. ММО, 5, ГИТТЛ, М., 1956, 203–268.

5. Сабитов К.Б. Задача Дирихле для двумерных уравнений в частных производных высших порядков. Уравнения смешанного типа и родственные проблемы анализа и информатики, Материалы международного Российско-Болгарского симпозиума (Нальчик–Хабез), Изд-во НИИ ПМА КБНЦ РАН, Нальчик–Хабез, 2010, 212–214.

6. Сабитов К.Б. Задача Дирихле для уравнений с частными производными высоких порядков и проблема Ферма. Труды математического центра им. Н.И. Лобачевского, Изд-во Казанского математического общества, Казань, 2010, 98–102.

7. Юлдашева А.В. 2015. Об устойчивости краевой задачи для уравнения четного порядка. Вестник КРАУНЦ, физ.-мат. науки. 1: 5–11.

8. Юлдашева А.В. 2016. О разрешимости краевой задачи для уравнения четного порядка. Вестник КРАУНЦ, физ.-мат.науки, 1(12): 16–22.

9. Алимов Ш.А. 1999. О разрешимости одной некорректной задачи. Узбекский математический журнал. 3: 19-29.

10. Алимов Ш.А. 1999. Об L_p-решениях одной некорректной краевой задачи. Докл. АНРУз-Ташкент. 9: 7–11.

11. Александрян Р.А. 1950. Об одной задаче Соболева для специального уравнения с частными производными четвёртого порядка. Докл. АН СССР. – Москва. 4 (43): 631–634.

12. Amanov Dj., Yuldasheva A.V., 2009. Solvability and spectral properties of boundary valueproblems for equations of even order, International Journal of Computer Mathematics, 38: 61–70.

13. Otarova J.A. Volterraness of boundary value problem for fourth order equation.// Fen Edebiyat Dergisi, 2007 yil, Volume 9 SayiVek (Sakarya University Faculty of Arts and Science) The Journal of Arts and Science , P 152–162.

14. Аманов Д., Юлдашева А.В. 2007. Разрешимость и спектральные свойства самосопряжённой задачи для уравнения четвёртого порядка. Узб. мат. журн. Ташкент.4: 3–8.

15. Аманов Д., Юлдашева А.В. 2009. Сильная разрешимость полупериодической самосопряженной задачи для уравнения четвертого порядка. Узб. мат. журн. – Ташкент. 2: 3–8.

16. Аманов Д. Вольтерровость краевой задачи для уравнения четвертого порядка. Спектральная теория дифференциальных операторов и родственные проблемы: Тр. межд. конф. 24–28 003.– Т.1. – Стерлитамак, 2003. – С. 78–82.

17. Аманов Д., Юлдашева А.В. Самосопряжённая задача для уравнения четвёртого порядка. СамДиф-2007. Дифференциальные уравнения и их приложения: Тр. конф. 29 января – 4 февраля 2007. – Самара, 2007. – С. 19–21.

18. Отарова Ж.А. 2008. Вольтеррова краевая задача для уравнения четвертого порядка. Докл. АНРУз-Ташкент. 6: 18–22.

19. Никольский С.М. 1977. Приближение функций многих переменных и теоремы вложения. Москва: Наука.

20. Шидловсский А.Б. 1987. Трансцендентные числа. Москва: Наука.

УДК 372.862

B. Blessing

LOGISTICS EDUCATION

Vorarlberg University of Applied Sciences

Vorarlberg University is a public post-secondary university maintained by the State of Vorarlberg. The employees create and convey knowledge and skills through degree programmes, research and continuing education. The University has a total of 16 degree programmes in the areas of business, engineering and technology, design and the social sciences. Nine of these degree programmes are at the bachelor's level and seven at the master's level; of these, seven are offered as professional degree programmes. Class sizes are small, and instruction is efficient and practice-oriented. The high level of didactic skills of instructors guarantees a universal quality of instruction throughout the campus. These quality standards are supported by state-of-the-art laboratories, in which students and staff work on current research topics. The immediate connection between research and studies means that students come into contact with current research questions early on.

At Vorarlberg University of Applied Sciences, scientists and students conduct research in the areas of Process and Product Engineering, Micro technology, User-Centred Technologies, Society and Socio-Economics and Mechatronics. The international focus of degree programmes and research provides students and staff with valuable experience abroad. [http://www.fhv.at/fhv-organisation/organisation-fh-vorarlberg, 22.01.2013]

Bachelor of Engineering and Management

The degree programme for a Bachelor of Engineering and Management at the FHV is an interdisciplinary simultaneous degree programme with approx. 60 % business and approx. 40 % technical content. The focus is on employed applicants with initial technical training and relevant professional experience who want to obtain qualifications at the interface of business and technology on an academic level.

After 3 years (6 semesters), students will graduate with an academic degree, which gives you the greatest possible flexibility for later decisions: You can have a career in your field, but you are also entitled to pursue a postgraduate master's degree. And you can do this not only at our university, but at any university or FH that has implemented a degree system in accordance with the "Bologna" system (bachelor and master).

The successful graduates of the programme in Engineering and Management can be found in almost all areas of the corporate sphere, but it is possible to identify preferred areas.

Graduates from Engineers and Management are frequently used in logistics, marketing, controlling and consulting. A large number of engineers with business administration also receive positions in corporate management, where both technological knowledge and business thinking are required. For this reason, there are many possible fields of work in various professions after graduating with a bachelor's degree. [http://www.fhv.at/fhv-studies/technology/ engineering-business-administration/ worth-knowing/mechatronics-at-vorarlberg-university-of-appliedsciences, 22.01.2013].

The degree programme for a Bachelor of Engineering and Management is divided in 5 divisions (see FIGURE 1).

Division 1: Economics

Division 2: Social Skills

Division 3: IT

Division 4: Engineering

Division 5: English

The divisions cover all 3 years. In the 3rd year the integration of the content of all divisions is the main focus. Students have to work on case studies, projects and topics which need perspectives from all divisions.

Logistics Education

Logistics is part of Division 1 and being taught in the 2nd and 3rd year. For the 3rd year students have to decide either for Marketing or for Logistics/Supply Chain Management. About 60 % of students select Logistics/ Supply Chain management. Logistics1, 3rd semester: Topics are Procurement, Operations Management and Warehousing. Didactics: Problem Based Learning, Simulations

Logistics2, 4th semester: Topics are Lean Manufacturing and KAIZEN. Didactics: Simulation and Gaming, Problem Based Learning.

Supply Chain Management, 5th semester: Topics are Supply Chain management and Warehousing. Didactics: Case studies and lectures.

Supply Chain Management, 6th semester: Topics are Global distributions; Transfer and integration of concepts. Didactics: Simulations, lectures, case studies.

Some of the lessons in Logistics and other subjects are being taught in English language by guest lectures from other countries (e.g. Finland, Netherlands)

Didactics

We are proud to have a great bunch of different didactics in teaching logistics. Front lecturing has a percentage of not more than 20 %. Main focus is on group work and on self-studying. Students have to learn how to learn outside the classroom in small groups.

a) Simulation & Gaming

Games and simulations are powerful tools. They are based on learning from experiences. [Jones, Ken: Games and Simulations Made Easy, Kogan Page, 1997, 1 p.]

We are using about 6 different types of simulations

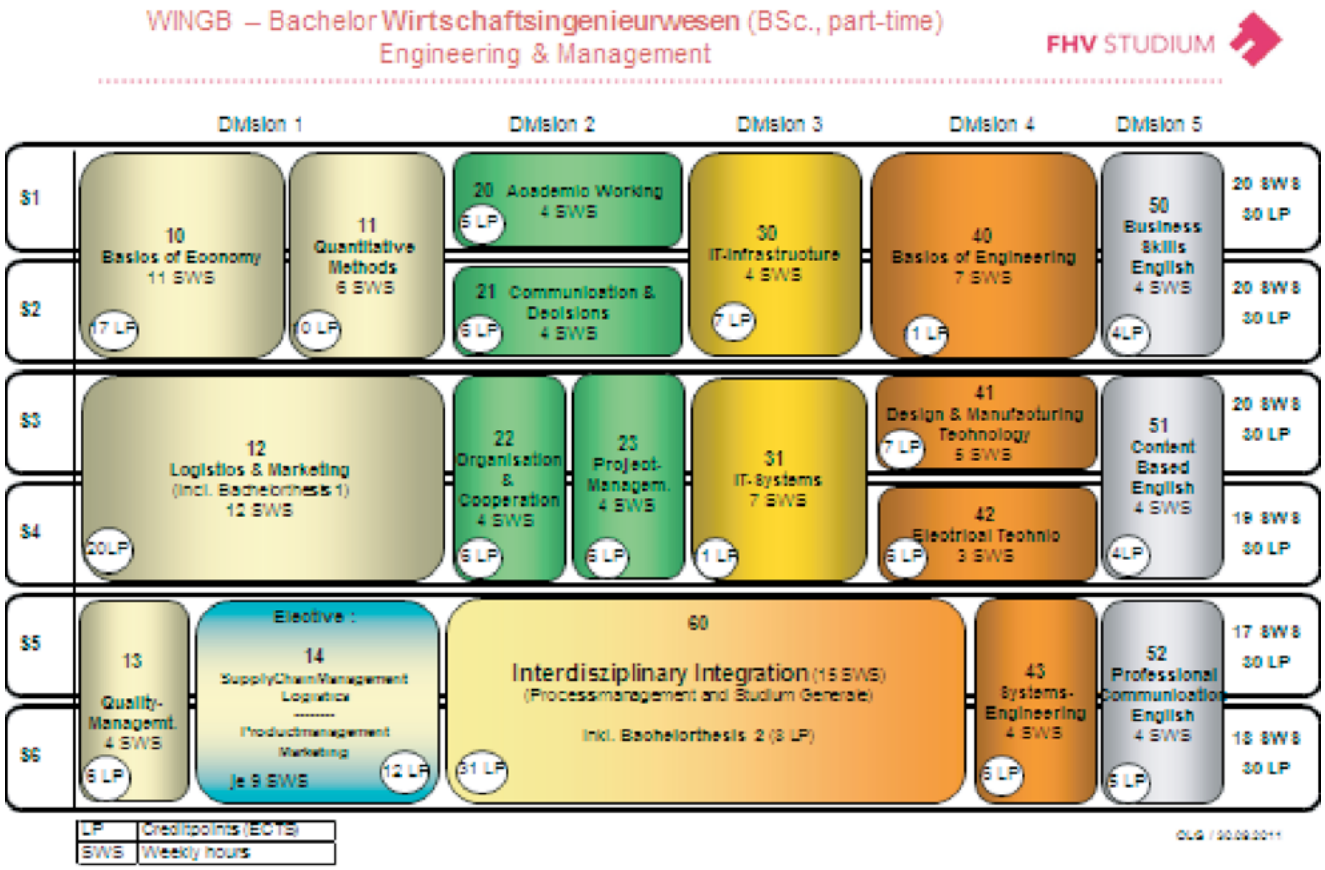


Fig. 1. Structure Bachelor of Engineering and Management

to teach various aspects of Logistics and Operations Management. In average there is one simulation about logistics per semester.

All simulations were improved, modified and adopted to our needs over the last 15 year. Usually we start with the board game version. IT-modules normally have to be designed and/or implemented by students.

Example: SCM Simulator

SCM-Simulator starts as a simple board game with very simple products (see: FIGURE 2). Lego bricks have to be completed to little towers.

Complexity can be easily increased by the instructor. After participants have analysed processes they can improve or redesign information systems and processes. First paper based. Transfer to IT is possible. As the game is dealing with physical goods (towers) material flow has to fit to information flow.

There is also a radio-frequency identification (RFID) module which can be included to the procurement process. Supplier and Customer get reading units. IT based product information can be stored and used by both the customer and the supplier. Like in real life the challenge is the implementation of the new processes and the new system while the business is going on (in the simulation).

Complexity can be increased easily. Even professionals reach their limits. The SCM-Simulator normally takes 4 full days for about 12 participants.

b) Problem Based Learning

Problem-based learning (PBL) is a method of learning in which the learners first encounter a problem, followed by systematic, student-centred enquiry process. [Barrows and Tamblyn, Problem-Based Learning: An Approach to Medical Education, Springer Publishing Company, 1980].

The enquiry process is divided into the "7-steps of PBL":

Step 1: Explain unknown wording, statements and concepts

Step 2: Define the problem(s)

Step 3: Brainstorm – analyse/try to explain the problem(s)

Step 4: Make a systematic inventory of explanations

Step 5: Formulate self-study assignments

Step 6: Perform self-study assignments

Step 7: Report and evaluate on self-study.

After each group meeting, the group formulates the next stage of the self-study assignments. For logistics we have been using PBL for more than 10 years. In 2007 we organized the first PBL conference in German language about PBL which was repeated in 2012 in Vienna.

Feedback

On regular base students give us feedback about their learning progress and about the different types of didactics we use.

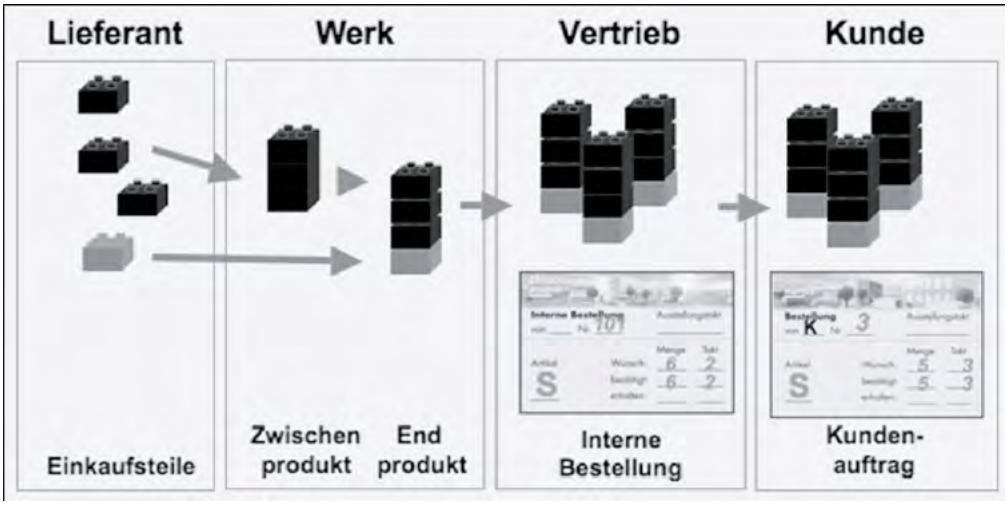


Fig. 2. SCM Simulation — Products

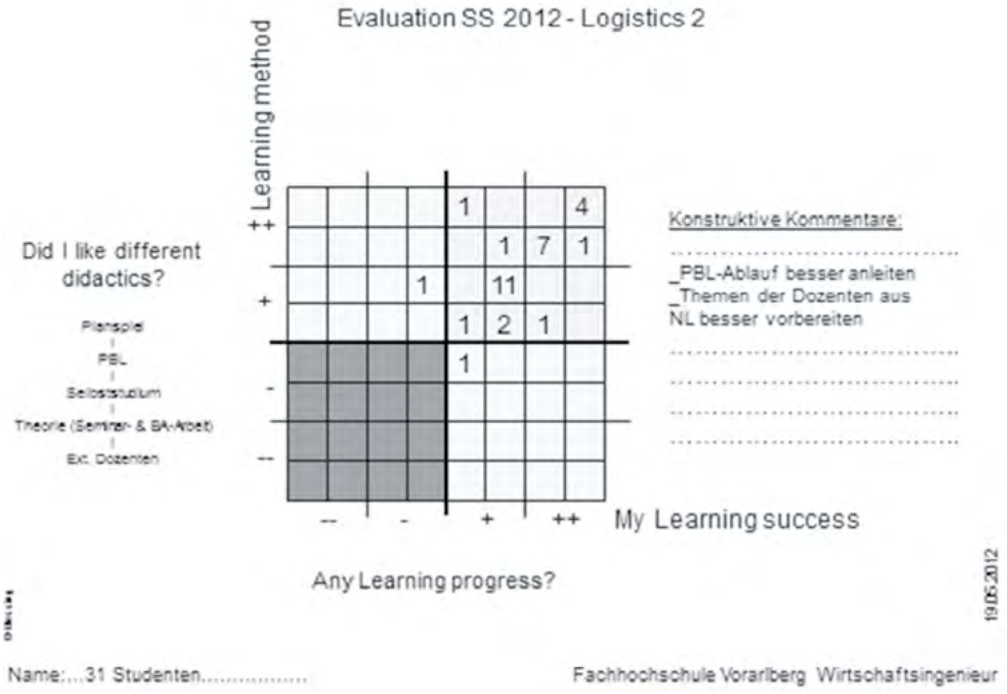


Fig. 3. Evaluation — one example

Of course not all students like all different types of didactics. This influences the learning progress of the individual student.

At the end of each semester students give a short written feedback. It influences structure and content of next year´s term.

On fig. 3 you can see the feedback results from June 2012.

The results are positive like in the previous years. But there are a few critical comments about the PBL-sessions.

They missed support in terms of structure.

Further a few students wrote that there should be a better briefing for guest lecturers from other Universities.

Obviously there was some repetition.

Future developments

Continuous improvement process in logistics education will be continued.

International aspects will become more emphasis. Summer school with focus on logistics was organized together with Chelyabinsk State University in 2011 and 2012. More than 40 students came from Austrian to South Ural to learn about logistics – and about Russia.

УДК 57.087

З.В. Максименко, И.А. Лакман, Р.Х. Бахитова

ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ ПОДДЕРЖКА ОЦЕНКИ И ПЛАНИРОВАНИЯ РАСХОДОВ НА ВЫХАЖИВАНИЕ И РЕАБИЛИТАЦИЮ НЕДОНОШЕННЫХ ДЕТЕЙ, РОДИВШИХСЯ С ОЧЕНЬ НИЗКОЙ И ЭКСТРЕМАЛЬНО НИЗКОЙ МАССОЙ ТЕЛА

Аннотация

В статье рассмотрена проблема оценки и планирования расходов на выхаживание и реабилитацию недоношенных детей, родившихся с очень низкой и экстремально низкой массой тела. Предложена концепция информационно-аналитической системы, позволяющая на основании интеллектуального анализа обновляемых данных проводить регулярный мониторинг проблемы и планирование расходов на уровне федерального/региональных бюджетов и фонда ОМС.

Ключевые слова: недоношенные новорожденные, медико-экономическая информационная система, экономическое бремя государства.

Введение

Недоношенность является основным фактором риска заболеваемости и смертности среди младенцев во всем мире и накладывают значительное бремя на здравоохранение, образование и социальные услуги, а также на бюджет домохозяйства семьи, обеспечивающей уход за такими детьми. Заболеваемость и смертность наиболее часто является следствием ранних преждевременных родов. Если проблему выхаживания детей с поздней недоношенностью как правило относят к неонатальному периоду (до 28 дней), то для глубоко недоношенных детей этот период затрагивает и постнатальный (до 1 года) и детский периоды, а иногда реабилитация продолжается и в более позднем возрасте.

В настоящее время не существует общих стандартов и протоколов, позволяющих оценить валовое экономическое бремя по уходу и реабилитации детей, родившихся с очень низкой (ОНМТ) и экстремально низкой массой тела (ЭНМТ). В связи с этим возникает необходимость в разработке общей методики по оценке и долгосрочному планированию таких расходов. Решение этой проблемы невозможно представить без применения современных информационных технологий. При планировании таких расходов следует закладывать прогноз выживаемости, а также вероятность инвалидизации глубоко недоношенных детей.

Проблеме оценки экономического бремени в Канаде с учетом прогноза выживаемости недоношенных, в том числе глубоко недоношенных, детей посвящена работа Джонсона и соавт. [1]. В работе рассматривается оценка экономического бремени на выхаживание и реабилитацию в течение первых 10 лет жизни ребенка. Причем при расчетах недоношенные дети делятся на 3 категории в соответствии с гестационным возрастом: менее 28 недель, 28–32 недели, 33–36 недель.

Оценка экономического бремени с учетом прямых и косвенных затрат, связанных с уходом за глубоко недоношенными детьми, в период до 1 года в Нигерии рассматривалась в работе Тонго и соавт. [2].

В исследовании Гилберта (США) [3] оценивались экономические затраты, связанные с преждевременными родами, включающими стоимость начальной госпитализации, стоимость любых хронических заболеваний, связанных с недоношенностью, и социальные издержки, в том числе потерю оплачиваемой работы членом семьи, ухаживающим за младенцем или ребенком, и потерю потенциальных будущих заработков.

Мураскас и Парси [4] исследовали целесообразность направления средств по выхаживанию недоношенных на повышение качества оказания дородовой помощи и уменьшения доли рождения детей с дефицитом массы тела.

Льюит и соавт. [5] показали, что с учетом дополнительных затрат по выхаживанию и адаптации до 15-летнего возраста недоношенных новорожденных, рожденных в 1988, общие расходы примерно на 6 млн. долларов выше, по сравнению с ситуацией, если бы дети родились доношенными.

Шмидт и соавт. [6] на основании данных о рождениях более 500 000 детей в штате Калифорния также приходят к выводу о непропорциональности неонатальных расходов по выхаживанию недоношенных новорожденных по отношению к расходам на новорожденных, родившихся в срок.

В России вопросы оценки экономического бремени преждевременных родов и выхаживания недоношенных новорожденных практически не исследовались. Затраты на лечение и выхаживание детей, родившихся с ОНМТ и ЭНМТ, в соответствии с высоким тарифом предусмотрены только на неонатальный период (до 28 дней) за счет выделяемого объема по высокотехнологичной медицинской помощи. Реальные расходы на лечение по оценкам различных медицинских учреждений, занимающихся выхаживанием и реабилитацией недоношенных детей (например, ГБУЗ РБ Городская детская клиническая больница № 17 г. Уфа), гораздо выше установленного тарифа [7].

Существующий подход к определению затрат не рассматривает ни дифференциацию по массе тела, ни по периоду необходимой интенсивной терапии. Так-

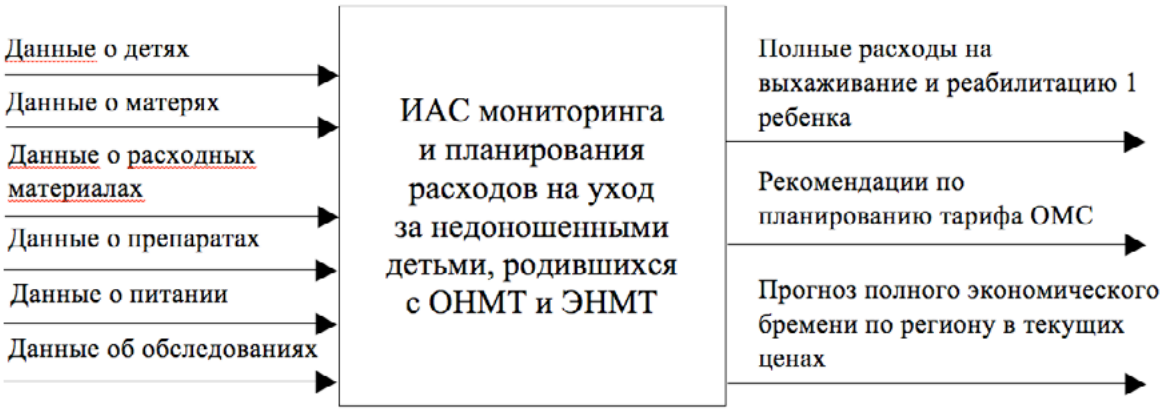


Рис. 1. Входные и выходные данные ИАС

же не учитываются риски возникновения инвалидности, ожидаемое качество жизни детей в будущем и расходы в отдаленной перспективе. Авторами помимо расходов непосредственно на выхаживание и реабилитацию на первом году жизни предлагается оценить прямые (например, затраты на лечение хронических заболеваний, связанных с недоношенностью) и косвенные (социальные издержки, связанные с возможной инвалидизацией и нетрудоспособностью ребенка в будущем) затраты.

Для упрощения расчетов, связанных с оценкой комплексного экономического бремени выхаживания и реабилитации недоношенных новорожденных, предлагается создание регистра детей, родившихся с ОНМТ и ЭНМТ, в виде базы данных в составе информационно-аналитической системы мониторинга и планирования расходов на уход за данной категорией детей. Такой подход для предупреждения смертности недоношенных детей с внутрижелудочковым кровоизлиянием был применялся в работе Хана и соавт. [8].

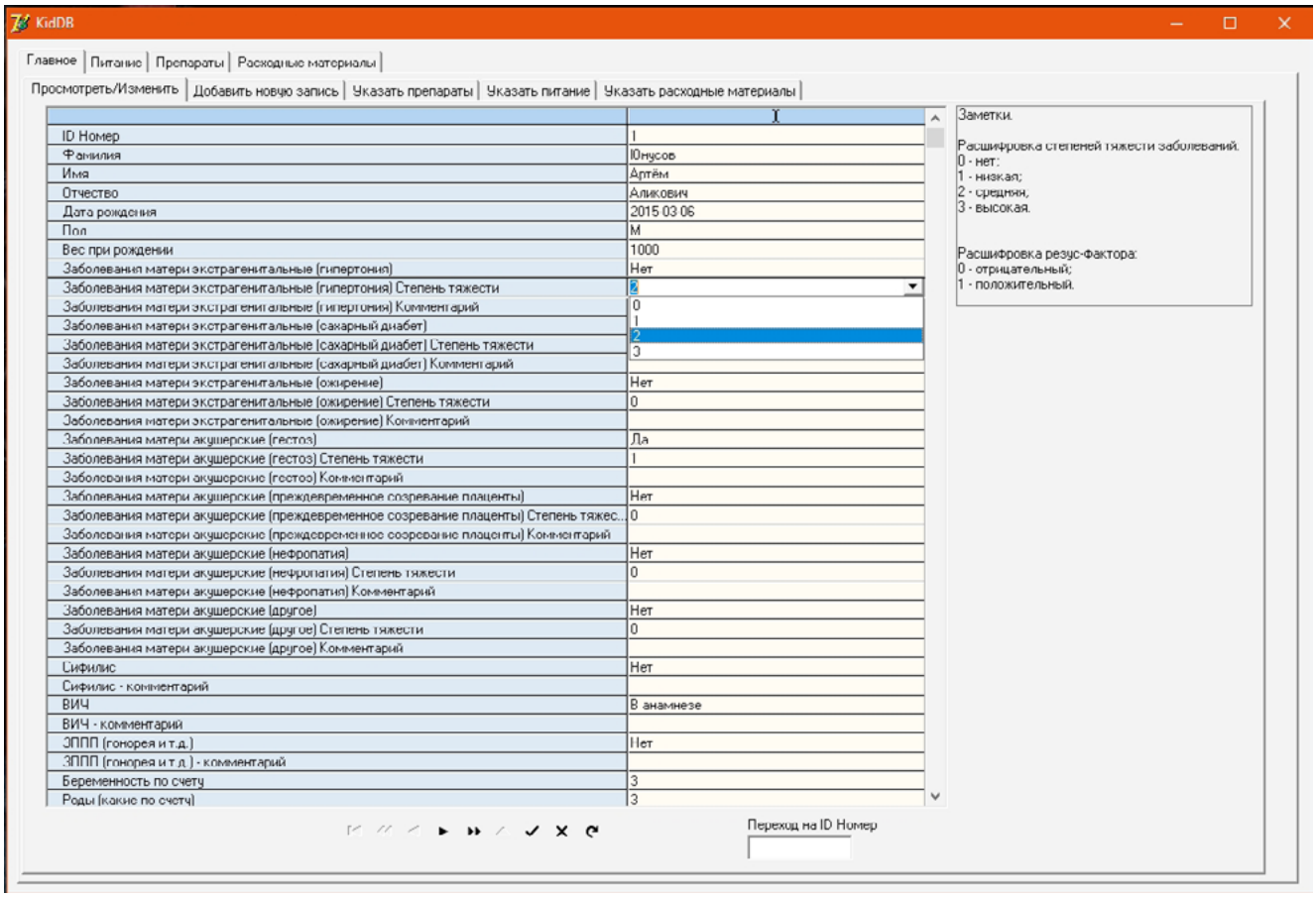


Рис. 2. Интерфейс доступа к базе данных

Методика

В рамках исследования проблемы экономических затрат, связанных с рождением глубоко недоношенных детей, для учета детальной информации о данной категории новорожденных, проведения сравнительной статистики и анализа наиболее частых видов патологии, авторами разработана реляционная база данных в СУБД MS SQL Server для информационно-аналитической системы (ИАС) медицинского мониторинга и учета расходов по выхаживанию и лечению новорожденных, родившихся с ОНМТ и ЭНМТ. Пользовательское приложение для доступа к базе данных и аналитический модуль разработаны в программной среде Delphi.

Описание ИАС

Предложена концепция информационно-аналитической системы, позволяющая на основании интеллектуального анализа обновляемых данных проводить регулярный мониторинг проблемы и планирование расходов на уровне федерального/региональных бюджетов и фонда ОМС.

Цели информационно-аналитической системы:

1. Сбор данных и ведение базы данных детей, родившихся с ОНМТ и ЭНМТ, их заболеваниях и полученном лечении, а также стоимости всех мероприятий по выхаживанию.

2. Расчет рисков выживаемости в неонатальном, постнатальном и детском периодах, риска инвалидизации, оценку качества жизни недоношенных детей, родившихся с очень низкой и экстремально низкой массой тела.

3. Формирование рекомендаций:

• по формированию тарифа ОМС на оказание гражданам медицинской помощи в части выхаживания таких детей;

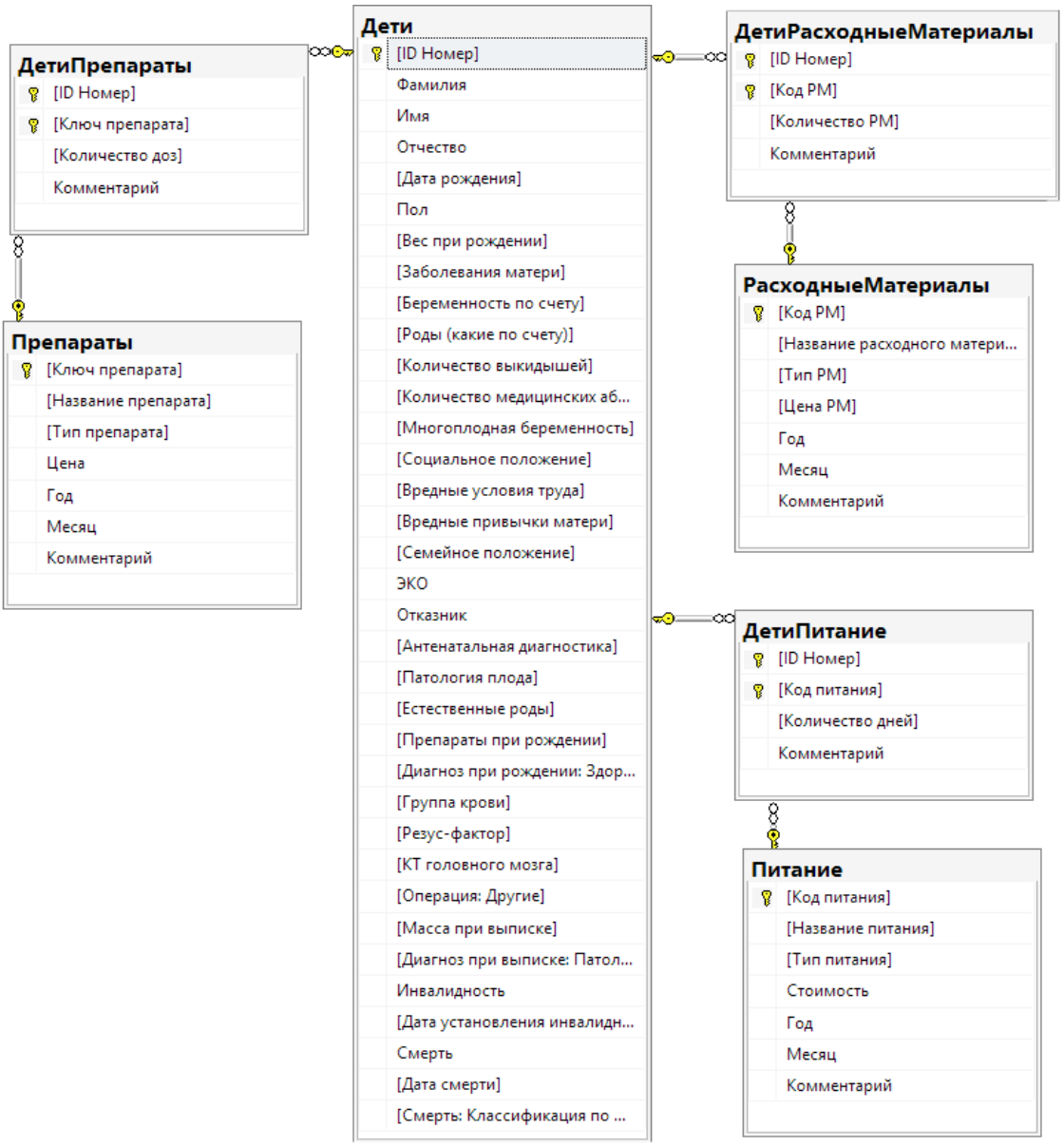


Рис. 3. Упрощённая логическая модель базы данных

• по планированию расходов ОМС на реабилитацию, поддержку здоровья и качества жизни недоношенных детей, родившихся с ОНМТ и ЭНМТ.

Входные и выходные данные ИАС представлены на рисунке 1.

В состав ИАС входят следующие компоненты: детализированный регистр в виде базы данных недоношенных новорожденных (всего более 200 показателей), приложение для пользовательского доступа к базе данных (рис. 2) и аналитический модуль.

Упрощенная схема таблиц базы данных представлена на рисунке 3. Логическая модель дает представление об общем информационном поле проводимого исследования и включает следующие таблицы [9]:

1) Дети;
2) Расходные Материалы;
3) Препараты;
4) Питание;
5) Дети: Расходные Материалы;
6) Дети: Препараты;
7) Дети: Питание.

В таблице «Дети» хранится основная информация по новорожденным и их матерям:

- пол ребенка;
- вес, состояние и диагноз ребенка при рождении;
- демографические и социальные характеристики матери;
- наличие и степень тяжести заболеваний матери;
- основные характеристики беременности и родов, патологии беременности и родов, результаты антенатальной диагностики;
- объем, виды и стоимость реанимационных мероприятий;
- обследования, консультации врачей, операции, проведенные при рождении и в период выхаживания, данные об их стоимости;
- препараты, применяемые при рождении и в период выхаживания;
- данные о парэнтеральном и энтеральном питании в период выхаживания;
- расходные материалы, в т.ч. для респираторной и инфузионной терапии;
- сумма, выплаченная ЛПУ страховой компанией за период выхаживания, а также сверхнормативные выплаты по пациенту;
- вес, состояние и диагноз ребенка при выписке, группа диспансеризации;
- сведения об установлении инвалидности;
- обследования и операции, проведенные в период реабилитации, данные об их стоимости;
- лечение, проведенное в период реабилитации (дорогостоящие лекарства, ЛФК, массаж, санаторно-курортное лечение и пр.), данные о его стоимости;
- экспертная оценка качества жизни ребенка на различных годах жизни;
- информация о смерти ребенка (включая классификацию по МКБ-10).

Таблицы «Расходные Материалы», «Препараты» и «Питание» используются для хранения информации по расходным материалам, препаратам и питанию – название, данные о стоимости и дополнительная информация.

Таблицы «Дети: Расходные Материалы», «Дети: Препараты», «Дети: Питание» используются для хранения информации о том, какому ребёнку какие расходные материалы, препараты и питание применялись при рождении и в период выхаживания и в каком количестве.

Аналитический модуль включает следующий инструментарий:

- модели выживаемости в неонатальном, постнатальном и детском периодах детей, родившихся с ОНМТ и ЭНМТ;
- модели определения риска возникновения инвалидности различной степени у недоношенных детей, родившихся с ОНМТ и ЭНМТ;
- модели ожидаемой длительности жизни детей, в зависимости от категории недоношенности и возникающих патологий;
- методику оценки экономической нагрузки на выхаживание и поддержание в течение всей жизни недоношенных детей, родившихся с ОНМТ и ЭНМТ.

Основные интеллектуальные технологии, применяемые в аналитическом модуле – технологии машинного обучения: модели анализа выживаемости (модель Кокса и модель ускоренной жизни (AFT)), логистическая регрессия, случайные леса решений.

Практическая значимость

Получено свидетельство о регистрации программы для ЭВМ [10]. В настоящее время ИАС внедрена в работу Детской городской клинической больницы № 17 г. Уфа, завершается процесс внедрения в Республиканском клиническом перинатальном центре. Проведенные к данному моменту расчеты основаны на данных представленной ИАС. Сбор данных для исследований продолжается.

Выводы

Впервые создана информационно-аналитическая система мониторинга и планирования расходов на уход за недоношенными детьми, родившимися с ОНМТ и ЭНМТ. Сформулированы требования, разработана концепция ИАС, выбраны методы и средства проектирования. В настоящее время реализованы компоненты ИАС – детализированный регистр в виде базы данных недоношенных новорожденных и приложение для пользовательского доступа к базе данных. В стадии доработки находится аналитический модуль, который позволит формировать прогноз полного экономического бремени на реабилитацию, поддержку здоровья и качества жизни недоношенных детей, родившихся с ОНМТ и ЭНМТ и рекомендации по формированию тарифа ОМС на оказание гражда-

нам медицинской помощи в части выхаживания недоношенных детей.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-010-00872 «Разработка методологии, математического и программного инструментария интеллектуальной оценки и планирования расходов на выхаживание, реабилитацию, поддержание здоровья и качества жизни недоношенных детей, родившихся с очень низкой и экстремально низкой массой тела».

Список литературы

1. Johnston K.M., Gooch K., Korol E., Vo P., Eyawo O. Bradt P., Levy A., 2014. The economic burden of prematurity in Canada. BMC Pediatrics, 14(1). Номер статьи 93.

2. Tongo O.O., Orimadegun A.E., Ajayi S.O., Akinyinka O.O., 2009. The economic burden of preterm/very low birth weight care in Nigeria. Journal of Tropical Pediatrics, 55(4):262–264.

3. Gilbert W.M., 2006. The cost of preterm birth: The low cost versus high value of tocolysis (Review). BJOG: An International Journal of Obstetrics and Gynaecology, 113(3):4–9.

4. Muraskas J. and Parsi K., 2008. The Cost of Saving the Tiniest Lives: NICUs versus Prevention. Virtual Mentor, 10(10):655–658. DOI: 10.1001/virtualmentor.2008.10.10.pfor1-0810.

5. Lewit E.M., Schuurmann Baker L., Corman H. and Shiono P.H., 1995. The Direct Cost of Low Birth Weight. The Future of Children Low Birth Weight, 5(1): 35–56. DOI: 10.2307/1602506.

6. Schmitt S.K., Sneed L. and Phibbs C.S., 2006. Costs of Newborn Care in California: A Population-Based Study. Pediatrics, 117(1):154–160. DOI: 10.1542/peds.2005-0484.

7. Бахитова Р.Х., Максименко З.В., Лакман И.А., 2018. Подход к оценке и планированию расходов на выхаживание глубоко недоношенных детей. Экономика и управление: научно-практический журнал, 5(143):144–148.

8. Han R.H., McKinnon A., CreveCoeur T.S., Baksh B.S., Mathur A.M., Smyser C.D., Strahle J.M., Olsen M.A., Limbrick D.D., 2018. Predictors of mortality for preterm infants with intraventricular hemorrhage: a population-based study. Child's Nervous System, 34(11):2203–2213.

9. Бахитова Р.Х., Гиндуллин Р.В., Максименко З.В. Структура базы данных для информационно-аналитической системы мониторинга и планирования расходов на уход за недоношенными детьми, родившимися с ОНМТ и ЭНМТ. В книге: Системное моделирование социально-экономических процессов аннотации к докладам 41-ой Международной научной школы-семинара имени академика С.С. Шаталина. Под редакцией В.Г. Гребенникова, И.Н. Щепиной. 2018. С. 32.

10. Свидетельство о регистрации программы для ЭВМ № 2018664152 «Информационно-аналитическая система мониторинга и планирования расходов на уход за недоношенными детьми, родившимися с ОНМТ и ЭНМТ» URL: http://www1.fips.ru/fips_servl/fips_servlet?DB=EVM&DocNumber=2018664152. Дата обращения: 10.02.2019 г.

УДК 004.048

В.А. Мельников, Н.Д. Харченко

ОЦЕНКА ПРЕДСКАЗАТЕЛЬНОЙ СПОСОБНОСТИ ВРЕМЕННЫХ РЯДОВ НА ВАЛЮТНОМ РЫНКЕ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Аннотация

В данной статье рассматриваются вопросы связанные с возможностью прогнозирования движения котировок валютных пар на основе использования временных рядов. Производится проверка слабой формы гипотезы эффективности рынков применительно к валютным парам валютного рынка FOREX, на основе оценки предсказательной способности предыдущих временных рядов.

Ключевые слова: финансовые рынки, гипотеза эффективности рынков, машинное обучение, прогнозирование движения рынков, рынок FOREX.

Введение

С началом двадцатого века началось активное вовлечение большой массы людей в торговлю на рынках ценных бумаг с целью обогащения. Особенно эта тенденция усилилась с появлением ЭВМ и позднее персональных ЭВМ. Теперь уже не требовалось обязательное присутствие на бирже, а достаточно было иметь удаленный доступ к бирже. Высокая вычислительная мощность ЭВМ, а также легкость присутствия на бирже дали мощный толчок в развитии методов прогнозирования движения на коротком промежутке времени на финансовых рынках.

Методы прогнозирования финансовых рынков подразделяются на два типа [1]:

1) Фундаментальный анализ [2].

2) Технический анализ (прогнозирование на основе предыдущих временных рядов) [3].

Фундаментальный анализ строит прогноз цены финансового инструмента на основе различных макро и микро экономических показателях исследуемого актива. Фактически на основе финансового анализа производится оценка реальной стоимости актива. Данный вид анализа сложен для освоения, и поэтому распространен в основном среди профессиональных трейдеров. Хорошо подходит для прогнозирования на длительный срок прогнозирования, так как часто требуется большой период времени прежде чем актив достигнет справедливой цены [2].

Технический анализ строит прогноз цены финансового актива на основе анализа предыдущих временных рядов данного актива. Технический анализ говорит о цикличности и повторяемости движения цены актива, что позволит выявить различные паттерны [3]. Данный вид анализ с использованием ЭВМ просто для проведения, и поэтому получил в последнее время массовое распространение среди трейдеров. Следует отметить, что данный вид анализа отлично подходит как для внутридневной торговли так и для прогнозирования на большие периоды. Прежде чем описать основное противоречие технического анализа, следует рассмотреть Гипотезу эффективного рынка [4].

Гипотеза эффективного рынка была выдвинута в 1970 году известным американским экономистом, Нобелевским лауреатом, Юджином Фама [4]. Согласно данной гипотезе вся существенная информация немедленно и в полной мере отражается в цене актива. Принято выделять 3 формы рыночной эффективности:

1) слабая форма эффективности, если стоимость рыночного актива полностью отражает прошлую информацию, касающуюся данного актива (общедоступная в настоящий момент времени информация о прошлом состоянии рынка, прежде всего по динамике курсовой стоимости и объемах торговли финансовым активом);

2) средняя форма эффективности, если стоимость рыночного актива полностью отражает не только прошлую, но и публичную информацию (текущая информация, которая становится общедоступной в настоящий момент времени, предоставленная в текущей прессе, отчётах компаний, выступлениях государственных служащих, аналитических прогнозах и т.п.);

3) сильная форма эффективности, если стоимость рыночного актива полностью отражает всю информацию – прошлую, публичную и внутреннюю (инсайдерская информация, которая известна узкому кругу лиц в силу служебного положения, или иных обстоятельств).

Как мы видим даже выполнение слабой формы эффективности уже позволяет говорить о несостоятельности технического анализа.

Обзор существующих исследований

К настоящему моменту были проведены ряд исследований направленных на опровержение применимости технического анализа для прогнозирования движения курсов акций. В частности следует отметить работы [5–7], результатом которых стал вывод о присутствии слабой формы гипотезы на рынках акций, и невозможности прогнозирования движения цены используя технический анализ. В то же время

есть работы [8–9], в которых указывается возможность использования технического анализа. В рамках нашего исследования мы проверим прогностическую способность временных рядов на международном валютном рынке (FOREX).

Описание эксперимента

В качестве инструментария для определения прогностической способности воспользуемся методами машинного обучения. Методы машинного обучения позволяют найти неявные зависимости между входными и выходными данными [10]. В рамках эксперимента будем производить бинарную классификацию, где 1 – цена вырастет в след периоде и 0 – цена упадет в след периоде. Для решение данной задачи классификации были выбраны 3 популярных алгоритма машинного обучения: метод ближайших соседей, случайный лес и метод стохастического градиентного спуска. В качестве метрик будем использовать Accuracy (доля правильных ответов алгоритма), F – мера, Precision (Точность). Обучение производилось на основе 60 % выборки с использованием k-fold cross validation, при k = 3, тестирование проводили на 40 % выборки.

Рассмотрим каждый из методов машинного обучения и его параметры:

1) Метод ближайших соседей (KNN) [11].

Данный метод определяет объект к тому классу, к которому относятся большинство его k-ближайших соседей. Близость высчитывается с помощью функции расстояния.

В данном методе перебирались следующие параметры:

• Количество ближайших соседей – параметр, определяющий количество объектов с которыми будет производится сравнение, перебирался в диапазоне 5,6,7,…,28,29,30.

• P – параметр для регулирования расстояния Минковского, перебирался в диапазоне 1, 2, 3, 4, 5.

2) Случайный лес (Random Forest) [12].

Является ансамблем решающих деревьев, вся выборка делится на под выборки, по каждой из которых строится отдельное дерево решений, при этом для построения каждого расщепления в дереве из общего количества признаков отбираются случайные и из них выбираются лучшие. Каждое дерево строится независимо друг от друга и в случае классификации решение принимается как голосование по большинству.

Параметры для перебора:

• Количество деревьев в ансамбле: 100,150,200,…,900,950,1000

• Максимальная глубина деревьев: None, 80, 90, …,140,150.

• Максимальное количество признаков: auto, sqrt, loge2, None.

• Минимальное количество объектов, судействуемых в разделении узла: 2,3,4,5,6,7,8,9.

• Минимальное число объектов в узле: 1,2,3,4,5

3) Стохастический градиентный спуск (SGDClassifier) [13].

Это метод нахождения минимума или максимума функции при помощи движения по градиенту.

Параметры:

• Альфа – константа умножаемая на член регуляризации: 1e-4, 1e-3, 1e-2, 1e-1, 1e0, 1e1, 1e2, 1e3.

• Функция потерь: hinge, log.

• Функция штрафа: l2,l1, elasticnet.

Во всех экспериментах, проводимых в рамках данного исследования в качестве целевой переменной, был избран факт повышения открывающей цены на один пункт или более в течение ближайших пяти минут. Выбор такой метки обусловлен тем, что выигрыш хотя бы в один пункт позволяет окупить комиссию на сделку, которая как правило составляет 0,7 пункта, и в тоже время предоставляет достаточно сбалансированную обучающую выборку с соотношением True и False как 40 % и 60 %. Исходные данные представляют из себя два миллиона последовательных минут, в ходе которых происходили торги валютной пары USD/JPY. В каждой минуте хранится информация о цене открытия, цене закрытия, наивысшей и наименьшей ценой в течение минуты.

Проведение экспериментов

Ниже представлены результаты экспериментов с различным набором входных признаков.

Таблица 1

В качестве признака использовали цену открытия предыдущего периода

	Accuracy	F- мера	Precision
RandomForest	0,532	0,464	0,449
SGDClassifier	0,511	0,428	0,402
KNN	0,527	0,454	0,45

В таблице 1 представлен результат эксперимента, где прогнозирование производилось на основании значения предыдущего периода. Лучше всего себя показал метод случайных лесов, с показателем precision 0,449 и F-мерой 0,464. В то же время точность угадывания явно недостаточна.

Таблица 2

В качестве признаков использовали цену открытия предыдущих 30 1-минутных периодов

	Accuracy	F-мера	Precision
RandomForest	0,538	0,463	0,453
SGDClassifier	0,507	0,424	0,398
KNN	0,513	0,452	0,439

В таблице 2 представлен результат эксперимента, где в качестве входных данных использовались

30 1-минутных периодов. Как видим метод случайных лесов и в этом эксперименте показал себя лучше других. Так же использование большего количества данных позволило поднять точность угадывания – 0,453 против 0,449. Но так же недостаточно для получения положительного результата торговли.

Таблица 3

В качестве признаков использовали процент изменения цены открытия предыдущих 30 1-минутных периодов по отношению к текущей цене открытия и направление этого изменения в виде бинарного признака

	Accuracy	F- мера	Precision
RandomForest	0,566	0,493	0,487
SGDClassifier	0,535	0,467	0,461
KNN	0,523	0,446	0,432

В таблице 3 представлен результат эксперимента, где в качестве входных параметров использовался относительный показатель изменения цены 30 1-минутных периодов по отношению к текущей цене открытия, а так же направление этого изменения в качестве бинарного признака. Как мы видим, и в этом эксперименте метод случайных лесов показал лучший результат среди других методов. При этом и все показатели оказались гораздо лучше, чем результаты прошлых двух экспериментов – точность 0,487 против 0,453 и F-мера 0,493 против 0,463. И хотя полученный результат явно превосходит результаты прошлых экспериментов, он все еще не позволяет получать прибыль от торговли за вычетом комиссии от сделки.

Таблица 4

В качестве признаков использовали процент изменения цены открытия предыдущих 30 1-минутных периодов по отношению к текущей цене открытия, направление этого изменения в виде бинарного признака и разницу между текущей ценой открытия и ценами открытия за последние 30 минут

	Accuracy	F- мера	Precision
RandomForest	0,521	0,455	0,431
SGDClassifier	0,487	0,411	0,402
KNN	0,613	0,397	0,321

В 4 эксперименте в качестве входных параметров мы использовали как процентное изменение цены, так и фактическое изменение цены последних 30 1-минутных периодов. Как видим показатели точности и F-меры ухудшились по сравнению с прошлым экспериментом. Очевидно, что добавление новых данных привнесло лишь шум. Лучше всего себя показал метод случайных деревьев. Отдельно хочется выделить метод ближайших случайны соседей, показатель accuracy которого оказался очень высоким. Причина в том, что данный метод лучше других определял плохие сделки, но крайне плохо определял правильные. В целом же следует отметить, что результат пока отрицательный.

Таблица 5

В качестве признаков использовали процентное изменение между текущей ценой и ценами открытия за последние 30 минут, и процентную разницу между каждой парой цен открытия за последние 30 минут.

	Accuracy	F- мера	Precision
RandomForest	0,569	0,492	0,489
SGDClassifier	0,541	0,471	0,465
KNN	0,577	0,426	0,424

В пятом эксперименте в качестве входных параметров использовали относительные показатели изменения цен к текущему периоду и показатель изменения парных цен. Как видно из таблицы 5 метод случайных лесов показывает лучший результат в сравнении с другими методами и показывает схожий результат, который был получен во время третьего эксперимента.

Таблица 6

В качестве признаков использовали 4 параметра предыдущего 1-минутного периода: цена открытия, цена закрытия, максимум, минимум

	Accuracy	F- мера	Precision
RandomForest	0.525	0.427	0.437
SGDClassifier	0,424	0,595	0,411
KNN	0,521	0,451	0,439

В шестом эксперименте будем использовать все показатели предыдущего 1-минутного временного периода, а именно – цена открытия, цена закрытия, максимум внутри периода, минимум внутри периода. Полученные результаты представлены в таблице 6. Как видим, в данном эксперименте метод ближайших случайных соседей показал максимальную точность угадывания положительных сделок – 0,439 при F-мере 0,451. Так же следует выделить метод стохастического градиентного спуска показавшего F-меру 0,595 при самой низкой точности угадывания положительных сделок. Очевидно, что метод обучился ставить больше убыточных сделок и за счет этого собрал столь высокий показатель F-меры. В целом же полученные результаты хуже экспериментов 3 и 5.

В седьмом эксперименте в качестве входных параметров мы использовали булевые переменные, отражающие вырос или упал соответствующий показатель прошлого периода к текущему. Если цена открытия была выше цены закрытия то ставился True в противном случае False. И так для каждого из 4х показателей.

Таблица 7

В качестве признаков использовали булевые переменные, отражающие поднялись или опустились значения наших признаков по сравнению с предыдущим периодом

	Accuracy	F- мера	Precision
RandomForest	0.528	0.425	0.44
SGDClassifier	0,424	0,595	0,424
KNN	0.521	0,45	0,439

Полученные результаты представлены в таблице 7. Как видим лучшим опять оказался метод ближайших соседей с показателем точности 0,439 и F-мерой – 0,45. Данные результаты незначительно отличаются от полученных нами в прошлом эксперименте.

Таблица 8

В качестве признаков использовали цену открытия, закрытия, максимум и минимум для последних 30 1-минутных периодов

	Accuracy	F- мера	Precision
RandomForest	0.511	0.405	0.412
SGDClassifier	0,426	0,573	0,465
KNN	0.533	0,458	0,464

В восьмом эксперименте в качестве входных параметров мы использовали значения открытия, закрытия, максимум и минимум предыдущих 30 1-минутных периодов. Результаты эксперимента представлены в таблице 8. Как видно, лучшим методом оказался метод стохастического градиентного спуска с показателем точности 0,465 и F-мерой 0,573. Полученный результат не превосходит результаты случайных деревьев из эксперимента 3 и 5.

В девятом эксперименте мы добавили к предыдущим признакам относительную составляющую изменения цен периодов к текущему периоду. Результат эксперимента представлен в таблице 9.

Таблица 9

В качестве признаков использовали цену открытия, закрытия, максимум и минимум для последних 30 1-минутных периодов и информацию о процентных изменениях цен

	Accuracy	F- мера	Precision
RandomForest	0.532	0.414	0.426
SGDClassifier	0,477	0,58	0,417
KNN	0.52	0,449	0,441

Как видно, лучшим методом опять казался метод случайных соседей с показателем точности 0,441 и F-мерой 0,449. Данный результат оказался хуже, чем результат полученный в прошлой таблице. Можно сделать вывод, что добавление новых признаков дало больше шума.

В рамках десятого эксперименты, мы попробовали оставить только относительные признаки изменения цен к ценам текущего периода. Результат представлен в таблице 10.

Таблица 10

В качестве признаков использовали процентное изменения цен открытия, закрытия, максимумов, минимумов к ценам текущего периода

	Accuracy	F- мера	Precision
RandomForest	0,501	0,41	0,454
SGDClassifier	0,525	0,461	0,471
KNN	0,593	0,486	0,437

Лучше всех себя проявил метод стохастического градиентного спуска с показателем точности – 0,471 и F-мерой – 0,461. Данный результат лучше результата полученного в рамках восьмого эксперимента, но все еще хуже результатов третьего и пятого экспериментов.

В одиннадцатом эксперименте в качестве входных параметров мы использовали относительные изменения 4х параметров периода как к текущему периоду, так и между ближайшими периодами. Результат эксперимента представлен в таблице 11.

Таблица 11

В качестве признаков использовали процентное изменения цен открытия, закрытия, максимумов, минимумов к ценам текущего периода, и процентную разницу между каждой парой цен за последние 30 1-минутных периодов

	Accuracy	F- мера	Precision
RandomForest	0,539	0,424	0,416
SGDClassifier	0,521	0,453	0,424
KNN	0,576	0,379	0,361

Как видим лучшим оказался метод стохастического спуска, но значения точности и F-меры оказались сильно хуже результатов, полученных в прошлых экспериментах.

Выводы

Как видим на основе проведенных экспериментов нам не удалось получить какую-либо статистически значимую прогностическую способность предыдущих временных рядов. Максимум что удалось получить, это точность равную 0,489 и F-меру 0,493 (Эксперимент 3 и 5 метод случайных лесов), что является недостаточным для получения положительного результата торговли. В целом же добавление и убавление различных параметров практически не влияет на конечный результат, либо ухудшает его значительно. При используемой выборке, где 40 % сделок позволяли заработать, а 60 % нет, результат с точностью 0,489 не даст положительного результата в

торгах. Таким образом нам не удалось опровергнуть выполнение слабой формы гипотезы эффективности рынков применительно для рынков FOREX на коротком периоде прогнозирования. Вместе с тем наличие слабой формы указывает на несостоятельность всего технического анализа и технических индикаторов, основанных на данных предыдущих временных рядов применительно к рынку FOREX для короткого периода прогнозирования.

Список литературы

1. Financial market analysis methods. Date views 31.01.2019 https://privatfinance.com/metodyi-analiza-finansovogo-ryinka/

2. Fundamental analysis of financial markets. Date views 31.01.2019 https://equity.today/fundamentalnyj-analiz-cennyx-bumag.html.

3. John J. Murphy. echnical Analysis of the Financial Markets. Date views 31.01.2019 https://ru.wit-invest.com/userfiles/file/dzhon_dzh.merfi_-_tehnicheskiy_analiz_fyuchersnyh_rynkov._teoriya_i_praktika.pdf.

4. Eugene F. Fama. 1970. Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2):383–417.

5. Purmonen, Sami, and Paul Griffin. 2015. Testing Stock Market Efficiency Using Historical Trading Data and Machine Learning. Date views 31.01.2019 http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-166583.

6. Harper, Kevin M. 2016. Challenging the Efficient Market Hypothesis with Dynamically Trained Artificial Neural Networks, UNF Graduate Theses and Dissertations, 718. Date views 31.01.2019 https://digitalcommons.unf.edu/etd/718.

7. Das SR, Mokashi K, Culkin R. 2018. Are Markets Truly Efficient? Experiments Using Deep Learning Algorithms for Market Movement Prediction. *Algorithms*, 11(9):138.

8. Rajashree Dash, Pradipta Kishore Dash. 2016. A hybrid stock trading framework integrating technical analysis with machine learning technique. The Journal of Finance and Data Science.2(1):42-57. Date views 31.01.2019 https://doi.org/10.1016/j.jfds.2016.03.002.

9. Shen, Shunrong, Haomiao Jiang and Tongda Zhang. 2012. Stock Market Forecasting Using Machine Learning Algorithms. Date views 31.01.2019 https://www.semanticscholar.org/paper/Stock-Market-Forecasting-Using-Machine-Learning-Shen-Jiang/b68e8d2f4d2c709bb5919b82effcb6a7bbd3db37.

10. Cosmas Wong. How Machine Learning Uncovers Hidden Correlations for Marketers. Date views 31.01.2019 https://www.cmswire.com/digital-experience/how-machine-learning-uncovers-hidden-correlations-for-marketers/

11. Mohtadi Bin Fraj. In Depth: Parameter tuning for KNN. Date views 31.01.2019 https://medium.com/@mohtedibf/in-depth-parameter-tuning-for-KNN-4c0de485baf6.

12. Probst, Philipp & Boulesteix, Anne-Laure & Wright, Marvin. 2018. Hyperparameters and Tuning Strategies for Random Forest. Date views 31.01.2019 https://www.researchgate.net/publication/324438530_Hyperparameters_and_Tuning_Strategies_for_Random_Forest.

13. Stochastic Gradient Descent. Date views 31.01.2019 https://code.i-harness.com/ru/docs/scikit_learn/modules/sgd.

УДК 004.9

Р.В. Мещеряков

ИНФОКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ УПРАВЛЕНИЯ ОКАЗАНИЯ УСЛУГ НАСЕЛЕНИЮ С ИСПОЛЬЗОВАНИЕМ ТЕЛЕВИДЕНИЯ

Аннотация

В статье представлен подход к системе оказания услуг населению с использованием современных телекоммуникационных технологий. Приведена обобщенная структура использования сети телевещания для построения сети оказания услуг населению. В качестве базы используется цифровое телевидение, для обратной связи используется сеть общего пользования Интернет. Рассматриваемая социо-киберфизическая система обладает ограничениями на прием обрабатываемой информации. Предлагается использование критерия структурной сложности для оценивания количества передаваемой и обрабатываемой информации в центрах оказания услуг и в абонентских пунктах. Приведены расчетный пример и практическая апробация, которые показывают возможность реализации и расширения функционала предлагаемой технологии для все страны. Предлагается использование технологии «интернет вещей» для обеспечения оказания услуг с учетом наличия сети телевещания в замкнутом контуре цифровых услуг. Делается предложение по развитию системы с учетом больших данных и информационной безопасности.

Ключевые слова: инфокоммуникационные технологии, управление, телевизионный поток, государственные услуги, имитационная модель, прогнозирование, социо-киберфизическая система, интернет вещей.

Введение

В настоящее время происходит быстрый рост инфокоммуникационных технологий (ИКТ). Несмотря это и разнообразие технических средств передачи информации, телевидение в России остается одним из главных источников информированности населения. Несмотря на переход в 2019 году на федеральном уровне к использованию современного формата вещания DVB-T2 [1, 2], телевидение нуждается в развитии новых форматов интерактивности и возможностей предоставления новых услуг на базе имеющейся инфраструктуры. На этом этапе требуется предусмотреть проблемы, которые могут возникнуть при этом развитии, и принять системные меры для их предотвращения. В связи с этим, органам управления образованием важно не только обеспечивать построение необходимой инфраструктуры – каналов связи, серверов, пользовательской аппаратуры, – но и уделять как можно больше внимания содержательной части инфокоммуникационных услуг.

В то же время в Российской Федерации активно развивается проект предоставления государственных услуг через специализированный портал [3]. В основу проекта положено распространение широкополосного доступа к сети Интернет на территории всей страны. Замедляющим фактором для данного процесса являются высокая степень затрат, большая территория и, как следствие, нерентабельность обеспечения отдаленных поселений скоростными каналами связи. Вместе с тем активно развивается программа «Цифровая экономика», которая предполагает развитие цифровой экономики, «в которой данные в цифровой форме являются ключевым фактором производства во всех сферах социально-экономической деятельности» [4].

Перспективным могло бы стать совмещение телевизионного вещания и сетевых технологий – гибридное телевидение, одной из основных возможностей которого является способность передавать в основном потоке вещания дополнительные данные [5] об интерактивных сервисах и информационных услугах.

Первая из проблем, которые могут возникнуть, заключается в нецелевом использовании ресурсов, затрачиваемых на построение инфраструктуры. Как известно, доступность развитых ИКТ не является достаточным условием успешного решения задач, для которых эти технологии внедряются. Как показывает практика, если процесс развития массовых ИКТ не контролируется, основная часть возникающих при этом возможностей используется не для решения социальных задач, а для предоставления услуг развлекательного характера. В связи с этим, без средств контроля и ограничений создаваемые в процессе информатизации возможности использоваться неэффективно. Этой проблеме было уделено много внимания на недавно прошедшей в Санкт-Петербурге встрече стран Азиатско-Тихоокеанского экономического сотрудничества (АТЭС), что нашло отражение в тексте Петербургской декларации, принятой по итогам совещания министров связи стран АТЭС [6].

К проблеме нецелевого использования ресурсов относится также наблюдаемое явление увеличения количества передаваемых данных в расчете на одну оказанную услугу. Стремительное развитие ИКТ часто дает иллюзию неограниченной емкости каналов связи, в результате чего вместе с действительно необходимой информацией передается большое количе-

ство данных, несущих рекламную или декоративную функцию (например, баннеры или многочисленные элементы графического оформления веб-страниц). По этой причине лишь малая часть передаваемого сегодня трафика несет полезную нагрузку.

В ходе аналитического обзора информационных источников установлено, что современной научно-технической проблемой является отсутствие способа встраивания в основной телевизионный поток дополнительной информации с последующим использованием абонентского устройства в качестве узла конвергентности инфокоммуникационных услуг [7]. Таким образом, актуальной представляется задача разработки технологии инкапсулирования в цифровой поток формата DVB-T2 данных стека протоколов TCP/IP [5, 7] для определения возможности предоставления государственных услуг в электронном виде посредством телевидения.

Система оказания услуг

В работе [8] представлены результаты встраивания услуг в цифровой поток, где программные потоки формируются на уровне первого мультиплексирования цифрового сигнала и рассчитаны на использование надежных носителей, например, DVD-дисков. Таким образом, можно использовать существующие стандартные технологии инкапсуляции дополнительных данных непосредственно в программный поток. Возможны варианты с «воздушным зазором», значительно увеличивающие время выполнения услуги. Обобщенная схема представлена на рисунке 1.

Вместе с тем остаются нерешенными вопросы при использовании технологий «последней мили» для потребителей продукции. Очевидно, что по [9, 10] имеются три режима работы сети – «точка-точка», «точка-множество», «множество-множество». Вместе с тем необходимо учитывать, что данная система является социо-киберфизической, поэтому необходимо гармонизировать не только по пропускной способности канала, но и по количеству информации, допустимой для обработки в узлах. В случае оказания услуг с использованием предложенного способа целесообразным представляется использование технологии «интернет вещей» [11].

Для оценивания возможности функционирования сети предлагается использование показателя энтропии для оценки трудоемкости выполняемых операций в системе оказания услуг [12]. Рассмотрим частный случай получения потребителем широковещательного сообщения по телевизионному каналу. Примем, что блок обработки информации, на вход которого поступает s потоков, в каждом из которых может быть q состояний, производит некоторое преобразование B, в результате которого на выходе полу-

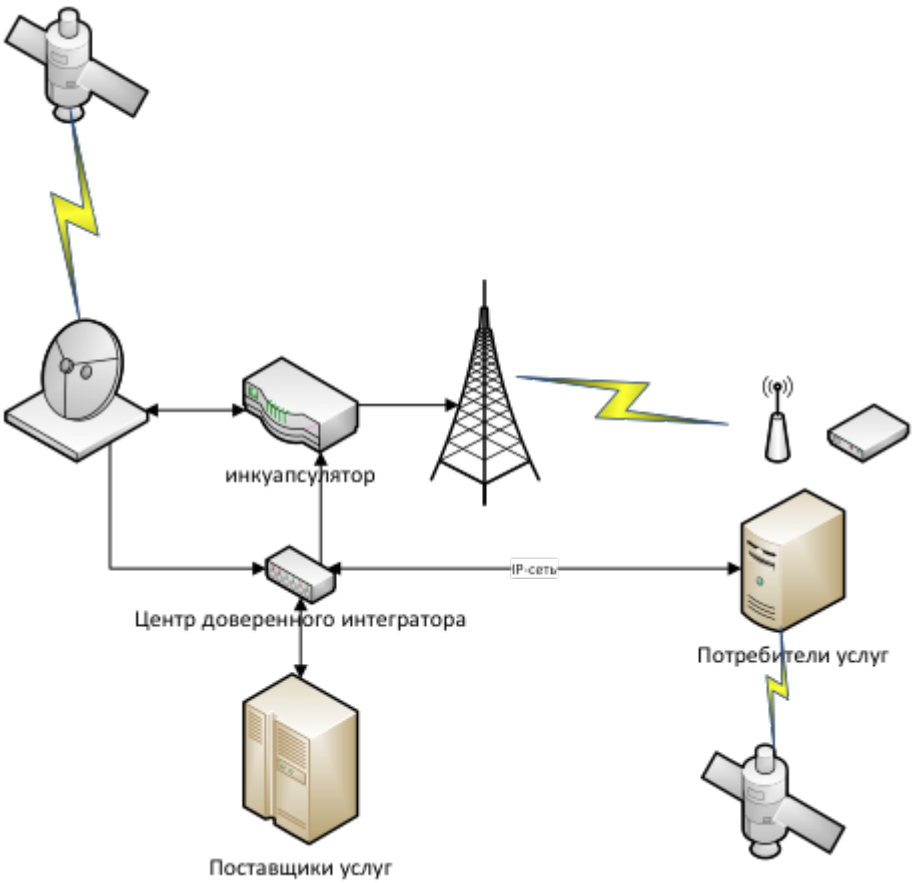


Рис. 1. Схема функционирования сети оказания услуг с использованием телевидения

чаем s' потоков, в каждом из которых может быть q' состояний: B: $s \times q \rightarrow s' \times q'$.

Оценим количество входящей обрабатываемой информации в битах, при этом сделаем допущение, что вероятность p нахождения каждого потока s в состоянии q одинакова. Тогда:

$$H = -\sum_{i=1}^{s} p_i \log_2 p_i = -s \times p \times \log_2(p) \quad (1)$$

Таким образом, постановка задачи разбиения и выполнения задачи оказания услуги сводится к оцениванию конечного устройства обработки информации с одновременным учетом человеческого фактора. При этом частная задача определения оптимального числа блоков обработки и количества уровней иерархии (при гетерогенной системе оказания услуг, например, в случае делегирования оказания услуг с федерального уровня на региональный, далее муниципальный и т.д.) представляет собой задание функционала:

$$\arg\min_{G \in \Omega} R(G)$$

где Ω – множество представлений процесса обработки информации (с представлением в иерархическом виде) с заданным функционалом R: $\Omega \rightarrow [0,+\infty)$. При этом определение функционала представляет собой постановку задачи определения ограничений на оптимизацию процесса обработки информации и по существу он представляет собой функционал стоимости.

Оценивание услуги и практическая реализация

Рассмотрим частный случай оказания услуги предоставления информации по телевизионному каналу и заказа дополнительных сведений.

Входные данные: Центр оказания услуг формирует широковещательное сообщение длительностью 1 минута (1080p – 155 Мб), количество предложений – 100, вариаций – 2, количество абонентов – 1 000, объем возвращаемого сообщения на услугу 4 Кб.

Тогда получаем количество выданной информации по (1) H=66, а каждый клиент получает информации 0,066. Таким образом, абонентский пункт пропускает через себя 155 Мб/минуту и выдает 0,066.

Предположим, что для формирования заявки сформируется запрос с ответом на 100 предложений. В качестве модельного примера предположим, что ответ на услугу сформировало всего 10 абонентов. Тогда количество возвращаемой информации составит 0,66 при общем входящем потоке в 40 Кб.

Для практической апробации создано 2 стенда «Исследование возможности инкапсуляции данных в программный поток» и «Исследование возможности инкапсуляции данных на уровне MPEG TS».

Транспортный поток MPEG (MPEG-TS или TS) является стандартным форматом для передачи и хранения аудио, видео и цифровых потоков, программ и систем данных. Он определяет формат упаковки элементарного потока данных с коррекцией ошибок и функциями синхронизации потока для поддержания целостности передачи в случае деградации сигнала. Транспортные потоки отличаются от одноименных программных потоков тем, что в первую очередь предназначены для менее надежной передачи, а именно для наземного или спутникового вещания. Кроме того, транспортный поток может включать несколько программ.

Был проведен ряд измерений в соответствии с ETSI TS 102 032 recommendation [7], которые подтвердили работоспособность исследуемых технологий. При этом измерения MER [11] проводились с учетом методики, использованной на других схемах инкапсуляции, благодаря чему удалось избежать ситуации, когда возможны краткие развалы картинок при нормальном MER.

В настоящем модельном примере можно выделить узкие места – при большом потоке сообщений последовательное получение центром оказания услуг может не обеспечивать требуемого качества оказания услуг с задержкой обработки для формирования ответов на запрашиваемые услуги.

Выводы

Очевидно, что в простейшем случае система не решает задачи реального времени, тем не менее, имеется возможность перспективного развития системы цифрового телевидения. Одним из наиболее важных результатов исследования является подтверждение возможности предоставления государственных услуг населению Российской Федерации посредством телевизионного вещания. Данное положение может послужить опорной ступенью в решении важных государственных задач, таких как государственная программа «Цифровая экономика» [2].

Важность проведенных исследовательских работ состоит в способности разработанных схем и технологий предоставить дополнительные сервисы и информационные услуги без доступа к сети Интернет. Данная задача является весьма актуальной для переходного периода развития доступности Интернета, особенно в условиях удаленных территорий, где целесообразность прокладки проводных линий связи не очевидна и экономически невыгодна, спутниковый доступ во всемирную сеть имеет высокую стоимость, что является существенным фактором для применения технологий для широких масс населения.

Внедрение предложенной технологии открывает новый этап в использовании телевидения в государственных масштабах; в построении независимого или резервного канала информационной поддержки государственного управления через доверенное пространство. Вместе с тем, остаются актуальными вопросы обеспечения гиперсвязности [13] и обработки больших данных [14], но и формирования действительно защищенной среды, обеспечивающей конфиденциальность, целостность и доступность в инфокоммуникационном пространстве [15].

Благодарности

Автор благодарит д.т.н., профессора, академика НАН РА Сарьяна В.К. за критику, обсуждение основных идей, изложенных в настоящей статье.Работа выполнена при частичной поддержке гранта РФФИ 19-08-00331 А.

Список литературы

1. Информация о стандарте DVB http://www.dvb.org/standards/dvb-t2.

2. Формат MPEG TS. Описание http://www.afterdawn.com/glossary/term.cfm/mpeg2_transport_stream.

3. Электронное правительство. Госуслуги http://www.gosuslugi.ru.

4. Программа «Цифровая экономика Российской Федерации», утверждена Распоряжением от 28 июля 2017 г. № 1632-р Правительством Российской Федерации.

5. Amerini I.A DVB-MHP web browser to pursue convergence between digital terrestrial television and Internet / I. Amerini, R. Becarelli, R. Caldelli, F. Filippini, G. Ballocca, R. Borri // Multimedia tools and applications. – 2010. – № 2 (50) – P. 381–414.

6. Saint Petersburg Declaration – Building Confidence and Security in the Use of ICT to Promote Economic Growth and Prosperity. – 2012. – http://www.apec.org/Meeting-Papers/Ministerial-Statements/Telecommunications-and-Information/2012_tel.aspx.

7. ETSI TS 102 032 recommendation. Digital Video Broadcasting (DVB); SNMP MIB for test and measurement applications in DVB systems http://www.etsi.org/deliver/etsi_ts/102000_102099/102032/01.01.01_60/ts_102032v010101p.pdf.

8. Исхаков С.Ю., Шелупанов А.А., Мещеряков Р.В., Беляков К.О., Ширшин В.А., Шум А.Л., Сарьян В.К. Гибридная система встраивания интерактивных услуг в цифровой телевизионный поток // Телекоммуникации. 2015. № 1. С. 11–19.

9. Мещеряков Р.В., Росошек С.К., Шелупанов А.А., Сонькин М.А. Криптографические протоколы в системах с ограниченными ресурсами // Вычислительные технологии. 2007. Т. 12. № S1. С. 51–61.

10. Росошек С.К., Мещеряков Р.В., Шелупанов А.А., Бондарчук С.С. Встраивание криптографических функций в систему связи с ограниченными ресурсами // Вопросы защиты информации. 2004. № 2 (65). С. 22–25.

11. Andrea Zanella ; Nicola Bui ; Angelo Castellani ; Lorenzo Vangelista ; Michele Zorzi Internet of Things for Smart Cities // IEEE Internet Of Things Journal, Vol. 1, No. 1, February 2014, 22–32 pp.

12. Мещеряков Р.В. Критерий структурной сложности информационных систем // Труды СПИИРАН. 2010. № 3 (14). С. 76–90.

13. The Internet of Things Creating a Hyperconnected World https://www.unit4.com/blog/2016/08/the-internet-of-things-creating-a-hyperconnected-world

14. Mohsen Marjani, Fariza Nasaruddin, Abdullah Gani, Ahmad Karim, Ibrahim Abaker Targio Hashem, Aisha Siddiqa, And Ibrar Yaqoob Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges // IEEE Access, March 29, 2017.

15. J M Blythe; S D Johnson The consumer security index for IoT: A protocol for developing an index to improve consumer decision making and to incentivize greater security provision in IoT devices // Living in the Internet of Things: Cybersecurity of the IoT – 2018, 2018, Page(s):1–7.

УДК 004.891.2

Ю.В. Петриченко, О.И. Сушкова

ВНУТРЕННЯЯ МИГРАЦИЯ МОЛОДЕЖИ В СОВРЕМЕННОЙ ЭКОНОМИКЕ: ИССЛЕДОВАНИЕ ПРИЧИН С ЦЕЛЬЮ СОЗДАНИЯ МОНИТОРИНГОВОЙ СИСТЕМЫ

Аннотация

Работа посвящена исследованию внутренних миграционных потоков молодежи на примере Челябинской области. Наибольшее внимание уделено миграции образованной молодёжи и выявлению причин, побуждающих её покидать место жительства, делая выбор в пользу крупных федеральных центров. В работе использовалась теория Э. Ли о делении факторов влияния на принятие решения о переезде на три категории, видоизменённую гравитационную модель и предположение о сведении их в общую мониторинговую систему для определения степени критичности миграционного положения. В ходе работы были намечены ключевые показатели, которые в дальнейшем будут необходимы для прогнозирования вероятности отъезда молодёжи из одного города в другой. Исследование проводилось на примере челябинской молодёжи в возрасте от 18 до 35 лет.

Ключевые слова: внутренняя миграция, миграция молодёжи, гравитационная модель, факторы влияния на миграцию, мониторинговая система.

Проблемы исследования молодёжной миграции

В ряде зарубежных стран вопрос внутренней миграции давно занимает значимое место при решении вопросов распределения ограниченных ресурсов. Учёные различных направлений, такие как Венхорст В.А., Флоринская Ю.Ф., Колумб С., Нельсен И., проводят исследования, изучая закономерности в процессах перемещений людских потоков для увеличения точности их прогнозирования [1, 2, 3, 4]. Чаще всего основным объектом исследования является трудовая миграция, как наиболее распространенная категория перемещающихся лиц. В лучшем случае переезды людей в возрасте 18–35 лет рассматриваются с точки зрения «образовательной миграции», например, в исследованиях социологов и демографов [5, 6, 7].

В настоящий период в Российской Федерации вопрос внутренней миграции населения только начинает исследоваться в новых реалиях. Развитие сферы услуг, повсеместное использование информационных технологий, переход к цифровой экономике обуславливает притяжение образованной молодёжи и молодых специалистов в федеральные центры. Данная категория считается наиболее мобильной группой населения, пока ещё не обременённой социальными и финансовыми обязательствами, а также более восприимчивой к изменениям. К таким выводам приходят и учёные в ходе своих исследований. Бегство в федеральные центры из периферийных городов молодых людей, на которых должны перекладываться обязательства по развитию города, обрекают последние на застой и деградацию.

На данный момент в России основным инструментом измерения внутренних миграционных потоков служит статистика регистрации граждан по месту жительства. Однако подавляющее большинство молодого населения при переезде не имеет даже временной регистрации и таким образом не учитываются органами статистики. Так, например, официально гражданин РФ может проживать на территории другой области как минимум в течение 90 дней, не предоставляя никаких документов в органы регистрационного учёта. А так как молодёжь зачастую не имеет средств для быстрого приобретения недвижимости, то точно отследить, когда именно прибыл или убыл такой гражданин, не представляется возможным. Это делает миграционные потоки неконтролируемыми и значительно искажает имеющиеся официальные данные, так как отражает только часть мигрантов.

Рассмотрение существующих систем прогнозирования

В качестве базиса для проведения анализа была взята эконометрическая теория Э. Ли, разработанная в 1960-х годах. Согласно ей, на каждой территории действуют различные группы факторов, влияющих на миграцию, которые условно можно разделить на три категории: выталкивающие, удерживающие и притягивающие. Каждую группу, в свою очередь, можно разделить ещё на несколько подгрупп. В теории отмечены и основные тенденции влияния каждого фактора. Например, притягивающие факторы окажут влияние с большей вероятностью на высокообразованных и имеющих чуть более высокий уровень жизни людей. Зачастую для них миграция служит инструментом продвижения по карьерной лестнице или повышения заработка. И наоборот, на население, находящееся на более низком социально-экономическом уровне, гораздо большее влияние будут оказывать выталкивающие факторы.

Однако сама по себе теория Э. Ли не может являться основополагающей в нашем исследовании, т.к. в ней отсутствует формула расчёта привлекательности конкретно взятого города, которая могла бы помочь отследить влияние того или иного фактора на интенсивность входящих и исходящих людских потоков [14].

Подобная формула подробно рассматривается в гравитационной модели миграции, разработанной в шестидесятых годах прошлого столетия и служащей основой для большинства социологических исследований нашего времени [12, 13]. В стандартной формулировке в ней говорится о прямо пропорциональной зависимости величины миграционных потоков между двумя населёнными пунктами и расстоянием между ними. Формула имеет следующий вид:

$$Mij = Pi*Pj/Dij^2*A$$

где: Mij – прогнозируемое число мигрантов из региона i в регион j; Pi – численность населения региона i; Pj – численность населения региона j; Dij – расстояние между регионами i и j; A – коэффициент соответствия.

Некоторые современные исследователи предлагаю модернизировать стандартную формулу, и вместо показателя «расстояние» использовать показатель «стоимость переезда из места А в место Б». Такой подход также может быть не совсем верен, так как большое расстояние от первичного места жительства до места планируемого переезда может являться, скорее, удерживающим психологическим фактором. При этом, в формуле не учитываются другие важные финансовые и социально-значимые показатели.

Для получения наиболее полной картины и осуществления эксперимента по слиянию теории о факторах влияния на принятие решения о переезде и гравитационной модели проведено анкетирование среди молодёжи города Челябинска в возрасте от 18 до 35 лет, уже получившей высшее образование или находящейся на финальной стадии его получения.

B. Экспериментальное анкетирование молодёжи Челябинска

Всего в исследовании приняли участие более 800 человек. Респондентами стали молодые люди в возрасте от 18 до 35 лет, студенты или выпускники вузов Челябинска.

Данные, полученные в ходе эксперимента

A. Притягивающие факторы

В ходе исследования был сделан вывод: притягивающие факторы актуальны только в случае прогноза количества мигрантов для двух конкретных регионов, т.к. отражают привлекательность региона притяжения. Для нас же интерес представляла именно Челябинская область.

В связи с этим, теория Э. Ли была подвергнута изменениям: притягивающие факторы были исключены из расчётов, а коэффициенты весомости значений были распределены между удерживающими и отталкивающими факторами в равных долях. Удерживающие факторы брались с положительным знаком, отталкивающие – с отрицательным.

B. Удерживающие факторы

Были выявлены следующие удерживающие от смены места жительства факторы:

1. Отсутствие финансовых возможностей для переезда в ближайшее время отметили 12 % всех опрошенных. Эквивалентом для оценки этого фактора может послужить стоимость переезда из одного город в другой, например, ближайший федеральный центр. Данный показатель учитывается и в европейских исследованиях, где отмечается его удерживающее влияние на принятие решения о переезде.

2. Наличие недвижимого или движимого имущества. О нём упоминают 10 % респондентов. Для оценки этой характеристики можно учитывать среднюю стоимости квадратного метра жилья исследуемого города, а в случае углубления расчётов посчитать среднюю стоимость покупки и эксплуатации автомобиля. Согласно открытым статистическим данным, чем выше популярность города, тем выше стоимость жилья на рынке.

3. Ещё одним удерживающим фактором оказалось соотношение между заработными платами компаний городов. 10 % респондентов ответили, что уже имеют конкретное предложение от работодателей на текущем месте пребывания.

Помимо указанных факторов было выяснено, что основным удерживающим фактором для молодёжи является психологический, т.е. социальная привязка к текущему мест жительства, наличие там близких людей. При этом, данное заключение применимо как к тем, для кого Челябинск является родным городом, так и для тех, кто приехал в Челябинск на учебу после окончания школы из области.

Также мы предположили причинно-следственную связь между упомянутыми факторами: зависимость от семьи и налаженных социальных связей отчасти обуславливается неполной финансовой самостоятельностью, которой подвержена большая часть наших респондентов.

C. Выталкивающие факторы

К категории выталкивающих факторов было отнесено следующее:

1. 55 % опрашиваемых ответили, что их не устраивают экологические условия жизни в городе. Критическая обстановка в городе активно освещается в СМИ на протяжении последних лет. Эквивалентом этого фактора можно считать позицию города в рейтинге благополучия России, составленном на основе оценок местного населения.

2. 30 % респондентов отметили отсутствие возможности для профессионального роста в Челябинске. Как и в случае с зарплатами, этот фактор может стать как притягивающим в другой регион, так и отталкивающим для текущего. Количество респондентов, выбравших этот вариант ответа, значительно превалирует над количеством тех, кто указал, что уже имеет конкретное предложение по трудоустройству на текущем месте жительства. С учётом возрастной категории исследуемой когорты населения, этот фактор весьма показателен, так как показывает одну из наиболее приоритетных проблем для молодёжи.

Таблица 1

Миграция Челябинской области 2002–2017

2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
-1491	-854	-939	646	1601	2740	3914	1569	2511	9974	8786	9499	11002	3746	2704	-3840

3. В эту категорию было отнесено и недостаточное разнообразие культурных мероприятий, которую замечают 24 % опрошенных. Её можно характеризовать количеством проводимых мероприятий в год и/или количество активнодействующих мест для крупных мероприятий.

4. Наконец, городской средой недовольны около 16 %. Сюда респонденты относят и дорожные ситуации, и решения жилищных вопросов, и отношения с компаниями, обслуживающими город. Реальным показателем этого может стать оценка деятельности губернатора региона. Кроме того, использование современных технологий, позволяющих отслеживать не только количество упоминаний конкретной темы, но и эмоциональный окрас каждого из таких упоминаний, может помочь оценить соотношение настроений жителей.

Можно заключить, что главными отталкивающими факторами, которые воздействуют на молодёжь Челябинска, является социальная ситуация в регионе: экологическая обстановка, политика местных властей, бедность культурно-массовой деятельности и невозможность профессиональной реализации на сформированном рынке труда.

Результаты исследования

Согласно проведенному нами исследованию, в качестве наиболее привлекательных районов для переезда выступают Центральная Россия (42,7 %), Южная Россия (27 %) и г. Екатеринбург (22,8 %). Результаты ожидаемы, а потому гипотеза была сформирована ещё до начала сбора данных и подтверждена в его ходе: Центральная Россия и Екатеринбург, ближайший уральский центр, притягивают молодёжь своими возможностями в разных сферах и большей восприимчивостью к современным изменениям. Юг же обещает более приятный для жизни климат.

Одним из утверждений, полученных в ходе исследования, стало наблюдение о возрастной градации отношения к освещаемой теме. Чем старше респондент, тем более осознанно он подходит к вопросу переезда. Например, более зрелые опрашиваемые реже отмечали южные регионы как привлекательные, ссылаясь на их недостаток возможностей для трудоустройства квалифицированной рабочей силы. При этом, они упоминают и высокую стоимость жизни на той территории. Также эти респонденты чаще признаются в недостатке крупных средств для переезда в ближайшее время, что может говорить о более серьёзном отношении к данному вопросу и наличии семьи.

Выявленные факторы легли в основу расчёта индекса критичности, который в модели гравитации эквивалентен коэффициенту различия. Как уже было отмечено, в формулу были взяты только удерживающие и отталкивающие факторы с соответствующим значением и коэффициентами в 0,5, в качестве гипотезы о равной силе влияния обеих групп факторов.

В качестве идеальной для города ситуации можно считать ту, при которой значение индекса критичности будет положительным, т.е. удерживающие факторы будут превалировать над отталкивающими. В противном случае можно предполагать, что в исследуемом городе существует повышенная вероятность массовой миграции молодёжи.

На основе проведённого исследования значения в формуле принимают следующий вид:

0,5 * (0,12 + 0,1 + 0,1) - 0,5 * (0,55 + 0,3 + 0,24 + 0,16) = 0,5 * 0,32 - 0,5 * 1,25 = 0,16 - 0,625 = -0,465

Даже при учёте небольших размеров выборки (1 000 респондентов) заметно, что значение индекса непросто не положительно, а достаточно значимо отрицательно. Можно предположить, что миграционная обстановка в городе носит критический характер.

Это можно увидеть и на примере статистических данных по миграции за прошедшие шестнадцать лет.

В 2017 году количество миграционного прироста впервые за более, чем 10 лет стало отрицательным, что означает снижение миграционной привлекательности населения региона. При этом, необходимо учитывать, что в официальных источниках отражены только те факты миграции, которые имеют регистрационное подтверждение. Вероятной же причиной такого резкого снижения как раз может стать увеличение исходящей миграции, т.е. эмиграции из Челябинской области.

В некоторых статистических исследованиях миграции отмечают также необходимость контроля такого явления как медианный возраст населения. В отличие от среднего, медианный возраст делит всё население чётко на две половины: нижняя половина населения младше значения, верхняя – старше.

Рис. 1. Медианный возраст населения Челябинской области

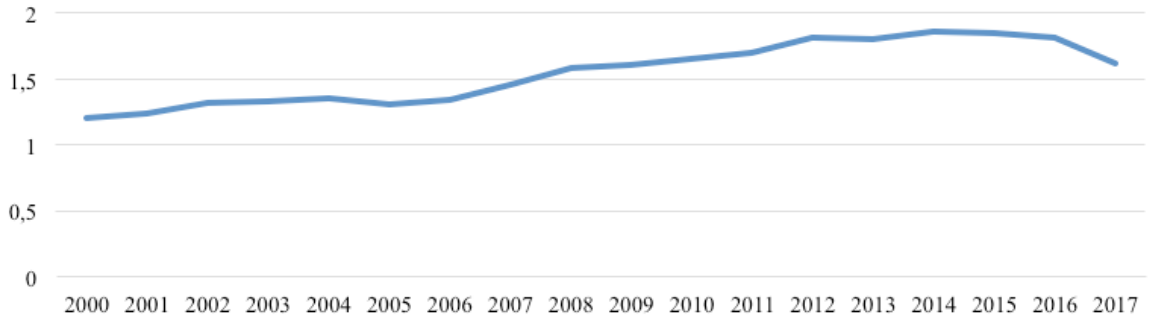


Рис. 2. Коэффициент рождаемости в Челябинской области

По графику видно, как за последние восемнадцать лет медианный возраст постепенно увеличивался, т.е. происходит процесс устаревания населения. Это явление также может быть связано с постепенным выездом за пределы области молодёжи и семей, имеющих несовершеннолетних детей. Такое предположение может подтвердить и график суммарных коэффициентов рождаемости. Он показывает, что количество новорождённых постепенно растёт.

Также в ходе исследования для сравнения официальной статистики были рассмотрены показатели соседних крупных центров: Екатеринбурга, Тюмени, Перми и Кургана. Нами было выявлено, что сбор статистической информации в регионах не унифицирован, что делает процесс сравнения невозможным. Даже информация о значениях миграционного прироста в каждом регионе рассчитывается по-разному.

По полученным данным можно увидеть, что снижение миграционного прироста за последний год отмечается во всех областях. Однако в большинстве случаев это снижение плавное, не такое резкое, какое было зафиксировано в Челябинской области в этот же период.

Создание мониторинговой системы на основе полученных данных

Все полученные в ходе исследования данные можно уложить в мониторинговую систему, которая позволила бы оценивать ситуацию в режиме реального времени.

Помимо уже выявленных факторов влияния, их состав можно будет видоизменять и увеличивать при помощи перманентно проводимого анкетирования среди населения. Если заложить в основу системы предложенную формулу, в результате можно получить не только значение индекса критичности, но и определить зоны риска.

Также такая система сможет затронуть все группы населения, анализируя не только результаты анкетирования и общую статистику, но и более целевые показатели: количество уезжающих на обучение в другие города школьников, рейтинг города по уровню счастья, показатели экологически опасных выбросов, среднюю заработную плату на рынке, количество открытых вакансий, количество проживающих в городе самозанятых предпринимателей и пр.

Главное, что необходимо учесть при использовании уже полученных результатов, в ходе текущего исследования были определены не все факторы влияния. Работа будет продолжаться, и, в первую очередь, основные усилия будут направлены на расширение спектра факторов и уточнение использования коэффициентов.

Выводы

В ходе исследования был проведён анализ миграционной ситуации Челябинской области с двух сторон: по данным официальной статистики и по ответам анкетирования на предмет миграционных настроений. При помощи такого синтеза удалось не только заметить и подтвердить критичность текущий ситуации, но и наметить предполагаемые сферы развития, которые в перспективе позволят сократить количество уезжа-

Таблица 2

Сравнение показателей миграционного прироста

	2010	2011	2012	2013	2014	2015	2016	2017
Челябинск	2 370	6 572	4 660	4 115	5 740	3 366	2 675	-3 840
Екатеринбург	-7 299	13143	6 742	1 884	4 581	1 524	326	214
Курган	-6 962	-9 952	-8 620	-6 824	-5 281	-5 536	-4 845	-5 113
Пермь	-3 826	-807	1 910	22	-1 058	-4 014	-3 201	-6 029
Тюмень	8 997	21 000	19 700	17 989	11 340	17 734	18 144	17 545

ющих. Модернизированная гравитационная модель, которая, в свою очередь, в дальнейшем послужит основой для разрабатываемой мониторинговой системы, помогла определить общую канву распределения коэффициентов значимости факторов влияния, выявленных в ходе анкетирования. Дальнейшее расширение и углубление исследования поможет дополнить уже намеченные факторы и уточнить спектр сфер, требующих развития, которые так или иначе были упомянуты при анализе полученных в анкетировании ответов.

Список литературы

1. Rafiq, S (Rafiq, Shuddhasattwa); Nielsen, I (Nielsen, Ingrid); Smyth, R (Smyth, Russell), 2013, The Encyclopedia of Global Human Migration Published, “Effect of internal migration on the environment in China”.

2. Mkrtchyan, NV (Mkrtchyan, N., V) ; Florinskaya, YF (Florinskaya, Yu F.), 2018, “Labor Migration in Russia: International and Internal Aspects”, “Трудовая мобильность населения России по данным всероссийской переписи 2010 года”.

3. Beine, M (Beine, Michel), Coulombe, S (Coulombe, Serge), 2018, “Immigration and internal mobility in Canada”.

4. Venhorst, VA (Venhorst, Viktor A.), 2013, “Graduate migration and regional familiarity”.

5. Ислакаева Г.Р., 2016, «Образовательная межрегиональная миграция: причины и экономические следствия» https://cyberleninka.ru/article/n/obrazovatelnaya-mezhregionalnaya-migratsiya-prichiny-i-sotsialno-ekonomicheskie-sledstviya.

6. Волох В.А., 2015, «Миграция населения как объект научного исследования: сущность, современные трактовки и классификация». https://cyberleninka.ru/article/n/migratsiya-naseleniya-kak-obekt-nauchnogo-issledovaniya-suschnost-sovremennye-traktovki-i-klassifikatsiya.

7. O. Busch, B. Weigert, “Where have all the graduates gone? Internal cross-state migration of graduates in Germany in 1984–2004”, Annals of regional science, 2010, vol. 44, issue 3, pp. 559–572.

8. J. van Geel, V. Mazzucato, “Conceptualising youth mobility trajectories: thinking beyond conventional categories”, Jornal of ethnic and migration studies, 2018, vol. 44, issue 13, pp. 2144–2162.

9. A. Bernard, M. Bell, “Educational selectivity of internal migrants: A global assessment”, Demographic research, 2018, vol. 39, art. 29.

10. Biagi B. (Biagi, Bianca), Faggian A. (Faggian, Alessandra); McCann P. (McCann, Philip), 2011, “Long and Short Distance Migration in Italy: The Role of Economic, Social and Environmental Characteristics”.

11. Busch O. (Busch, Oliver), Weigert B. (Weigert, Benjamin), 2010, “Where have all the graduates gone? Internal cross-state migration of graduates in Germany 1984–2004”.

12. JE Anderson, “Theoretical foundation for the gravity equation”, American Economic Review, 1979, vol. 69, issue 1, pp. 106–116.

13. A.Arpaia, A. Kiss, B. Palvolgyi, A. Turrini, “The effects of European integration and the business on migration flows: a gravity analysis”, Review of World Economics, 2018, vol. 154, issue 4, pp. 815-834.

14. Huttunen K. (Huttunen, Kristiina), Moen J. (Moen, Jarle), Salvanes K.G. (Salvanes, Kjell G.), 2018, “Job Loss and Regional Mobility”.

15. Tarasyev A.A. (Tarasyev, Alexandr A.), Agarkov G.A. (Agarkov, Gavriil A.), Hosseini S.I. (Hosseini, Seyed Iman), 2017, “Machine Learning in Labor Migration Prediction”.

16. JJ. Lewer, H. Van den Berg, “A gravity model of immigration”, Economics Letters, 2008, vol. 99, issue 1, pp. 164–167.

17. Gryshai V. (Gryshai, Vladimir), Gafiatulina N. (Gafiatulina, Natalya), Kasyanov V. (Kasyanov, Valery), Velikodnaya I. (Velikodnaya, Irina), Kosinov S. (Kosinov, Sergei), Lyubetsky N. (Lyubetsky, Nikolay), Samygin S. (Samygin, Sergey), 2018, “Social health of youth in the context of migration processes in russia: assessment of the threat to national security”.

18. O.I. Nikonov, A.A. Tarasyev, “Dynamic modelling of multi-regional migration processes: Ural Federal District case study”, IFAC-Papers online, vol. 48, issue 25, pp. 45-49, published 2015.

19. S. Sardadvar, E. Vakulenko, “A model of interregional migration under the presence of natural resources: theory and evidence from Russia”, Annals of Regional Science, 2017, vol. 59, issue 2, pp. 535–569.

20. Челябинская область в цифрах, 2018, unpublished.

УДК 004.4'2

Е.В. Деркач

ФУНКЦИОНАЛЬНОЕ ТЕСТИРОВАНИЕ ИСПОЛНИТЕЛЬНОГО ЯДРА ПРОГРАММНО РЕАЛИЗОВАННОГО КОНТРОЛЛЕРА НА ЭТАПЕ РАЗРАБОТКИ ПРОГРАММНОЙ СОСТАВЛЯЮЩЕЙ

Аннотация

Статья посвящена практическим аспектам процесса тестирования систем логического управления на этапе разработки программного обеспечения. В качестве системы логического управления рассматривается инновационный продукт в области автоматизации технологических процессов – программно реализованный логический контроллер (SoftPLC). Рассмотрены: особенности работы контроллера при его программной реализации, его отличие от традиционных ПЛК, состав прикладного программного обеспечения и его архитектурная модель. В архитектуре программно реализованного контроллера выделена среда программирования, работающая в машинном времени и ядро контроллера, работающее в режиме жесткого реального времени. Проведена систематизация видов тестирования и построена интеллект-карта. Среди видов тестирования выделено функциональное тестирование, которое определяет соответствие разработанного программного обеспечения исходным функциональным требованиям заказчика. Рассмотрены практические аспекты функционального тестирования программно реализованного контроллера, в частности разработан набор тестовых сценариев в среде проектирования и отладки программ логического управления FBEditor (разработка кафедры компьютерных систем управления, ФГБОУ ВО «МГТУ «СТАНКИН») на языке функциональных блоков (FB – functional block). Набор тестовых сценариев охватывает тестирование логических функций, реализуемых контроллером, и может работать в автоматическом и полуавтоматическом режимах.

Ключевые слова: автоматизация, система логического управления (СЛУ), тестирование, программируемый логический контроллер (ПЛК), программно реализованный контроллер, прикладное программное обеспечение, среда разработки, интеллект-карта, функциональное тестирование, тестовый сценарий.

Введение

Применение вычислительной техники в автоматическом управлении – важнейшая черта технической инфраструктуры современного общества. Промышленность, транспорт, системы связи и защиты окружающей среды существенно зависят от компьютерных систем управления. Особенности работы систем управления сложными объектами требует применения специализированных средств, которые, в отличие от универсальных ЭВМ, обладают функциональной компонентой, прямо ориентированной на процессы управления объектами в режиме реального времени [2]. Такими системами являются системы логического управления (СЛУ), реализованные на базе программируемых логических контроллеров (ПЛК).

Промышленные контроллеры широко применяются в машиностроении, станкостроении, в области связи, в химической промышленности, используются в автоматизации складов, производстве продуктов питания, а также в энергетике, в частности, в системах противоаварийной защиты, управления технологическими процессами, в системах диспетчеризации, телемеханики и учета энергоресурсов, в системах распределения сетей среднего напряжения, для автоматизации блочно-модульных котельных установок и т.д. [9], [10], [11].

Программно реализованный логический контроллер

Новым направлением в области автоматизации стало развитие программно реализованного логического контроллера (в англоязычной литературе SoftPLC), который имеет ряд преимуществ по сравнению с традиционным ПЛК:

- SoftPLC не является дополнительным оборудованием, поэтому техническая поддержка и сопровождение осуществляется совместно с обслуживанием программно-аппартного обеспечения системы управления в целом;
- программно реализованный логический контроллер является программно-математическим обеспечением в рамках общей вычислительной системы, а значит, имеет возможность тесного взаимодействия как с другими задачами управления, так и со всеми модулями системы [17];
- возможность быстрой модернизации системы без длительной остановки и наладки аппаратной компоненты СЛУ за счет обновления программного обеспечения;
- при программной реализации появляется возможность диагностики, установки обновлений и устранения ошибочных ситуаций посредством удаленной работы через Internet [4], [6], [7], [8].

Одним из вариантов реализации SoftPLC является программно реализованный промышленный контроллер, разработанный на кафедре «Компьютерные системы управления» в ФГБОУ ВО «МГТУ «СТАНКИН». Структура контроллера включает следующие

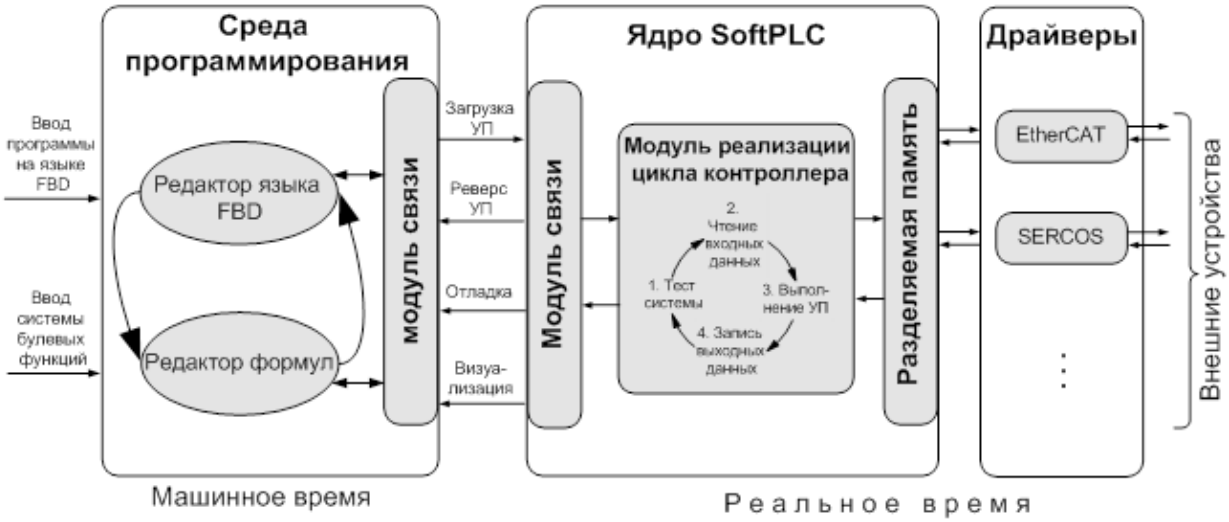


Рис.1. Архитектура программно реализованного логического контроллера

компоненты (рис. 1): среду разработки, где на основе технического задания (ТЗ) формируется управляющая программа (УП) и конфигурация оборудования; ядро, где происходят вычислительные преобразования модели; драйверы внешних устройств, с помощью которых командные сигналы передаются непосредственно на физические устройства (исполнительные механизмы объекта управления) [3], [5].

В состав прикладного программного обеспечения терминальной части входит:

• среда разработки и отладки программ логического управления FBEditor, основным языком программирования которой является язык FBD – Function Block Diagram (входит в стандарт МЭК 61131-3);

• среда конфигурирования оборудования, предназначенная для настройки основных параметров доступа к аппаратным средствам с помощью драйверов внешних устройств и одного из стандартных протоколов связи (например, SERCOS или EtherCAT).

В настоящее время в ФГБОУ ВО «МГТУ «СТАНКИН» продолжается разработка программно реализованного контроллера. В связи с этим возникает необходимость многократного проведения тестовых испытаний как программной составляющей, так и системы логического управления в целом.

Виды тестирования

По данным из разных источников процесс тестирования в среднем занимает от 30 до 60 % от общей трудоемкости всего проекта в целом [1], [12], [14]. В печатной литературе и электронных источниках существует огромное количество вариантов систематизации существующих методов тестирования. На рисунке 2 приведена интеллект-карта классифика-

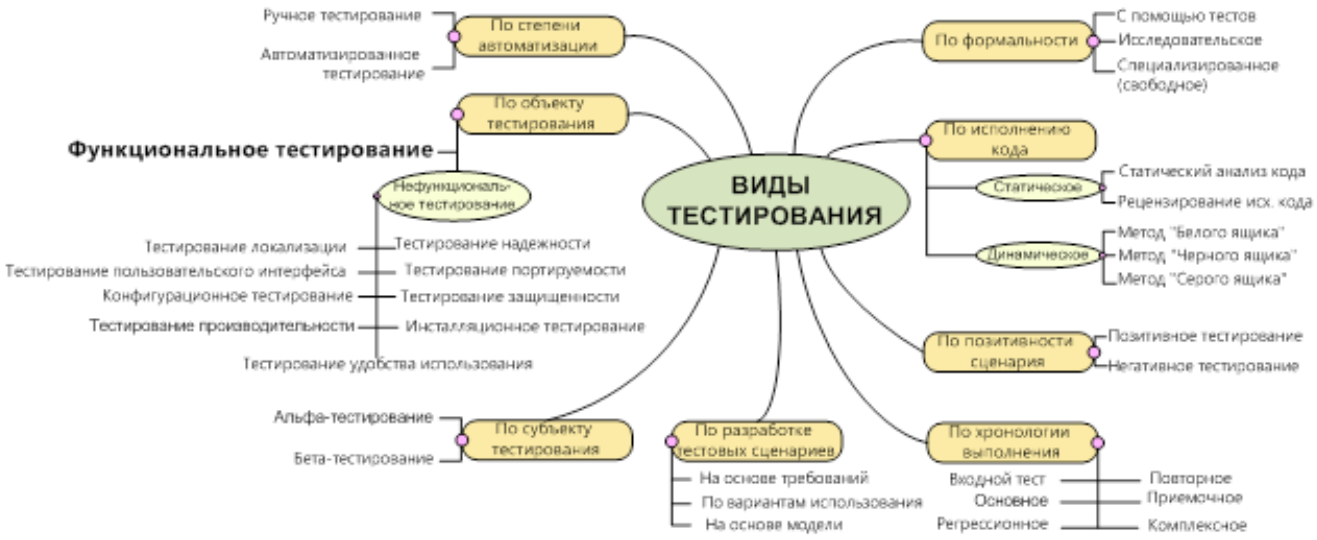


Рис.2. Классификация видов тестирования

ции видов тестирования по различным показателям. Предложенная структура не является единственной. Почти каждая группа способов тестирования, объединенная по какому-либо признаку, имеет, в свою очередь, по множеству разнообразных разбиений по особенностям тестирования, которые в той или иной мере могут пресекаться между собой.

Разные виды тестирования применяются для различных целей, и в каждом конкретном случае рациональнее всего использовать не весь спектр существующих тестов, а только такой набор, который обеспечит необходимое качество ПО и его соответствие выдвигаемым требованиям.

Функциональное тестирование

Одним из ключевых видов тестирования является функциональное тестирование, задача которого – проверить способность программного обеспечения (ПО) в определённых условиях решать задачи, нужные пользователям. Как правило, полноценное функциональное тестирование является самым трудоемким процессом и может занимать до 90 % от общего времени тестирования [16]. Поэтому остановимся на нем более подробно.

Процесс разработки ПО носит итеративный характер, который включает в себя как добавление вновь разработанных, дополнительных, модулей, так и корректировку уже существующих. Программа за время жизни претерпевает многочисленные изменения [13]. В разных моделях технологии программирования понятие жизненного цикла определяется и трактуется немного по-разному, но, в общем, близко к схеме, представленной на рисунке 3. Важно подчеркнуть, что за время своей жизни программный продукт неоднократно проходит метаморфозы, т.е. это именно цикл, который повторяется несколько раз. Это, в свою очередь, приводит к тому, что появляется необходимость проверки на наличие ошибок не только в новых блоках программного продукта, но и там, где уже было проведено многократное тестирование. Поэтому процедура тестирования должна проводиться каждый раз после очередного выпуска программного продукта.

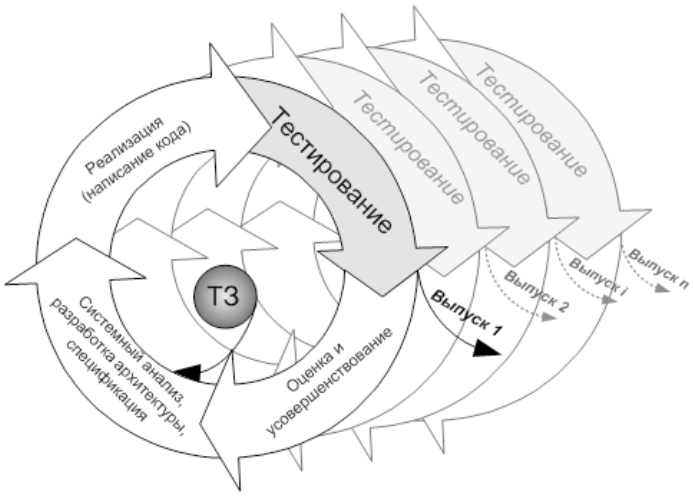


Рис. 3. Итеративная модель жизненного цикла программного обеспечения

Выпуск – это характерная точка в жизни программного продукта, отмечающая прохождение одного полного цикла. Для этого используется набор тестовых сценариев, которые могут работать как в автоматическом, так и в ручном режиме [15].

Создание тестовых сценариев

Рассмотрим разработку набора тестовых сценариев на примере тестирования ядра программного реализованного логического контроллера. Являясь, по существу, программным обеспечением, оно находится в процессе постоянной доработки существующего функционала и добавления новых функций. Тестирование будем проводить в среде FBEditor на языке программирования FBD.

Язык FBD является графическим языком и наиболее удобен для программирования процессов прохождения сигналов через функциональные блоки. Программа на языке FBD представляет собой совокупность функциональных блоков (functional flocks, FBs), которые соединяются линиями связи (connections). Каждый блок представляет собой математическую операцию (сложение, умножение, триггер, логическое «или» и т.д.). Начальные значения переменных задаются с помощью специальных блоков – входов или констант, выходные цепи могут быть связаны либо с физическими выходами контроллера, либо с глобальными переменными программы.

Для осуществления процесса тестирования составляется управляющая программа на языке FBD, которая, в общем случае, состоит из основной программы, хранящейся в файле с расширением «*.fbv», и множества библиотек пользовательских подпрограмм, хранящихся в файлах с расширением «*.fbl». В ходе работы системы, после проведения ряда системных проверок, программа логического управления объединяется с входящими в нее пользовательскими подпрограммами (в своем изначальном виде), происходит компоновка программы. В подсистему логического управления загружается объединенная управляющая программа и после успешного прохождения еще одного этапа проверок производится запуск программы в цикле работы логического контроллера.

В качестве примера реализации автоматического функционального тестирования используем логические блоки среды FBEditor, а конкретно, операцию логического умножения AND. Учитывая бинарный характер входов и выходов (принимают значения либо 0, либо 1) логических операций, а значит и не слишком большой набор возможных сочетаний входов-выходов блока, имеем возможность автоматизировать процесс тестирования логических блоков программно реализованного логического контроллера.

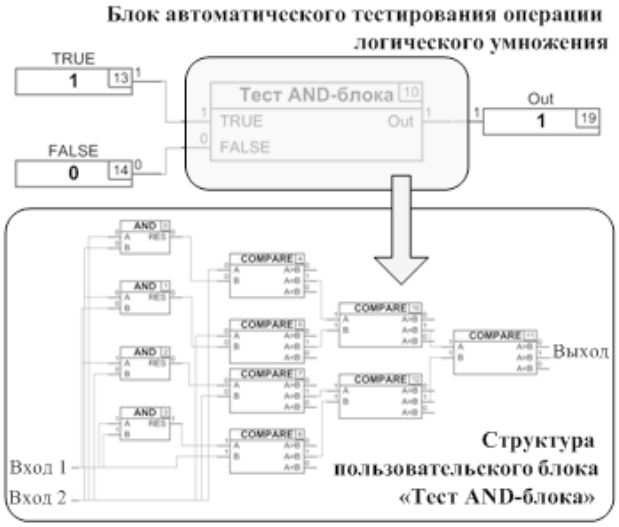


Рис. 4. Тестирование AND-блока

Рис. 5. Тестирование OR-NOT-блока

Рис. 6. Тестирование AND-NOT-блока

Рис. 7. Тестирование блока ИСКЛ-ИЛИ

Для этого в среде программирования однократно составляется управляющая программа из логических блоков, объединенная в единый пользовательский блок, которому назначаются определенные входы и выходы. Структура его показана на рисунке 4, и в дальнейшем он может быть использован как самостоятельный элемент. Для инициализации процесса тестирования необходимо однократно подать на вход блока два управляющих сигнала и получить значение на выходе. Успешное тестирование фиксируется при достижении логической 1 на выходе пользовательского блока (рис. 4). Если же значение на выходе равно 0 – тестовый сценарий считается непройденным, блок работает с ошибкой.

Аналогичным образом тестируются логические блоки AND-NOT (И-НЕ), OR-NOT (ИЛИ-НЕ), XOR (ИСКЛ-ИЛИ) (рис. 5–7).

Важно отметить, что автоматическое тестирование предполагает минимальное участие оператора в процессе тестирования. Это позволяет уйти от рутинной работы и исключить «человеческие» ошибки, связанные, например, с неправильной интерпретацией действия той или иной команды. При этом не требуется абсолютно никаких знаний внутреннего содержания тест-блока, нет необходимости в высокой квалификации оператора или владения какими-либо специальными знаниями.

Для примера реализации ручного тестирования рассмотрим тестирование блоков арифметических операций среды FBEditor, а именно, сложение, вычитание и умножение. Тестирование сложения и вычитания выполним в едином блоке путем различных комбинаций этих двух операций, а для умножения составим отдельный (рис. 8, 9). В данном случае использование ручного тестирование является наиболее приемлемым, так как значения входов и выходов блоков арифметических операций имеют большое количество граничных значений и классов эквивалентности (положительные, отрицательные числа, ноль и их комбинации).

Тестирование в ручном режиме, по сравнению с автоматическим, дает больше возможностей и вариантов для более полной и всесторонней проверки работы блоков. Однако здесь потребуются минимальные знания и навыки оператора, понимание некоторых основ работы в среде, а также займет больше времени.

Для выполнения всестороннего тестирования формируется набор тестовых сценариев, фрагмент которого приведен в таблице 1, которые последовательно отрабатываются оператором.

На входы программы логического управления (A, B, C, D на рис. 8, 9) подаются значения из таблицы (табл. 1), которые обрабатываются в соответствии с заданным алгоритмом. Затем выходное значение сравнивается с заранее рассчитанным (вход T_V). Если они совпадают, то на выходе управляющей программы фиксируется 1 – тестирование пройдено успешно. В противном случае – если на выходе получается 0 – тест не пройден, в работе блока фиксируется наличие ошибки.

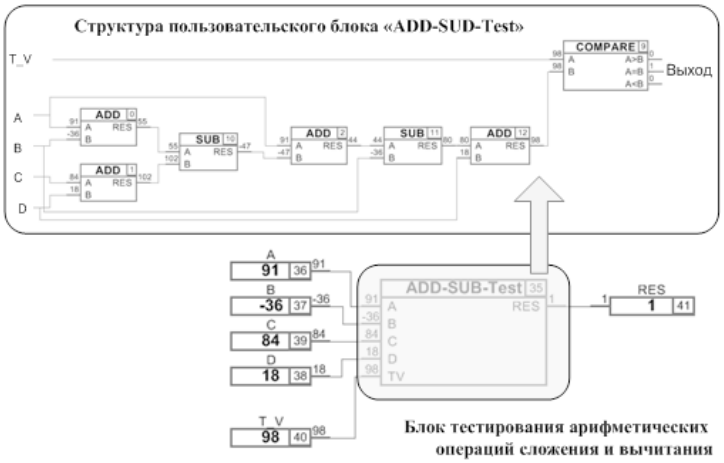


Рис. 8. Тестирование блоков арифметического сложения и вычитания

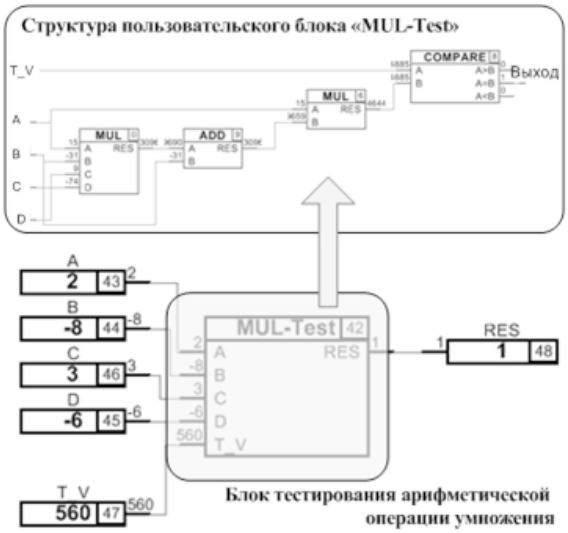


Рис. 9. Тестирование блока арифметического умножения

Таблица 1

Пример создания тестовых сценариев для ADD-SUB-Test и MUL-Test (фрагмент)

ADD-SUB-Test						MUL-Test				
A	28	521	3	609	91	2	21	3	-11	91
B	63	-124	51	225	-36	-8	-124	51	-18	-36
C	17	715	-14	850	84	3	-15	-14	-469	84
D	42	-359	-92	476	18	-6	108	-92	-75	18
T_V	39	327	20	368	98	560	88 585476	591 345	-76 610 952	-450 754 668

Выводы

Формализован подход к проведению тестирования систем логического управления на этапе проектирования программного обеспечения. Сформулированы варианты тестовых сценариев для тестирования в ручном и автоматическом режимах на примере функционального тестирования ядра программно реализованного логического контроллера.

Предложенный подход дает возможность значительно сократить количество ошибок, возникающих на начальной фазе разработки программного продукта. Описанные тестовые сценарии позволяют выявить несоответствия разработанного ПО заявленным функциональным требованиям заказчика, исправление которых на более поздних этапах разработки системы логического управления может стать причиной поломки дорогостоящего оборудования, либо привести к значительным экономическим затратам.

Список литературы

1. Гленфорд Майерс. Искусство тестирования программ, 3-е издание. / Гленфорд Майерс, Том Баджетт, Кори Сандлер. – Москва: Диалектика, 2012. – 272 с.

2. Мартинов Г.М. Современные тенденции развития компьютерных систем управления технологического оборудования / Г.М. Мартинов // Вестник МГТУ «Станкин». – 2010. – № 1. – С.119–125.

3. Мартинов Г.М. Декомпозиция и синтез программных компонентов электроавтоматики / Г.М. Мартинов, Н.В Козак. // Приборы и системы. Управление, контроль, диагностика. – 2006. – № 12. – С. 4–11.

4. Мартинов Г.М. Развитие систем управления технологическими объектами и процессами / Г.М. Мартинов // Мир техники и технологий. – 2009. – № 6. – С. 34–35.

5. Мартинов Г.М. Специфика построения редактора управляющих программ электроавтоматики стандарта МЭК 61131 / Г.М. Мартинов, Р.А. Нежметдинов, П.А. Никишечкин // Вестник МГТУ «СТАНКИН». – 2014. – № 4 (31). – С. 127–132.

6. Минаев И.Г., Программируемые логические контроллеры [Текст]: практическое руководство для начинающего инженера / И.Г. Минаев, В.В. Самойленко. – Ставрополь: АГРУС, 2009. – 100 с. – ISBN 978-5-9596-0609-1.

7. Нежметдинов Р.А., Повышение эффективности функционирования электроавтоматики станков с ЧПУ на основе реализации регулярных моделей архитектуры программно реализованных контроллеров типа SoftPLC: диссертация на соискание ученой сте-

пени кандидата технических наук : 05.13.06 / Нежметдинов Рамиль Амирович; [Место защиты: Моск. гос. технол. ун-т «Станкин»]. – Москва, 2009. – 159 с.: ил. РГБ ОД, 61 09-5/2856

8. Нежметдинов Р.А. Подход к построению систем логического управления технологическим оборудованием для реализации концепции «Индустрия 4.0» / Р.А. Нежметдинов, П.А. Никишечкин, И.А. Ковалев, Н.Ю. Червоннова // Автоматизация в промышленности. – 2017. – № 5. – С. 5–9.

9. Официальный сайт интернет портала «МирПром» [электронный ресурс]: офиц. сайт // интернет портал «МирПром». – Режим доступа: http://mirprom.ru/public/sovershenstvovanie-sistem-chpu.html (дата обращения 15.01.2017).

10. Официальный сайт компании «Балт-Систем» [электронный ресурс]: офиц. сайт // компания «Балт-Систем». – Режим доступа: https://www.bsystem.ru (дата обращения 03.07.2017).

11. Официальный сайт фирмы «3S – Smart Software Solution» [электронный ресурс]: офиц. сайт // компания «3S – Smart Software Solution». – Режим доступа: http://www.codesys.com (дата обращения 03.10.2014).

12. Рекс Блэк. Ключевые процессы тестирования. Планирование, подготовка, проведение, совершенствование. / Рекс Блэк. – Москва: Лори, 2006. – 544 с.

13. Селевцов Л.И. Автоматизация технологических процессов: учебник для студ. учреждений сред. проф. образования / Л.И. Селевцов, А.Л. Селевцов. – 3-е изд., стер. – М.: Издательский центр «Академия», 2014. – 352 с.

14. Сэм Канер. Тестирование программного обеспечения. Фундаментальные концепции менеджмента бизнес-приложений: Пер. с англ. / Сэм Канер, Джек Фолк, Енг Кек Нгуен. – Москва: ДиаСофт, 2001. – 544 с.

15. Тюленева И.А., Новоселова О.В. Исследование методов тестирования программного обеспечения по уровню, объектам и автоматизации тестирования // Студенческий форум: электрон. научн. журн. 2017. № 21(21). URL: https://nauchforum.ru/journal/stud/21/29564 (дата обращения: 31.01.2019).

16. Фредерик Брукс мл. Мифический человеко-месяц или как создаются программные системы. 2-е издание, юбилейное. Addison-Wesley Longman, Inc. (ISBN 0-201-83595-9), 1995. Издательство Символ-Плюс (ISBN 5-93286-005-7), 2000, 2005.

17. Шемелин, В.К. Применение технологии клиент-сервер при проектировании контроллера типа SoftPLC для решения логической задачи в рамках систем ЧПУ / В.К. Шемелин, Р.А. Нежметдинов // Автоматизация и современные технологии. – 2010. – № 3. – С. 31–37.

УДК 517.988.6 338.23

С.В. Костюченко, Н.А. Черемисин, А.Н. Емельянов

ЛОКАЛИЗАЦИЯ ТЕКУЩИХ ИЗВЛЕКАЕМЫХ ЗАПАСОВ НЕФТИ И ПРЯМОЙ РАСЧЕТ ОХВАТА ВЫТЕСНЕНИЕМ В ГЕОЛОГО-ТЕХНОЛОГИЧЕСКИХ МОДЕЛЯХ НЕЛИНЕЙНОЙ ФИЛЬТРАЦИИ

Аннотация

В статье описан комплекс методов интеллектуального управления извлечением запасов нефти: локализация текущих извлекаемых запасов, расчет коэффициента охвата запасов нефти вытеснением, нефтяных пластов. Объектами исследований являются месторождения на поздних стадиях разработки и с трудно-извлекаемыми запасами. Особенность методов – применение цифровых гидродинамических моделей с нелинейными законами фильтрации пластовых флюидов в пластах. В статье обобщено понятие коэффициент охвата вытеснением и выделены накопленный и текущий коэффициент вытеснения, описан метод динамического расчета относительных фазовых проницаемостей, приведены примеры карт распределения запасов, охваченных и неохваченных вытеснением.

Ключевые слова: зрелые месторождения нефти, локализация текущих запасов нефти, коэффициент охвата вытеснением, динамические фазовые проницаемости, нелинейная фильтрация, цифровые гидродинамические модели.

Введение

Стратегические планы развития нефтедобывающих компаний России, в т.ч. ПАО «НК «Роснефть» связаны не только с освоением новых нефтегазовых месторождений, но и с интенсификацией разработки и стабилизацией добычи жидких углеводородов «зрелых» месторождений, с разработкой залежей нефти с трудноизвлекаемыми запасами (ТРИЗ).

Разработка зрелых месторождений осложнена высокой обводненностью скважин; несбалансированностью отборов, закачки и заводнения; бессистемными малоэффективными ГТМ. Внедрение цифровых технологий должно обеспечить реалистичные расчеты 3D-распределений текущих извлекаемых запасов нефти, количественный анализ и реализацию адресных программ управления заводнением: новых очагов заводнения, контроль компенсации отбора закачкой и потоков пластовых флюидов; охват вытеснением текущих запасов нефти, извлечение и доизвлечение запасов.

Для расчета уровней добычи нефти, коэффициентов извлечения и 3D-распределений запасов УВС применяют цифровые геолого-технологические модели. Модели дают детальную информацию для формирования технологических и геологических карт, для обоснования проектных решений по разработке месторождений и для выбора способов извлечения и доизвлечения запасов УВС.

Вместе с тем, современные методы создания моделей и интерпретации результатов моделирования имеют немало нерешенных проблем и нереализованных возможностей. Может быть, решение этих проблем и есть один из реальных способов интеллектуализации разработки месторождений и извлечения запасов нефти? В этой статье авторы дают свой ответ на этот вопрос.

Локализация текущих запасов нефти с применением моделей нелинейной фильтрации. В геолого-технологических моделях, основанных на линейных законах фильтрации Дарси, коэффициент извлечения нефти (КИН) слабо зависит от плотности сетки скважин и от систем разработки, а коэффициент охвата вытеснением для таких моделей близок к единице. Для этих моделей характерна невысокая подтверждаемость (особенно на поздних стадиях разработки) расчетных карт текущих запасов нефти по результатам бурения уплотняющих скважин и боковых стволов к ним. Поэтому и проектные решения по разработке месторождений УВС и совершенствования систем разработки на основе таких моделей могут быть не самыми лучшими.

Альтернативой таким моделям являются модели с потоками пластовых флюидов, отличающихся от линейного закона Дарси [1–6]. Суть процессов такой нелинейной фильтрации:

1. При малых градиентах пластовых давлений (grad P) скорость фильтрации флюидов нелинейно зависит от grad P, существует начальный градиент сдвига (НГС);

2. При малых скоростях фильтрации нефть в пористой среде м.б. неподвижна, а охват вытеснением может стремиться к нулю;

3. Объемы реальных нефтяных залежей с малой скоростью фильтрации и с низким охватом вытеснением, – м.б. велики.

Эти принципы позволяют создавать реалистичные модели потоков пластовых флюидов, распределений текущих запасов нефти и их охват процессами вытеснения.

Известны программные системы – симуляторы, в которых имеются возможности создания таких моделей (табл. 1). Вместе с тем, моделирование процессов фильтрации с отклонениями от закона Дарси в современных симуляторах пока затруднено или невозможно и поэтому требует развития.

Таблица 1

Симуляторы с возможностями моделирования потоков пластовых флюидов, отличающихся от линейного закона Дарси

Симулятор	Метод расчета	Примечания
STARS	Относительные фазовые проницаемости – функции от капиллярного числа	Зависимость от капиллярного числа – интерполяция дискретных наборов данных
Eclipse300 tNavigator		Реализовано только для ОФП по нефти и газу
MORE Eclipse100	Вязкость нефти – функция компонент grad P	Зависимость от grad P только для нефти с неньютоновскими свойствами
РН-КИМ	Поправки к абсолютной проницаемости – функция от grad P	Отсутствует независимое задание нелинейности для фильтрующихся фаз
FrontSim	Не реализован	

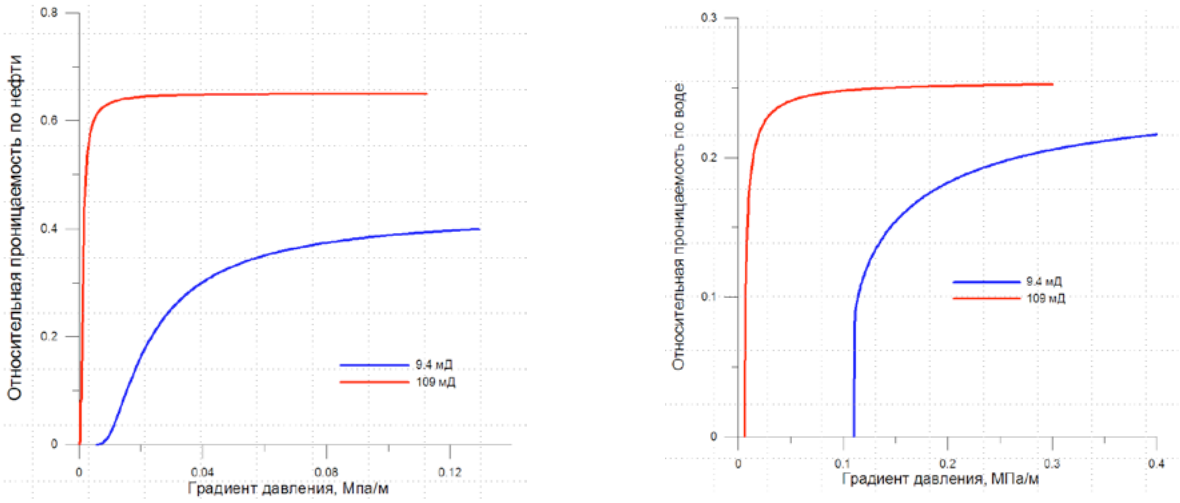
Одним из перспективных методов моделирования таких фильтрационных процессов является метод динамического расчета относительных фазовых проницаемостей (ОФП) [1, 3, 5, 6]: критической нефтенасыщенности, фазовых проницаемостей нефти и воды – как функций капиллярного числа *N* и скорости фильтрации:

$$N = \frac{k}{\sigma} * \left| gradP \right| = \frac{\mu}{\sigma} * \left| V \right|,$$

где *grad(P)* – градиент давлений, Па/м; *k* – проницаемость воды, м^2; *V* – скорость фильтрации, м/с; М – динамическая вязкость, Па*с; σ – поверхностное натяжение на границе раздела фаз, Н/м. На рисунке 1 приведен пример таких зависимостей, рассчитанных по методике [6]. Динамические свойства ОФП (т.е. их зависимость от *grad(P)*) показывают, что область нелинейности и величина НГС смещается в сторону увеличения градиентов давления с уменьшением проницаемости коллектора.

Метод динамического расчета ОФП может быть реализован с использованием любого из известных «линейных» симуляторов: РН-КИМ, Eclipse, FrontSim и др. Технология их применения предполагает итерационный пересчет ОФП и критических насыщенностей по специальному алгоритму с использованием процедур рестартов этих симуляторов. При этом симуляторы рассчитывают потоки флюидов, поля пластовых давлений и насыщенностей, масштабируют ОФП, а управление временными шагами и итерациями, расчет скоростей фильтрации, капиллярных чисел и «кубов» концевых точек ОФП выполняет специальный программный модуль, внешний для симулятора.

Расчет и анализ текущего охвата вытеснением запасов нефти в настоящее время необоснованно фактически исключен из анализа текущего состояния разработки нефтяных залежей. Коэффициент охвата вытеснением (E_v) – отношение нефтенасыщенного объема продуктивного пласта, охваченного процес-



а) б)

Рис. 1. Зависимости ОФП по нефти при остаточной воде – а) и по воде при остаточной нефти – б) от градиента порового давления

сом вытеснения, ко всему нефтенасыщенному объему пласта в выбранном расчетном контуре – имеет важное прикладное значение. Он входит в формулу: $E_r=E_d*E_v$ (обозначения в табл. 2), количественно характеризует эффективность разработки: текущее состояние и динамику, позволяет сравнить варианты разработки одного или разных объектов, может быть положен в основу построения карт охвата вытеснением текущих запасов и запасов, охваченных и неохваченных вытеснением.

Однако традиционные «линейные» симуляторы не позволяют выделить дренируемые и недренируемые расчетные блоки и поэтому коэффициент охвата вытеснением E_v^{end} приходится рассчитывать «обратным» способом (табл. 2). Рассчитанный таким образом коэффициент охвата характеризует эффективность разработки всей залежи и только на конец разработки – не годится для оценки текущего состояния разработки и динамики охвата вытеснением. По нашему мнению [7], в рамках линейных моделей фильтрации прямой расчет текущего коэффициента охвата вытеснением выполнить в принципе невозможно.

В таблице 2 приведены наши обобщения понятия коэффициента охвата вытеснением. Эти формулы основаны на расчете критической нефтенасыщенности $SOWCR_i$ в каждом расчетном блоке i цифровой модели – как функции от капиллярного числа N и скорости фильтрации методом динамического расчета ОФП. В таблице 2 обозначено: E_v – коэффициент охвата; E_r – коэффициент извлечения нефти; E_d – коэффициент вытеснения; e_v – текущий коэффициент охвата вытеснением; w – текущий коэффициент вовлечения воды; N_p – накопленная добыча нефти; N – геологические запасы нефти; W – геологические запасы воды; M_p – подвижные запасы нефти, охваченные вытеснением; M_w – запасы воды, вовлеченные в разработку; init – начальные; rec – текущие; end – конечные.

Прямой расчет коэффициентов охвата вытеснением E_v, e_v, w_v может быть выполнен по формулам таблицы 2 с применением объемного метода, где для каждого i-блока текущий коэффициент вытеснения

$$e_{v,i} = \frac{S_{oil,i} - SOWCR_i}{S_{oil,i} - SOWCR_0}$$

(если $S_{oil,i}$-$SOWCR_i < 0$, то $e_{v,i}=0$). Аналогично можно рассчитать новые параметры и карты, необходимые для оценки эффективности извлечения и доизвлечения запасов нефти: текущий коэффициент охвата вытеснением, потенциально достижимый КИН и запасы нефти, не охваченные вытеснением (табл. 3 и рис. 4).

На рисунке 3 приведены расчетные модель и карты плотности начальных, текущих и конечных подвижных запасов нефти по состоянию на начало разработки, на 6-й год и конец разработки – при достижении обводненности 98 % всеми скважинами. В линейной модели запасы нефти вырабатываются равномерно, Кохв стремится к 1,0 и равен 0,801 при обводненности скважин 98 %. В нелинейной модели остаточные запасы локализованы в межскважинном пространстве и Кохв=0,593. Срок разработки составляет 79 лет, что на 50 лет меньше, чем в линейной модели. Расчетный КИН в линейной модели составляет 0,463, в нелинейной – 0,250. На рисунке 2 приведена динамика накопленного и текущего коэффициентов охвата вытеснением E_v^{rec}, коэффициента извлечения нефти, e_v – текущего коэффициента охвата вытеснением для этой модели.

Таблица 2

Обобщение определения коэффициента охватавытеснением

Наименование параметра	Расчетная формула
Накопленный коэффициент охвата на конец разработки [1]	$E_v^{end} = \frac{E_r^{end}}{E_d} = \frac{N_p^{end}}{N^{init}} \cdot \frac{1}{E_d} = \frac{N_p^{end}}{N^{init} \cdot E_d}$
Накопленный коэффициент охвата вытеснением извлекаемых запасов нефти на текущую дату [2]	$E_v^{rec} = \frac{N_p^{rec} + M_p^{rec}}{N^{init} \cdot E_d}$ [3]
Текущий коэффициент охвата вытеснением извлекаемых запасов [2]	$e_v = \frac{M_p^{rec}}{N^{rec} \cdot E_d}$
Текущий коэффициент вовлечения воды в разработку [2]	$w_v = \frac{M_w^{rec}}{W^{rec}}$

Примечания: [1] E_v^{end} *- традиционный "коэффициент охвата вытеснением", рассчитываемый "обратным" способом;* [2] E_v^{rec}, e_v, w_v *– предложено авторами;* [3] $E_v^{rec}=E_v^{end}$ *в конце разработки, когда* M_p^{rec} *стремится к нулю.*

Таблица 3

Параметры и карты для оценки эффективности извлечения запасов нефти

Параметры	Традиционные	Предлагаемые
1. Запасы нефти	Запасы нефти: начальные, текущие и остаточные	Запасы нефти: охваченные, неохваченные вытеснением
2. Коэффициент извлечения нефти (КИН)	КИН текущий и конечный	Потенциально достижимый КИН для реализованной системы разработки и для текущего коэффициента охвата вытеснением
3. Коэффициент охвата вытеснением	Накопленный коэффициент охвата на конец разработки	Коэффициенты охвата вытеснением: 1. Накопленный и текущий на заданную дату; 2. Текущий на заданную дату и его динамика; 3. Коэффициент вовлечения воды в разработку
5. Карты	Карты: плотности текущих подвижных запасов нефти	1. Карты плотности текущих геологических, извлекаемых запасов нефти; охваченных, неохваченных вытеснением; 2. Карта охвата запасов нефти вытеснением; 3. Карта вовлечения запасов воды в разработку

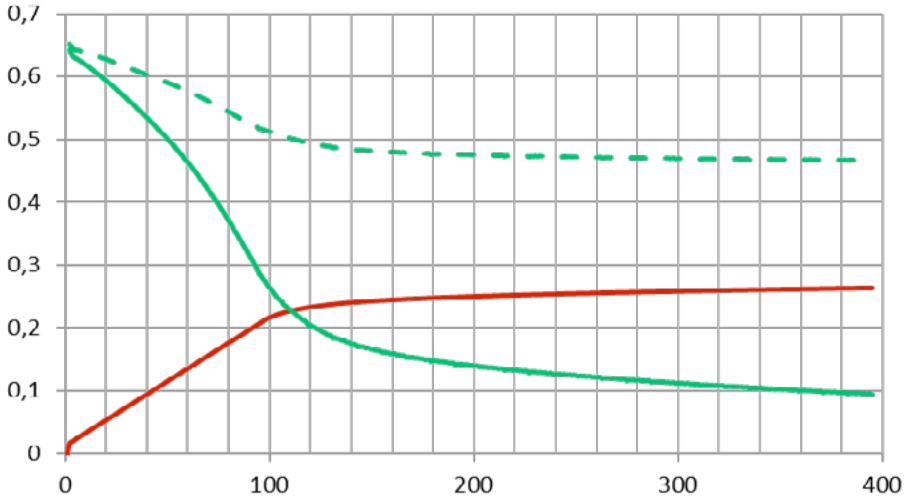
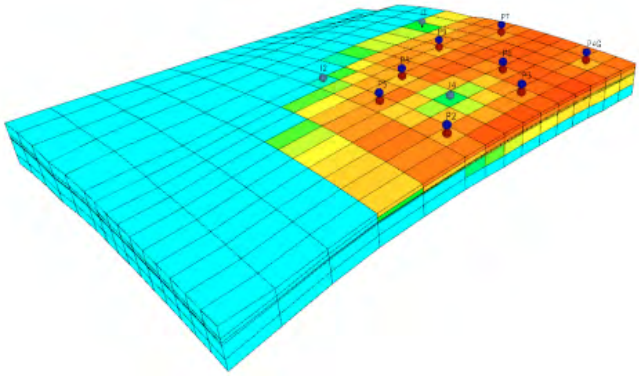
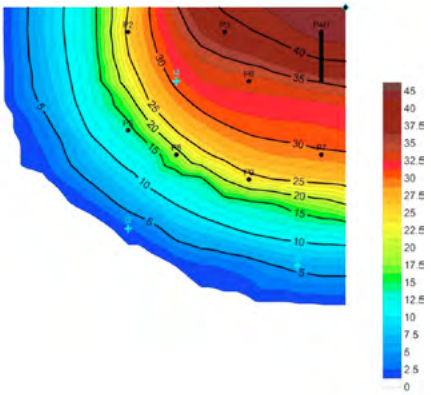


Рис. 2. Динамика накопленного и текущего коэффициентов охвата вытеснением:
1 – текущий коэффициент охвата вытеснением, e_v;
2 – сумма накопленного и текущего коэффициентов охвата вытеснением, Evrec;
1 – коэффициент извлечения нефти.

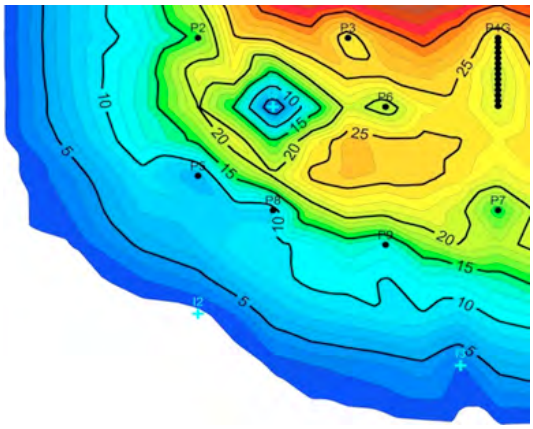
а) Расчетная модель:



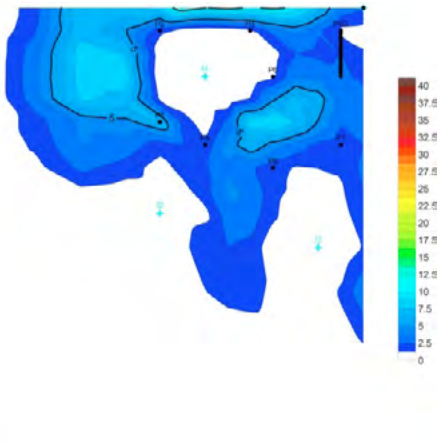
б) Начальная нефтенасыщенность:



Линейная фильтрация:

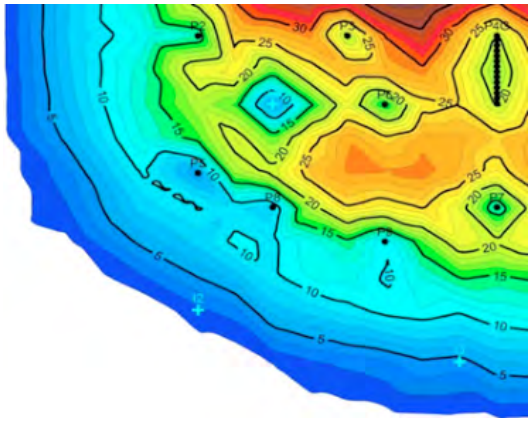


в)

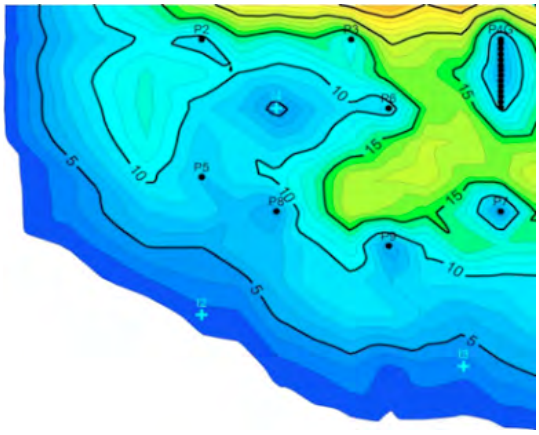


д) КИН=0.463 – 129 год

Нелинейная фильтрация:



г)



е) КИН=0.250 – 70 год

Плотность подвижных запасов,тыс.т/га

Рис. 3. Результаты моделирования процессов извлечений запасов нефти при линейном законе фильтрации Дарси и при отклонениях от него: расчетная модель – а), плотности подвижных запасов нефти: начальная – б), через 6 лет разработки – в) и д), на конец разработки – г) и е)

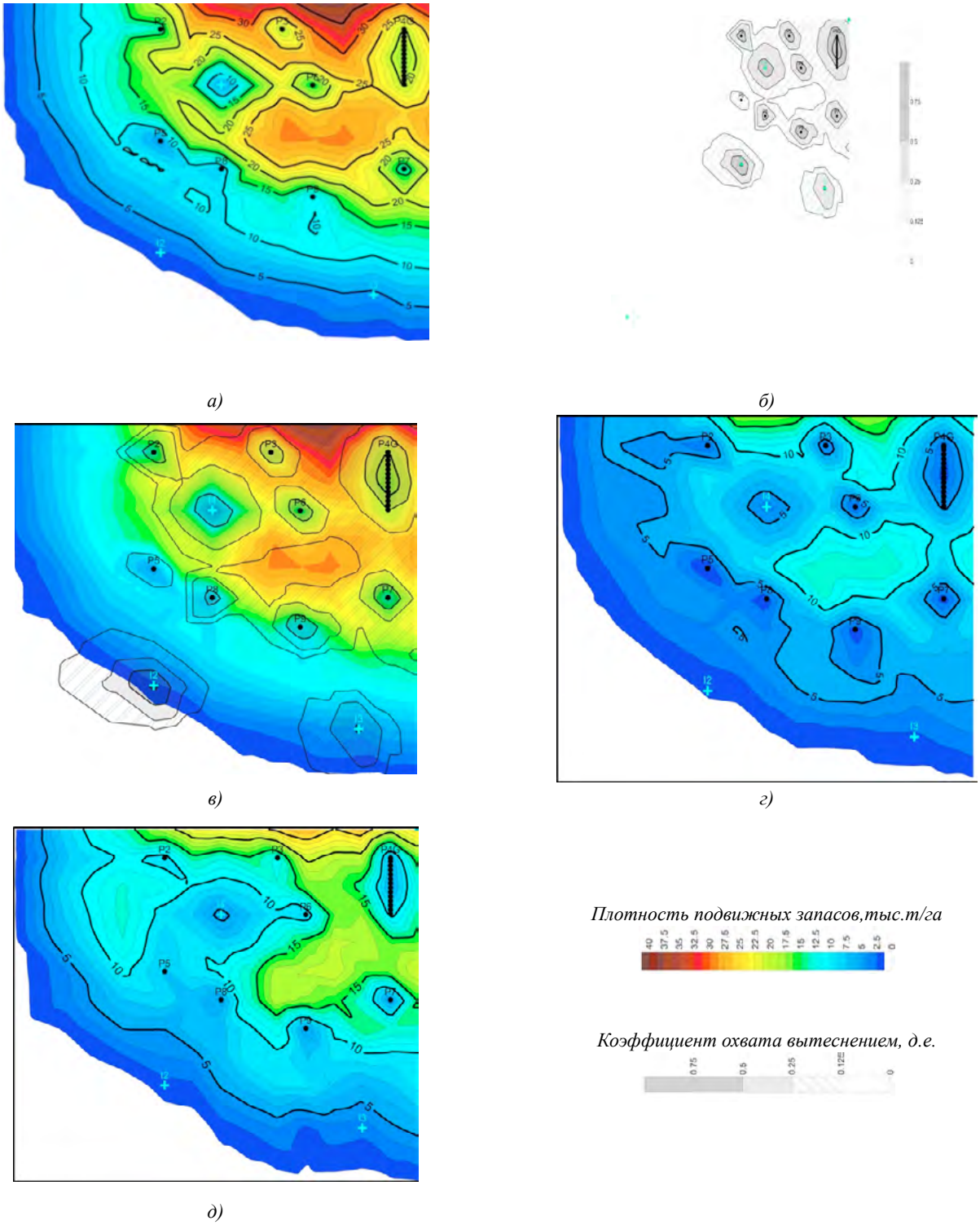


Рис. 4. Примеры карт охвата вытеснением:
а) Плотность подвижных запасов нефти на 6-й год разработки, б) охват вытеснением,
в) запасы с наложенным охватом вытеснения, г) не охваченных вытеснением
д) прогноз плотности остаточных подвижных запасов на 79-й год разработки

Выводы

1. Проблема адекватной локализации текущих подвижных запасов и эффективное управление извлечением и доизвлечением запасов нефти зрелых месторождений и с ТРИЗ может быть решена в рамках концепции: скорость фильтрационного течения пластовых флюидов нелинейно зависит от градиента давления, остаточная нефтенасыщенность коллектора зависит от плотности сетки скважин и темпов отбора.

2. Для цифровых моделей нелинейной фильтрации понятие текущего коэффициента охвата вытеснением может быть существенно обобщено: становится возможным выполнять прямой расчет накопленного, текущего коэффициента охвата вытеснением и их динамики, формировать новые карты охвата вытеснением текущих запасов нефти и карты запасов нефти, не охваченных процессами вытеснения.

3. Цифровые геолого-технологические модели необходимо создавать с учетом результатов интерпретации трассерных исследований и гидропрослушивания, с учетом результатов изучения фактической интерференции скважин.

Список литературы

1. S.M.P. Blom, J.Hagoort How to Include the Capillary Number in Gas Condensate Relative Permeability Functions. 1998. SPE 49268.

2. Черемисин Н.А., Сонич В.П, Батурин Н.Е., Медведев Н.Я. Физические основы повышения эффективности разработки гранулярных коллекторов. Нефтяное хозяйство, 2002, № 8, с. 38-42.

3. Попков В.И., Зацепина С.В., Шакшин В.П. Использование зависимости относительных фазовых проницаемостей от капиллярного числа в задачах трехмерного гидродинамического моделирования залежей нефти и газа. Матем. моделирование, 2005, том 17, номер 2, с. 92–102

4. Байков В.А. и др. Нестационарная фильтрация в сверхнизкопроницаемых коллекторах при низких градиентах давлений. Нефтяное хозяйство. 2013.№ 10. с. 52–56.

5. Михайлов Н.Н., Полищук В.И., Хазигалеева З.Р. Моделирование распределения остаточной нефти в заводненных неоднородных пластах. Нефтяное хозяйство. 2014. № 8. с. 36–39.

6. Костюченко С.В. Прямой расчет коэффициента охвата вытеснением при геолого-гидродинамическом моделировании. «Нефтяное хозяйство», 2006, 10, с. 112–115.

7. Костюченко С.В., Зимин С.В. Количественный анализ эффективности систем заводнения на основе моделей линий тока. «Нефтяное хозяйство», 1, 2005, с. 56–60.

8. Черемисин Н.А., Сонич В.П., Батурин Ю.Е., Дроздов В.А. Условия формирования остаточной нефтенасыщенности в полимиктовых коллекторах при их заводнении // Нефтяное хозяйство. – 1997.– № 9. – С. 40–45.

УДК 004.9 620.9. 004.82

Л.В. Массель, А.Г. Массель

ЦИФРОВАЯ ТРАНСФОРМАЦИЯ ЭНЕРГЕТИКИ: МЕТОДЫ И ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ НАУЧНОГО ОБОСНОВАНИЯ СТРАТЕГИЧЕСКИХ РЕШЕНИЙ

Аннотация

Рассматриваются проблемы перехода к цифровой энергетике. Выделяется проблема недостаточного использования результатов, имеющихся в научных организациях энергетического профиля. Обращается внимание на необходимость интеллектуальной поддержки обоснования и принятия стратегических решений по развитию технологической инфраструктуры (цифровой трансформации энергетики). Предлагается подход к решению проблемы интеграции программных средств, информационного обеспечения и интеллектуальных информационных технологий, разработанных в ИСЭМ СО РАН. Рассмотрены архитектура многоагентной интеллектуальной среды, обеспечивающей эту интеграцию, и методический многоагентный подход к ее разработке.

Ключевые слова: цифровая энергетика, интеллектуальные информационные технологии, математическое и семантическое моделирование, управление знаниями, фрактальная стратифицированная модель, многоагентные системы.

Введение

В связи с распространением концепций интеллектуальной (Smart Grid) [1–2] и цифровой [3] энергетики необходимо учитывать, что при решении проблем перехода к интеллектуальным энергетическим системам (ИЭС) и цифровой энергетике выделяют две взаимосвязанные области: технологическая инфраструктура и информационно-телекоммуникационная инфраструктура. Успех создания ИЭС и цифровой трансформации энергетики во многом зависит от успешного применения современных информационных технологий. В свою очередь, говорить о применении последних имеет смысл при наличии развитой современной технологической инфраструктуры. Решения по развитию технологической инфраструктуры, безусловно, относятся к классу стратегических решений. Для обоснования и поддержки принятия таких решений целесообразно привлечение интеллектуальных информационных технологий. Это в первую очередь технологии семантического моделирования и управления знаниями, которые коллектив под руководством авторов развивает и использует для создания интеллектуальных систем поддержки принятия стратегических решений в энергетике. Предложен подход к построению интеллектуальной системы поддержки принятия стратегических решений по развитию энергетики (многоагентной инструментальной среды), интегрирующей разработанные в Институте систем энергетики им. Л.А. Мелентьева (ИСЭМ) СО РАН математические и семантические методы и модели и инструментальные средства их поддержки.

Проблемы перехода к цифровой энергетике

Актуальность научного обоснования стратегических решений по цифровой трансформации энергетики обусловлена необходимостью решения ряда возникающих при этом проблем.

Как отмечается в экспертно-аналитическом докладе Центра стратегических разработок «Цифровой переход в электроэнергетике России» [4], во многих развитых странах мира реализуются сценарии, трансформирующие электроэнергетику на базе клиентоцентричных распределенных архитектур энергосистем (данное направление принято называть energy transition – энергетический переход). Данный переход ориентирован на масштабное использование распределенной возобновляемой энергетики, на вовлечение частных инвестиций и формирование децентрализованных рынков. Он также подразумевает интеллектуализацию инфраструктуры и переход потребителей к активным, просьюмерским моделям поведения: потребители превращаются в поставщиков электроэнергии. Вызовы для российской электроэнергетики имеют свои особенности, что связано с избытком традиционных топливно-энергетических ресурсов, большой и протяженной территорией с низкой плотностью сети населенных мест, специфическими социально-экономическими факторами. Но в то же время они перекликаются с глобальными вызовами и вызовами для стран с близкими условиями. Речь идет о переходе к новой технологической парадигме в электроэнергетике, представляющей организацию энергоснабжения в розничном секторе как экосистему производителей и потребителей энергии, которые беспрепятственно интегрируются в общую инфраструктуру и обмениваются энергией. Такой подход по аналогии осуществляемых взаимодействий также получил название «Интернет энергии» (Internet of Energy) [4].

В то же время существует ряд проблем, возникающих при переходе к цифровой энергетике. Все проблемы, связанные с цифровой трансформацией энергетики, т.е. с реализацией и применением в России интегрированных интеллектуальных энергетических систем в целом и интеллектуальных информационно-телекоммуникационных технологий в частности, можно условно разделить на три класса: когнитивные

и управленческие проблемы, научные и технологические ограничения и кадровые проблемы [3, 5].

В упоминавшемся выше экспертно-аналитическом докладе [4] предлагается, в качестве одной их мер институционального и организационного характера, «создать российское агентство передовых исследований и разработок в сфере энергетики (по аналогии с ARPA-E в США». При этом игнорируются опыт системных исследований энергетики и результаты исследований, как ИСЭМ СО РАН, так и других научно-исследовательских организаций энергетического профиля.

Системные исследования энергетики

Институт систем энергетики им. Л.А. Мелентьева (ИСЭМ) СО РАН является одним из лидеров в области системных исследований в энергетике России [6].

Основным инструментом исследований до последнего времени являлись математическое моделирование и вычислительный эксперимент. В связи с новыми трендами развития российской энергетики (интеллектуальная энергетика (Smart Grid) и цифровая энергетика) большое внимание уделяется развитию и применению интеллектуальных технологий. Реализуемая в России программа «Цифровая экономика» сейчас развивается и дорабатывается. В частности, авторы считают, что в федеральном проекте «Цифровая энергетика» уделяется недостаточно внимания таким направлениям, как интеллектуальная поддержка принятия стратегических решений по развитию технологической инфраструктуры энергетики и обеспечение кибербезопасности критически важных энергетических объектов. Ниже рассмотрим подробнее первое направление. Большую роль в принятии стратегических решений должно играть их научное обоснование, для чего могут быть использованы научные наработки института.

Традиционно в ИСЭМ СО РАН принята иерархическая схема исследований, в которой согласовываются исследования и математические модели топливно-энергетического комплекса (используются экономико-математические модели) и отраслевых систем энергетики (применяются физико-математические модели) (рис. 1). Исследования прогнозирования развития ТЭК выполняются на верхнем уровне с учетом результатов, полученных при исследованиях направлений развития отраслевых систем энергетики на следующих уровнях. Каждому блоку приведенной схемы соответствует совокупность математических методов, моделей и программные комплексы, с помощью которых выполняются вычислительные эксперименты с использованием этих моделей [6].

Для использования результатов этих исследований при обосновании стратегических решений по цифровой трансформации энергетики требуется выполнить формализованную интеграцию программных средств и информационного обеспечения с целью совершенствования иерархической технологии обоснования развития энергетики в целом и ее отраслевых и территориальных составляющих, при этом основное внимание надо уделить разработке программно-информационных интерфейсов между решаемыми задачами в горизонтальном (между отдельными системами энергетики) и вертикальном (системы энергетики – ТЭК – внешние условия) разрезах.

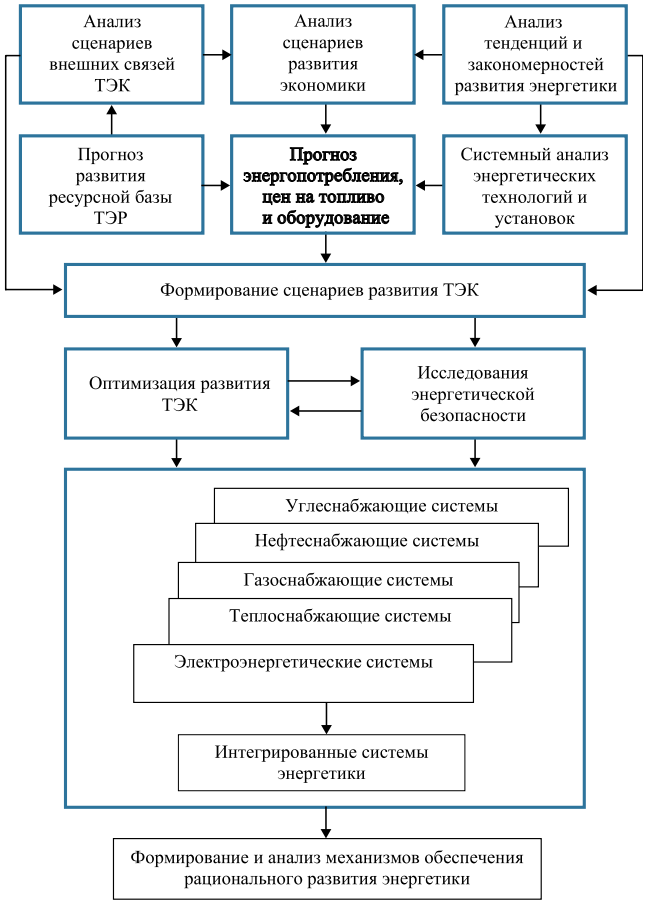


Рис. 1. Общая схема исследований по обоснованию развития энергетики

Предлагаемый подход к решению проблемы интеграции программных средств, информационного обеспечения и интеллектуальных информационных технологий

Реализация предлагаемых интеграционных возможностей может быть обеспечена на базе следующих информационных технологий: а) единой информационно-коммуникационной среды программных компонентов; б) семантической интеграции данных, знаний и программных компонентов; в) инструментальных средств ситуационного управления и семантического моделирования. Предлагается реализация единой информационно-коммуникационной среды взаимодействия программных компонентов в виде облачного сервиса, что позволит обеспечить сетевой доступ по требованию к общему пулу конфигурируемых вычислительных ресурсов (например, к серве-

рам, устройствам хранения данных, приложениям и сервисам и пр.). Для обеспечения необходимого уровня безопасности целесообразно выполнять реализацию информационно-коммуникационной среды в виде корпоративного облака [7]. Для семантической интеграции данных, знаний и программных компонентов предлагается использовать концепцию управления знаниями и применить, как методологический, фрактальный подход к структурированию знаний [8, 9].

Для обоснования стратегических решений по цифровой трансформации энергетики предлагается интегрировать средства математического и семантического моделирования. Ниже рассмотрим понятия и содержание **семантического моделирования** [10–12].

Под **семантической моделью** в обобщенном виде понимается информационная модель, отражающая понятия предметной области и отношения между ними. Авторы рассматривают семантическое моделирование на примере онтологических, когнитивных, событийных и вероятностных (на основе Байесовских сетей доверия) моделей [13–15]. В таблице 1 приведено сравнение технологий семантического моделирования.

Семантические модели разрабатываются на основе знаний экспертов и позволяют использовать как явные, так и неявные знания, основанные на опыте, эрудиции и интуиции экспертов. Например, когнитивные модели, позволяющие отображать причинно-следственные связи, могут быть применены для описания и анализа сценариев внешних связей ТЭК, сценариев развития экономики и развития ТЭК. Событийные и вероятностные модели позволяют рассмотреть варианты развития различных ситуаций, определяемых выбранными сценариями. После экспертной оценки различных вариантов развития с использованием семантических моделей привлекаются традиционные программные комплексы, реализующие математические модели отраслевых систем энергетики и ТЭК и решаются задачи оптимизации для обоснования рекомендуемых решений.

Концепция ситуационного управления используется авторами в соответствии с работами Д.А. Поспелова и его учеников [16]. В последнее время предлагается использовать эту концепцию для оперативного управления, но авторы считают, что возможно и целесообразно применить ее в области обоснования стратегических решений. Используется современная трактовка ситуационного управления, сформулированная в [17]. Применение авторами концепции ситуационного управления для обоснования и поддержки принятия решений по обеспечению энергетической безопасности рассмотрено, в частности, в [18–19].

Для обоснования стратегических решений в энергетике предлагается интегрировать средства математического и семантического моделирования [20]. При этом применяются как базовые технологии: агентно-ориентированные и облачные вычисления, так и проблемно-ориентированные: семантическое и математическое моделирование и **двухуровневая технология исследований,** интегрирующая семантическое и математическое моделирование и поддерживающая ее интеллектуальная ИТ-среда, включающая инструментальные средства семантического моделирования и обеспечивающая возможность интеграции с традиционными программными комплексами [11]. Для схемы, приведенной на рисунке 1, выделяются следующие уровни (этапы) исследований и поддерживающие их инструментальные средства (рис. 2). Интеграция инструментальных средств осуществляется с помощью языка управления знаниями.

Для поддержки принятия стратегических решений в энергетике с использованием предложенного методического подхода и научных прототипов инструментальных средств разработана архитектура многоагентной интеллектуальной среды (МАИС) (рис. 3).

Таблица 1

Сравнение технологий онтологического, когнитивного, событийного и вероятностного моделирования

Технология	Назначение	Аппарат для формализованного представления	Использование в исследованиях энергетической безопасности
Онтологическое моделирование	Для описания декларативных фрагментов знаний	Онтологии (Специальные языки (OWL, RDF, XML и др.))	Для выявления, классификации и спецификации концептов (основных понятий в исследованиях энергетики)
Когнитивное моделирование	Для выявления причинно – следственных связей концептов	Когнитивные карты (теория графов)	Для анализа угроз ЭБ
Событийное моделирование	Построение поведенческих моделей. Выявление динамики развития ЧС	Событийные карты (Теория Joiner-сетей)	Для анализа развития и последствий ЧС
Вероятностное моделирование	Построение вероятностных моделей. Оценка риска реализации угроз	Аппарат Байесовских сетей доверия	Для оценки рисков возникновения ЧС

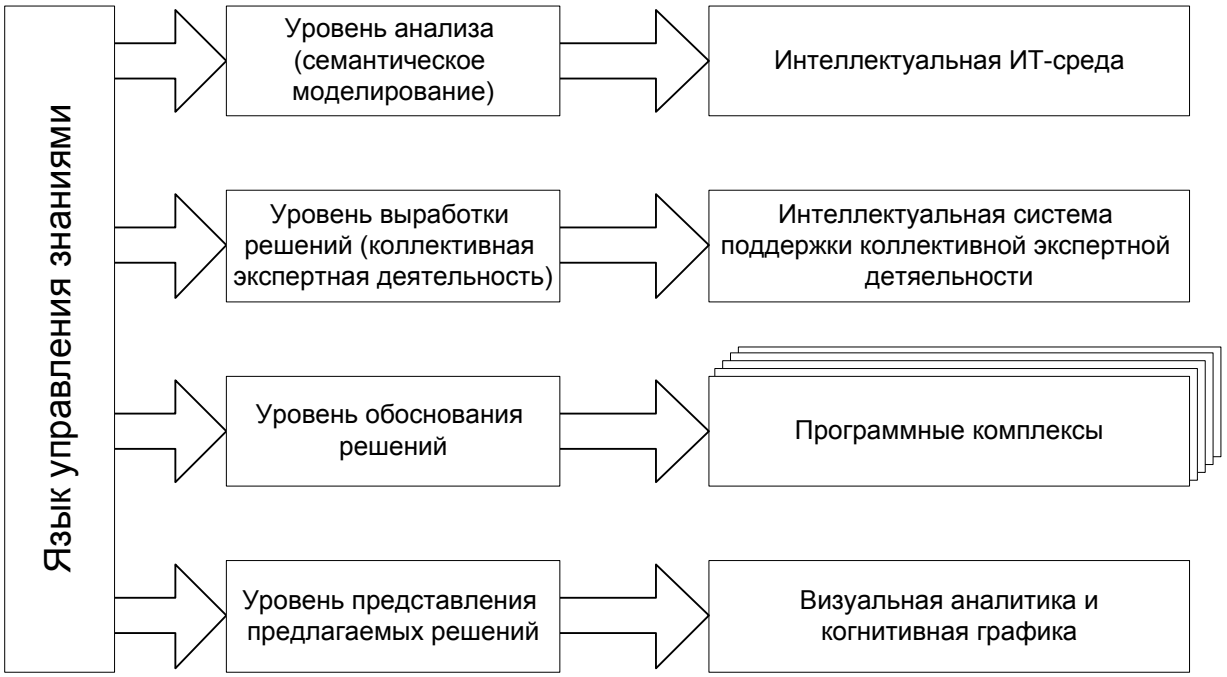


Рис. 2. Уровни (этапы) исследований и поддерживающие их инструментальные средства

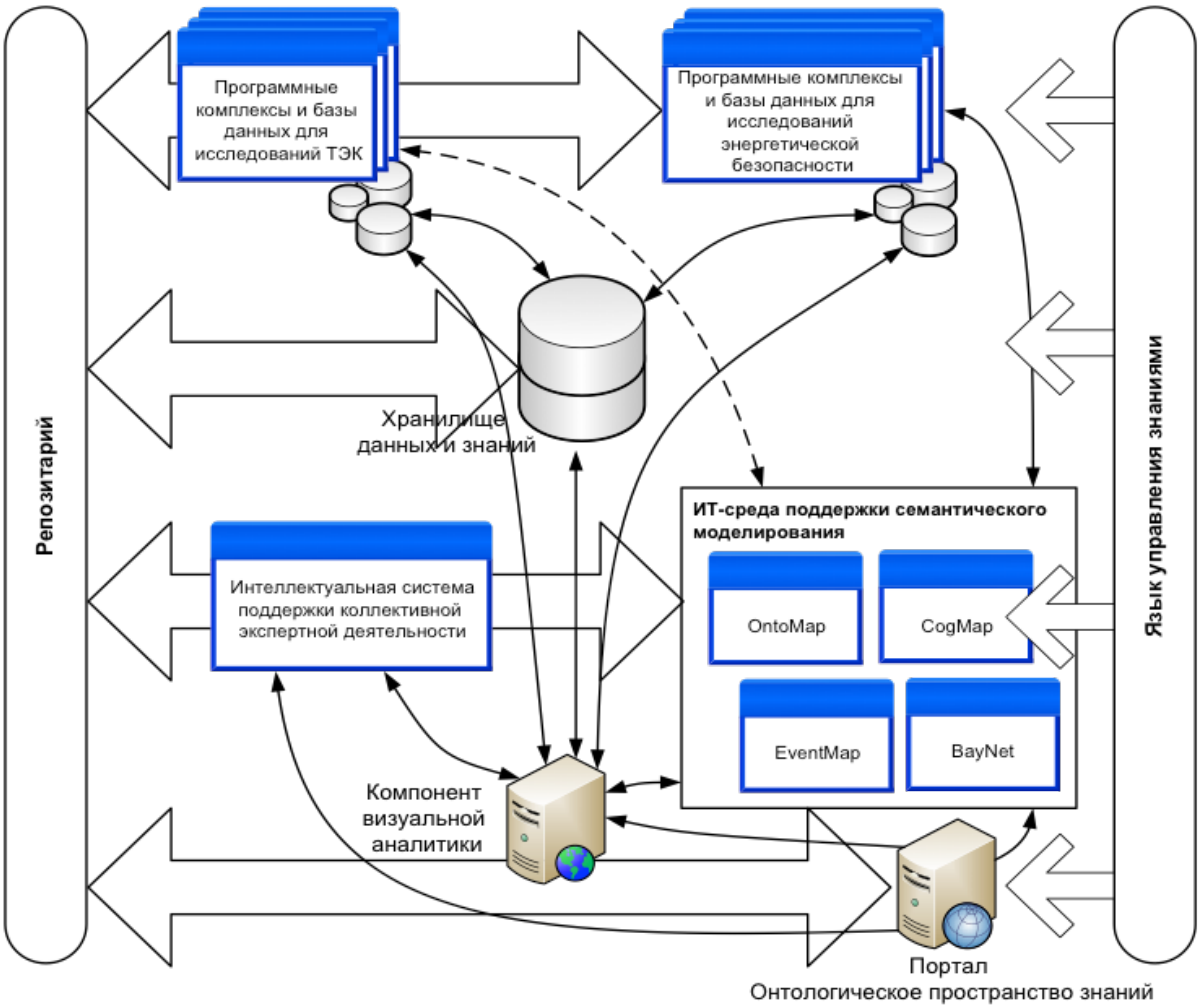


Рис. 3. Архитектура многоагентной интеллектуальной среды (МАИС)

Для реализации МАИС предлагается использовать авторский подход к построению многоагентных систем [22]. Новизна его определяется тем, что предлагается метод управления взаимодействием агентов на основе алгебраических сетей, для реализации которого разрабатываются событийные модели сценариев взаимодействия агентов. Этот подход апробирован, под руководством авторов, при разработке многоагентной системы оценивания состояний электроэнергетических сетей. В настоящее время коллективом, возглавляемым авторами, разработан ряд научных прототипов базовых компонентов, которые могут быть применены при реализации МАИС.

Выводы

Рассмотрены проблемы перехода к цифровой энергетике, одной из которых является, с одной стороны, дефицит научных исследований и доступных прикладных технологических решений в этой области, а с другой, недостаточное внимание к существующим заделам в области исследований энергетики. Отмечается, что в концепции цифровой энергетики не уделяется внимание таким направлениям, как как интеллектуальная поддержка принятия стратегических решений по развитию технологической инфраструктуры энергетики и обеспечение кибербезопасности критически важных энергетических объектов. Предлагается для устранения этих недостатков использовать результаты исследований ИСЭМ СО РАН, направленные на разработку многоагентных интеллектуальных систем для обоснования и поддержки принятия стратегических решений. Авторы полагают, что для их реализации целесообразна интеграция существующих заделов в области математического и семантического моделирования, ситуационного управления, агентных, облачных и интеллектуальных вычислений.

Благодарности

Результаты получены в рамках выполнения проекта по госзаданию ИСЭМ СО РАН № АААА-А17-117030310444-2, отдельные аспекты прорабатывались в рамках проектов, поддержанных грантами РФФИ № 19-07-00351, №18-57-81001 и № 18-07-00714.

Список литературы

1. Кобец Б.Б., Волкова И.О. Инновационное развитие электроэнергетики на базе концепции Smart Grid / М.: ИАЦ Энергия, 2010. – 208 с.

2. Воропай Н.И., Стенников В.А. Интегрированные интеллектуальные энергетические системы // Изв. РАН. Энергетика. – № 1. – 2014. – С. 64–78.

3. Воропай Н.И., Губко М.В., Ковалев С.П., Массель Л.В., Новиков Д.А., Райков А.Н., Сендеров С.М., Стенников В.А. Проблемы развития цифровой энергетики в России // Проблемы управления. № 1. 2019. – С. 2–14. DOI: http://doi.org/10.25728/pu.2019.1.1

4. Экспертно-аналитический доклад «Цифровой переход в электроэнергетике России» https://www.csr.ru/issledovaniya/tsifrovoj-perehod-v-elektroenergetike-rossii/ (дата доступа 10.08.2018).

5. Материалы 2-й отраслевой конференции «Цифровая трансформация электроэнергетики России», Москва, октябрь 2017 : http://digitenergy.ru/ (дата доступа 13.11.2017).

6. Системные исследования в энергетике: Ретроспектива научных направлений СЭИ–ИСЭМ / отв. ред. Н.И. Воропай. – Новосибирск: Наука, 2010. – 686 с.

7. Массель Л.В., Грибова В.В., Копайгородский А.Н. «Облачная» структура энергоинформационных систем // В кн.: Инновационная электроэнергетика-21 / Под ред. Батенина В.М., Бушуева В.В., Воропая Н.И. – М.: ИЦ «Энергия», 2017. – С. 556–577.

8. Массель Л.В. Фрактальный подход к структурированию знаний и примеры его применения / Онтология проектирования. – 2016. – Т. 6, № 2 (20). – С. 149–161. – DOI: 10.18287/2223-9537-2016-6-2-149-161.

9. L.V. Massel. Fractal Approach to Constructing Ontological Knowledge Space / Published in: 2018 3rd Russian-Pacific Conference on Computer Technology and Applications (RPC) Publisher: IEEE. Pp. 1-5. DOI: 10.1109/RPC.2018.8482138. https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=8469127

10. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СпБ: Питер, 2001. – 384 с.

11. Трахтенгерц Э.А. Компьютерная поддержка принятия решений. М.: СИНТЕГ, 1998. – 376 с.

12. Столяров Л.Н. Философия событийного моделирования на примере сценария энергетической катастрофы // Международная конференция «Информационные технологии в науке, образовании, телекоммуникации и бизнесе»: труды. Украина. Гурзуф, 2010. – С. 197–200.

13. Массель Л.В., Массель А.Г. Семантические технологии на основе интеграции онтологического, когнитивного и событийного моделирования // III международная научно-техническая конференция OSTIS-2013: труды. Беларусь. Минск. БГУИР, 2013. – С. 247–250.

14. Массель Л.В., Массель А.Г. Интеллектуальные вычисления в исследованиях направлений развития энергетики // Известия Томского политехнического университета. 2012. – Т. 321. – № 5. Управление, вычислительная техника и информатика. – С. 135–141.

15. Массель Л.В., Пяткова Е.В. Применение байесовских сетей доверия для интеллектуальной поддержки исследований проблем энергетической безопасности // Вестник ИрГТУ. – № 2. – 2012. – С. 13.

16. Поспелов Д.А. Ситуационное управление. Теория и практика. М. Наука. 1986. – 284 с.

17. Васильев В.И., Ильясов Б.Г. Интеллектуальные системы управления. Теория и практика. М. Радиотехника. 2009. – 392 с.

18. Пяткова Н.И., Массель Л.В., Массель А.Г. Методы ситуационного управления в исследованиях проблем энергетической безопасности. Известия Академии наук. Энергетика. – № 4. – 2016. – С. 156–163.

19. L. Massel, A. Massel. Intelligent support tools for strategic decision-making on Smart Grid development / International Conference “Green Energy and Smart Grids” (GESG 2018): Proceedings. Pp. 1–8. DOI: https://doi.org/10.1051/e3sconf/20186902009 https://www.e3s-conferences.org/articles/e3sconf/abs/2018/44/contents/contents.html.

20. Массель Л.В. Интеграция семантического и математического моделирования в исследованиях проблем энергетической безопасности// Международная конференция «Моделирование–2012»: труды. Киев. ИПМЭ НАН Украины, 2012. – С. 270–273.

21. Массель Л.В., Гальперов В.И. Разработка многоагентных систем распределенного решения энергетических задач с использованием агентных сценариев / Известия Томского политехнического университета. – Т. 326. – № 5. – 2015. – С. 45–53.

УДК 004.822

К.В. Миронов, С.В. Тришин, А.Р. Махмутов, В.М. Картак

О ПРИМЕНЕНИИ СИСТЕМ РАСПРЕДЕЛЕННОГО РЕЕСТРА В ЭНЕРГЕТИКЕ

Аннотация

Рассматривается задача обеспечения информационной безопасности в информационной системе, работающей по технологии Интернета вещей. Такая система включает сенсоры и устройства, являющиеся источниками информации, точки доступа, через которые информация от сенсоров и устройств передается в сеть Интернет, и серверы, на которых хранится полученная информация и которые предоставляют доступ к ней пользователям системы. При хранении информации на сервере и предоставлении доступа к ней необходимо обеспечить ее целостность и доступность для пользователей. Для этого предлагается применить технологию распределенного реестра, которая позволяет обеспечить целостность информации, находящейся в общем доступе. В качестве примера защищаемой информационной системы рассматривается система обработки и хранения данных о производстве и потреблении электроэнергии в децентрализованной электросети. Проведен обзор существующих на настоящий момент проекты, касающиеся применения технологий распределенного реестра в энергетике. Важным препятствием на пути применения систем распределенного реестра в сетях Интернета вещей является противоречие между, с одной стороны, высокими требованиями к памяти и вычислительным ресурсам, которые предъявляет технология распределенного реестра, и, с другой стороны, ограниченным ресурсам устройств в сетях Интернета вещей. Предлагаются дальнейшие направления исследований, связанных с преодолением этого препятствия при применении технологий распределенного реестра в энергетике.

Ключевые слова: системы распределенного реестра, интернет вещей, низкоресурсная криптография, блокчейн, Smart Grid.

Введение

Технологии распределенного реестра, в частности, блокчейн, в последние годы вызывают все больше внимания. Системы распределенного реестра включены в список девяти основных сквозных цифровых технологий программы «Цифровая экономика Российской Федерации» [1]. Начиная с возникновения системы биткоин в 2008–2009 годах большинство проектов в области блокчейна создается энтузиастами и бизнес-структурами. При этом почти не проводится фундаментальных исследований по оценке применимости данной технологии для тех или иных задач, преимуществ и проблем, которые могут возникнуть в ходе реализации и эксплуатации систем распределенного реестра. В связи с этим значительная часть источников, на которые мы ссылаемся – это не научные статьи, а сопроводительная документация инициативных проектов (так называемые white papers).

Данная статья посвящена применению технологии распределенного реестра для обеспечения целостности данных в информационной системе, работающей по технологии Интернета вещей. В качестве примера защищаемой информационной системы рассматривается система обработки и хранения данных о производстве и потреблении электроэнергии в децентрализованной электросети. Целью работы является построение системы защиты информации, поступающей с «умных» счетчиков электроэнергии с использованием технологии распределенного реестра. Как известно существует 3 основных аспекта защиты информации: обеспечение конфиденциальноти, целостности и доступности информации. В рамках нашего проекта большее внимание уделяется двум последним аспектам.

При распределенном производстве электроэнергии по технологии Smart Grid к одной электросети может быть подключено множество потребителей и производителей электроэнергии (при централизованном производстве производитель только один). Учет производства и потребления энергии ведется с помощью «умных» счетчиков. Обработка показаний с множества счетчиков позволяет определить, сколько должны заплатить за электроэнергию конкретные потребители и каким образом эта сумма должна быть распределена между производителями. В США и странах Евросоюза распределенные платформы для взаимодействия производителей и потребителей электроэнергии активно развиваются [2]. В России данное направление менее актуально в силу того, что производством электроэнергии занимается ограниченное число производителей, которые слабо заинтересованы в децентрализованной системе. Нормативная база под подобные решения только формируется, основой для этого служит Федеральный закон № 261-ФЗ «Об энергосбережении и о повышении энергетической эффективности» [3]. Однако, для потребителей в отдаленных районы, в которые поставка электроэнергии связана с неоправданно большими затратами или невозможна, предлагаемая концепция построения распределенных систем генерации энергии актуальна. Примерами таких районов являются отдаленные населенные пункты Сибири, Дальнего

Востока и Севера России. В рамках нашей работы в распределенный реестр предполагается вносить информацию о производстве и потреблении электроэнергии, фиксируемую «умными» счетчиками.

Использование технологии распределенного реестра позволяет осуществлять обмен информацией при отсутствии доверия между производителем и потребителем и без организации дополнительной среды взаимодействия. Во втором разделе статьи анализируются существующие проекты, основанные на применении технологии распределенного реестра. В третьем разделе приводится обоснование имеющихся технических ограничений.

Системы распределенного реестра в энергетике

Исторически первой формой распределенного реестра является блокчейн. Блокчейн как структура для хранения информации характеризуется следующими особенностями:

1. Хранимые данные представляют собой множество записей, каждая из которых подписана электронной подписью автора.

2. Множество записей состоит из подмножеств – блоков, упорядоченных хронологически, причем каждый следующий блок содержит криптографическую хеш-функцию от предыдущего.

3. Синхронизированные копии блокчейна хранятся на множестве узлов сети. Данное условие не является обязательным, если оно не выполняется, то говорят о «централизованном блокчейне». Однако централизованный блокчейн, строго говоря, не является распределенным реестром.

Такая структура позволяет обеспечить целостность информации: чтобы изменить некоторый блок, потребуется изменить все последующие блоки в цепи, поскольку хеш-функция каждого блока зависит от хеш-функции предыдущего блока. Таким образом, для подмены некоторого блока потребуется вступить в сговор со всеми пользователями, которые отправляли свои записи в блокчейн после данного блока, поскольку хеш-функции каждой записи подписывается отправителем. Так как блокчейн является распределенной структурой, благонамеренные узлы откажутся подтверждать внесение изменений. Чтобы осуществить подобные изменения злоумышленнику потребуется получить контроль над более чем половиной всех узлов сети.

Альтернативным способом построения распределенного реестра является концепция DAG (Directed acyclic graph – направленный ацикличный граф). В данной концепции узлами графа являются блоки, а направленными дугами – их хеш-функции блоков. Основное отличие от блокчейна заключается в том, что в блокчейне каждый блок ссылается строго на предыдущий, а в DAG каждый блок может ссылаться на произвольные блоки реестра. При записи нового блока его автор смотрит список тех блоков реестра, хеш-функции которых не содержатся ни в одном другом блоке. Далее выбирает из блоков несколько (обычно 2) и добавляет их хеш-функции в свой блок. По сравнению с блокчейном, в DAG менее жесткие требования к синхронизации сети. Если два узла на разных концах сети одновременно генерируют новый блок, то при использовании блокчейна это приведет к конфликту двух версий блокчейна (верной выбирается та, которую первой примет большее число пользователей). При использовании DAG оба блока будут успешно записаны в реестр. По этой причине концепция DAG представляется предпочтительной для сетей Интернета вещей, в которой пропускная способность каналов связи ограничена, а значит, возможны значительные задержки в передаче данных между узлами. На основе технологии DAG функционирует криптовалюта iota [4], которая позиционируется как криптовалюта для интернета вещей.

Тем не менее, все известные нам проекты построения распределенных реестров в энергетике применяют технологию блокчейн. Далее мы рассмотрим их подробнее.

Авторы работы [5] выделяют 9 основных направлений применения блокчейна в энергетике:

1. Распределенная торговля
2. Зарядные станции для электромобилей
3. Вознаграждение за производство возобновляемой энергии
4. Дифференцированные платежи
5. Оплата электроэнергии криптовалютой
6. Оптовые продажи энергии
7. Стабилизация нагрузки на сеть
8. Торговля экологическими активами
9. Платформы для обмена данными

Для нашего проекта наиболее интересными являются направления 1 и 9. Цель создания распределенных торговых площадок – сформировать децентрализованную систему, при помощи которой небольшие производители, например владельцы солнечных панелей или ветрогенераторов, смогут продавать электроэнергию близлежащим потребителям напрямую, не привлекая крупных поставщиков энергии, а потребители в свою очередь получают возможность выбирать поставщика и покупать энергию по наиболее экономичному тарифу. Как правило платформы создаются в пределах квартала или района, например проекты BrooklynMicrogrid и AllgauMicrogrid компании LO3 Energy [6] использует распределённую торговую площадку для жителей небольших общин. Многие компании создают собственные «умные» счетчики которые фиксируют потребление энергии и связываются с блокчейном для осуществления транзакций. Так, например счетчик, разрабатываемый компанией Grid+ [7], сочетает в себе технологии блокчейна и машинного обучения, он способен эффективно планировать потребление электроэнергии, покупать и продавать по её наиболее выгодной цене. Другие подобные про-

екты: PowerLedger [8], Greeneum [9], Wepower [10], Electrify.Asia [11].

Проекты по созданию платформ для обмена данными об использовании энергии используют блокчейн для создания механизмов, при помощи которых производители, регуляторы, коммунальные службы и потребители смогут взаимодействовать друг с другом и передавать информацию об электроэнергии. Подобные платформы создаются в рамках проектов Energy Web Foundation [12] платформа GridSingularity, Clearwatts [2], ElectriCChain [13]. Еще в 2016 году компания Qiwi в сотрудничестве с «Таврида электрик» планировала использовать блокчейн для хранения данных о контрактах на поставку энергии [14], однако на данный момент никаких новых сведений о развитии проекта нет. В таблице 1 представлены основные сведения о наиболее известных проектах. Как можно заметить Ethereum является наиболее популярной платформой, более 40 % [2] проектов используют его в качестве платформы. Благодаря механизму смарт контрактов и виртуальной машине Ethereum (Ehtereum Virtual Machine) становится возможным создавать различные распределенные приложения и использовать блокчейн в разных областях. Однако публичный тип блокчейна и майнинг построенный на «доказательстве работы» ограничивают возможности его применения.

Tobalaba [12] – это блокчейн-платформа созданная Energy Web Foundation. Она представляет собой закрытый блокчейн с разрешениями разного уровня (Private Permissioned Blockchain) основанный на Ethereum и предназначенный для использования компаниями в энергетическом секторе. HyperledgerFabric [16] реализует технологию распределенного реестра для создания платформы закрытых блокчейнов управляемых организациями объединения LinuxFoundation. Платформа использует смарт контракты, называемые Chaincode. Tendermint [17] – это протокол состоящий из двух основных компонентов: механизма достижения консенсуса и универсального интерфейса приложений. Право на создание нового блока присваивается псевдослучайным образом, а добавление блока происходит путем многоэтапного голосования. BigchainDB [18] – это платформа, которая объединяет свойства блокчейна (децентрализацию, неизменяемость, контролируемые активы) и возможности баз данных (высокая скорость обработки транзакций, низкие задержки, возможность индексирования и запроса структурированных данных). BigchainDB был представлен в 2016, а в 2018 обновлен до версии 2.0. Новая версия использует протокол Tendermint, решающий задачу византийских генералов и ускоряющий работу сети.

Применимость технологии в реальных сетях

Основные ограничения по использованию устройств Интернета вещей в качестве узлов распределенного реестра связаны с затратами памяти на хранение реестра и затратами вычислительных ресурсов на расчет хеш-функций и постановку электронных подписей. В существующих работах и проектах по применению распределенного реестра в

Таблица 1

Основные сведения о проектах

Название проекта	Сфера деятельности	Блокчейн платформа	Алгоритм хеширования и электронной подписи
LO3 Energy (Brooklyn Microgrid) [6]	Распределенная торговая площадка	Ethereum	Keccak-256 ECDSA
Grid+ [7]	Распределенная торговая площадка	Ethereum	Keccak-256 ECDSA
PowerLedger [8]	Распределенная торговая площадка	Ethereum	Keccak-256 ECDSA
Greeneum [9]	Распределенная торговая площадка	Ethereum/ Hyperledger/ Tendermint	Keccak-256 / SHA-256 ECDSA
Wepower [10]	Распределенная торговая площадка	Ethereum	Keccak-256 ECDSA
Electrify.Asia [11]	Распределенная торговая площадка	Ethereum	Keccak-256 ECDSA
Energy Web Foundation [12]	Платформы для обмена данными об использовании энергии	Tobalaba	Keccak-256 ECDSA
Clearwatts [2]	Платформы для обмена данными об использовании энергии	BigchainDB	Keccak-256 ECDSA
Enerchain [15]	Оптовые продажи энергии	Tendermint	SHA-256

энергетике вопросы этих ограничений как правило не рассматриваются. Проекты либо не перешли на работу с умными счетчиками на данном этапе развития (например в рамках дошедшего до практической реализации проекта LO3 Energy [6] данные о показаниях счетчиков вводятся потребителями вручную, что сводит на нет основное преимущество блокчейна – необязательность доверия между участниками системы; большинство других проектов и вовсе не дошло до практического применения), либо требуют применения более дорогих умных счетчиков с высокими вычислительными ресурсами.

В зависимости от вычислительных возможностей конечных устройств («умных» счетчиков) можно предложить три варианта их взаимодействия с распределенным реестром:

1. Счетчики не обладают ни памятью, ни вычислительной мощностью для работы с реестром, поэтому в сети назначены подтверждающие узлы, обладающие достаточными. В список задач узлов также входит хранение полной копии реестра, осуществление проверки транзакций, формирование и подтверждение блоков. Предполагается, что счетчики и узлы полностью доверяют друг другу и обмениваются информацией по каналам связи, защищенным средствами низкоресурсной криптографии.

2. Счетчики вычисляют хеш-функции и ставят свои электронные подписи на записи в реестре, но копию всего реестра не хранят.

3. Счетчики выступают полноценными узлами реестра: хранят в своей памяти копию реестра и занимаются вычислением хеш-функций и электронных подписей.

Третий вариант не представляется оправданным: встроенной памяти контроллеров недостаточно для хранения всех записей реестра, если он будет использоваться большой системой в течение длительного периода времени: например, размер открытого блокчейна Ethereum превысил 176 ГБ [19]. Вариантом, позволяющим реализовать данную альтернативу, является использование в составе «умного» счетчика микрокомпьютера по типу Raspberry Pi с возможностью использования карт памяти. Однако, такой вариант повысит стоимость конечного устройства в несколько раз; при этом значительно увеличатся физические размеры счетчиков и их собственное энергопотребление.

Второй вариант содержит существенную сложность, которая заключается в высоких требованиях по вычислительным мощностям при самостоятельном вычислении умным счетчиком электронных подписей. Считать электронные подписи на самих счетчиках можно, однако это существенно повысит их стоимость, так как потребуется замена аппаратного обеспечения более мощным.

Постановка электронной подписи состоит из двух операций – расчет хеш-функции от подписываемого файла и расчет электронной подписи на основе этой хеш-функции. Обе операции являются достаточно ресурсоемкими: аппаратная реализация хеш-функции SHA-256 [20, 21] занимает около 10 000 логических элементов, как и аппаратная реализация эллиптической криптографии с длиной ключа, эквивалентной 113-битному ключу при симметричном шифровании (при этом вычисление подписи в такой компактной реализации будет занимать значительное время [22]). Функция Keccak является быстрой на 64-разрядных системах (12,5 циклов на байт на системах с CPU Intel Core 2 Duo) [23], однако вычислительные возможности контроллеров в «умных» счетчиках гораздо ниже. Функция Scrypt основана на SHA-256, однако значительно требовательнее по ресурсам и оперативной памяти в том числе [24] для усложнения подбора значения хэш-функции перебором, в связи с чем, использование самой SHA-256 является наиболее оправданным. При использовании системы в России в соответствии с законодательными нормами необходимо построение системы на основе хеш-функции Стрибог и электронной подписи ГОСТ Р 34.10-2012.

Выводы

Технологии распределенного реестра могут применяться в распределенной энергетике для обеспечении целостности и доступности информации о показаниях «умных» счетчиков потребления энергии. Проведенный обзор источников по применению технологий распределенного реестра в энергетике показал, что хотя технология DAG является более удобной для построения реестра, существующие проекты используют технологию блокчейн. При этом как правило используется блокчейн системы Эфириум, или основанный на нем. Большинство проектов не дошли до стадии широкого практического применения, а те, что дошли, часто предъявляют завышенные требования к мощности и стоимости оборудования. В связи с этим разработка технологии, не предъявляющей высокие требования к вычислительной мощности и памяти счетчиков представляется актуальной. Дальнейшая разработка системы может вестись по следующим направлениям:

1. Разработка технологии защищенной передачи данных от счетчика к доверенному узлу, который записывает показания в реестр. Защита данных основана на использовании низкоресурсной криптографии.

2. Разработка модели счетчика, ресурсов которого достаточно для постановки электронной подписи на результаты измерений.

3. Разработка модели распределенного реестра для хранения данных о показаниях счетчиков на основе технологии DAG.

Благодарности

Работа выполнена при поддержке РФФИ, грант 19-07-00972.

Список литературы

1. Program “Digital Economy of the Russian Federation” of 28.07.2017 #1632-r.

2. Montemayor, L., Boersma, T. and van Dorp, T. (2018). Comprehensive Guide to Companies involved in Blockchain & Energy. Solarplaza. Date Views 29.01.2019 https://ipci.io/wp-content/uploads/2017/12/Energy-Blockchain-Report.compressed.pdf.

3. Federal Law “On energy saving and energy efficiency improvements and on amendments to certain Legislative Acts of Russian Federation” of 23.11.2009 #261-FL.

4. IOTA (MIOTA) – Whitepaper Date Views 29.01.2019 https://ipci.io/wp-content/uploads/2017/12/Energy-Blockchain-Report.compressed.pdf https://whitepaperdatabase.com/iota-miota-whitepaper/.

5. Edeland C., Mörk T., 2018. Blockchain Technology in the Energy Transition: An Exploratory Study on How Electric Utilities Can Approach Blockchain Technology. Date Views 29.01.2019 https://kth.diva-portal.org/smash/get/diva2:1235832/FULLTEXT01.pdf.

6. Exergy electric power technical whitepaper. Date Views 29.01.2019 https://exergy.energy/wp-content/uploads/2017/12/Exergy-Whitepaper-v8.pdf.

7. Grid+ White paper. Date Views 29.01.2019 https://gridplus.io/assets/Gridwhitepaper.pdf.

8. PowerLedger Whitepaper.Date Views 29.01.2019 https://cdn2.hubspot.net/hubfs/4519667/Documents%20/Power%20Ledger%20Whitepaper.pdf.

9. Greeneum Whitepaper. Date Views 29.01.2019 https://www.greeneum.net/whitepaper.

10. Wepower White Paper. Date Views 29.01.2019 https://icorating.com/upload/whitepaper/CDt62AS6IjlVUY8CNdHk0aYJycsT7ez8Yp5HgKod.pdf.

11. Electrify ICO Whitepaper. Date Views 29.01.2019 https://electrify.asia/whitepaper/

12. The Energy Web Chain: AcceleratingtheEnergyTransitionwithanOpen-Source, Decentralized Blockchain Platform. Date Views 30.01.2019 https://energyweb.org/wp-content/uploads/2018/10/EWF-Paper-TheEnergyWebChain-v1-201810-FINAL.pdf.

13. SolarCoin A blockchain-based solar energy incentive. Date Views 30.01.2019 https://solarcoin.org/wp-content/uploads/SolarCoin_Policy_Paper_EN-1.pdf

14. Qiwi is going to implement blockchain into the energy sector of the Crimea.Date Views 31.01.2019 https://forklog.com/qiwi-eyes-annexed-crimea/

15. Potential of the Blockchain Technology in Energy Trading. Date Views 30.01.2019 https://ponton.de/downloads/mm/Potential-of-the-Blockchain-Technology-in-Energy-Trading_Merz_2016.en.pdf.

16. Androulaki E. et al., 2018. Hyperledger fabric: a distributed operating system for permissioned blockchains. Proceedings of the Thirteenth EuroSys Conference. – ACM, 2018. – С. 30.

17. Tendermint Documentation Date Views 29.01.2019https://media.readthedocs.org/pdf/tendermint/v0.21.0/tendermint.pdf.

18. BigchainDB 2.0 The Blockchain Database May 2018 Paper version 1.0 Date Views 29.01.2019 https://www.bigchaindb.com/whitepaper/bigchaindb-whitepaper.pdf.

19. Ethereum (ETH) prices and statistics Date Views 25.01.2019 https://bitinfocharts.com/ru/ethereum/

20. M. Feldhofer and C. Rechberger. A Case Against Currently Used Hash Functions in RFID Protocols. In First International Workshop on Information Security—IS 2006, volume 4277 of Lecture Notes in Computer Science, pages 372–381. Springer-Verlag, 2006.

21. Poschmann A. Lightweight Cryptography - Cryptographic Engineering for a Pervasive World, 2009. Ph.D. Thesis, RuhrUniversityBochum.

22. Zhukov A.Y. Lightweight cryptography, 2015. Part 2 Voprosy kiberbezopasnosti, 2(10).

23. Manifavas C., Hatzivasilis G., Fysarakis K., Rantos K. Lightweight Cryptography for Embedded Systems – A Comparative Analysis, SETOP’2013.

24. GuidoBertoni, JoanDaemen, MichaëlPeeters, GillesVanAssche. TheKeccakreference, Version 3.0, January 14, 2011. – С. 1–69.

УДК 551.34 551.58

И.Н. Муратов, М.С. Терентьева

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ СТАТИСТИЧЕСКИХ СВОЙСТВ ПОЛЕЙ ТЕРМОКАРСТОВЫХ ОЗЕР В КРИОЛИТОЗОНЕ ЗАПАДНОЙ СИБИРИ

Аннотация

Проведено исследование статистических свойств полей термокарстовых озер в зоне мерзлоты Западной Сибири на основе космических снимков. Исследования проводились в границах сплошной, прерывистой и островной криолитозон методом дистанционного зондирования Земли использованием космических снимков высокого разрешения Канопус-В, отобранных за сравнительно короткий промежуток времени (2013–2015 гг.). В исследовании использован метод ключевых (тестовых) участков. На территории исследования было заложено 78 тестовых участков, на которых были определены поля озер, рассчитаны их основные статистические характеристики. Анализ данных показал, что при увеличении площадей распространения многолетней мерзлоты происходит уменьшение значений характеристик полей озер. Выдвинуто предположение, что данная зависимость может быть вызвана значениями среднегодовой температуры воздуха и амплитудой колебания температуры почвы, которые имеют более высокие значения в зоне распространения островной мерзлоты. Проведено исследование формы береговых границ озер. Для оценки формы береговых границ использовался коэффициент извилистости. Показано, что при увеличении площади мерзлоты уменьшается извилистость озер. Исследование извилистости береговых границ озер также подтверждает, что термокарстовые процессы быстрее протекают в зоне островной мерзлоты.

Ключевые слова: дистанционные методы, спутниковые снимки, геоинформационные системы, вечная мерзлота, Западная Сибирь, термокарстовые озера.

Введение

В последние десятилетия пристальное внимание исследователей занимает проблема парникового эффекта, рассматриваемого в качестве одного из факторов повышения среднегодовой температуры на планете – глобального потепления. Важность этих исследований подтверждается решениями Парижского климатического саммита (2015) о разработке мероприятий, позволяющих не допустить повышение среднегодовой температуры на Земле более 2 °С. Парниковый эффект вызывается ростом объёмов поступления в атмосферу парниковых газов, главными из которых являются углекислый газ и метан. Метан является одним из наиболее важных парниковых газов благодаря высокому потенциалу глобального потепления, по величине которого он, по данным [1], превышает углекислый газ. В работах [2–5] показано, что на территориях севера наибольшее количество метана выделяется в атмосферу из болот и озерно-термокарстовых равнин. Доказано [6], что ¼ часть мировых запасов метана законсервирована на территории Западной Сибири, основную часть которой занимают болотно-озерные ландшафты.

По оценкам [7, 8], наиболее интенсивное выделение метана в арктических и субарктических районах происходит из малых озер термокарстового происхождения. Вследствие труднодоступности территорий в зоне мерзлоты эти исследования малых термокарстовых озер проводятся дистанционными методами с применением спутниковых снимков. К территориям, на которых проведены исследования озер термокарстового происхождения с использованием дистанционных методов, относятся: Сибирь, Скандинавия, Аляска и другие регионы Арктической зоны [9–12]. В основном исследования проводились по снимкам Landsat среднего пространственного разрешения, на которых выделяются только крупные озера. Данные снимки позволили изучить флуктуации численности и размеров озер в интервале нескольких десятилетий [13–15] и распределения озер по размерам, необходимые для оценки запасов метана в озерах [16–19].

Однако на снимках Landsat не видны малые озера (с размерами менее 0,05–0,1 га), концентрация метана в воде которых, по данным [7, 8, 19, 20], более чем на порядок превышает его концентрацию в больших озерах. Это приводит к необходимости использовать снимки высокого пространственного разрешения [21, 22] для изучения полей малых термокарстовых озер.

Для оценки вклада миллионов малых озер в общий объем запасов метана на обширных территориях Северной Евразии необходимо при моделировании динамики термокарстовых озерных полей учитывать особенности распределения полей термокарстовых озер в разных зонах мерзлоты с использованием наземных и дистанционных методов исследования (по космическим снимкам высокого разрешения). Из-за труднодоступности территорий Севера Западной Сибири наземные исследования очень трудоемки и дорогостоящи, в то время как дистанционные исследования менее затратные. Насколько нам известно, дистанционные исследования особенностей распределения полей термокарстовых озер в разных зонах мерзлоты по снимкам высокого разрешения ранее не проводились, что и определило цель данной работы.

Методика

Исследования проводились в границах сплошной, прерывистой и островной криолитозон Западной Сибири методом дистанционного зондирования Земли (ДЗЗ) с использованием космоснимков высокого разрешения Канопус-В (2,1 м), отобранных за сравнительно короткий промежуток времени (2013–2015 гг.). Выбирались снимки, осуществленные в летний сезон (июнь – август) для уменьшения роли сезонных изменений уровня воды в водоемах. Также, в летний сезон отсутствует лед на озерах, препятствующий их дешифрированию на космических снимках.

Исследования проведены на совокупности тестовых участков, карта расположения которых на территории разных зон многолетней мерзлоты Западной Сибири приведена на рисунке 1. Выбор тестовых участков осуществлялся на основе общепринятого в подобных исследованиях подхода, в местах концентрации термокарстовых озер, т.е. в местах, где возникают условия для развития и формирования очагов термокарста [22]. На каждом тестовом участке (ТУ) по космоснимкам дешифрировалось от нескольких сотен до нескольких тысяч термокарстовых озер. Общая характеристика тестовых участков и их распределение по зонам мерзлоты даны в таблице 1.

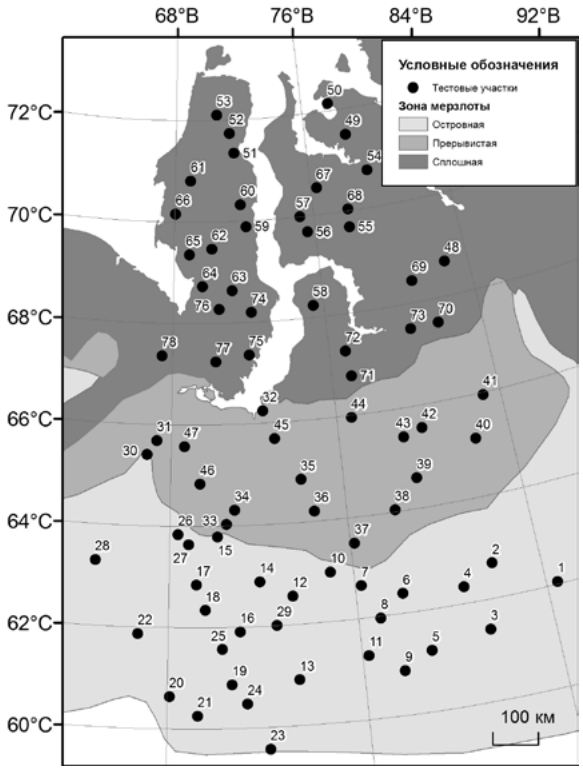


Рис. 1. Схема расположения тестовых участков

Таблица 1

Характеристика тестовых участков

Криолитозона	Количество ТУ	Номера ТУ	Средняя площадь ТУ
Островная	29	1-29	2845
Прерывистая	18	30-47	3125
Сплошная	31	47-78	3109

Космические снимки высокого разрешения Канопус-В дешифрировались с использованием метода бинарной классификации. Обрабатываемые панхроматические снимки имеют сильный контраст между водной поверхностью и окружающей растительностью, что позволяет визуально определить пороговые значения для классификации воды. Метод основывается на визуальном определении порогового значения, с помощью которого производится автоматическое выделение воды. Пороговые значения определяются отдельно для каждого тестового участка, так как эти значения варьируются из-за различий в параметрах съёмки. Обработка, дешифрирование и векторизация данных ДЗЗ проводилась с использованием лицензионного программного обеспечения ArcGIS 10.3.

Результаты экспериментальных исследований

Для изучения закономерностей изменения свойств полей малых термокарстовых озер в зависимости от типа криолитозоны Западной Сибири была построена таблице 2. В таблице использованы средние значения исследуемых характеристик, рассчитанных как среднеарифметические величины по всем ТУ, находящимся в каждой криолитозоне.

Указанные характеристики при переходе из одной зоны мерзлоты в другую проявляют заметные изменения, причем они демонстрируют сходный характер зависимости, а именно: уменьшение значений характеристик при увеличении площадей распространения многолетней мерзлоты, что проиллюстрировано на рис 2.

Данная зависимость может быть вызвана действием климатического фактора. В зоне островной мерзло-

Таблица 2

Характеристики полей термокарстовых озер в разных криолитозонах

Зона мерзлоты	Число озер	Общая площадь озер, га	Заозеренность, %	Извилистость
Островная	950	701	25	1,95
Прерывистая	1 163	664	21	1,85
Сплошная	451	540	17	1,72

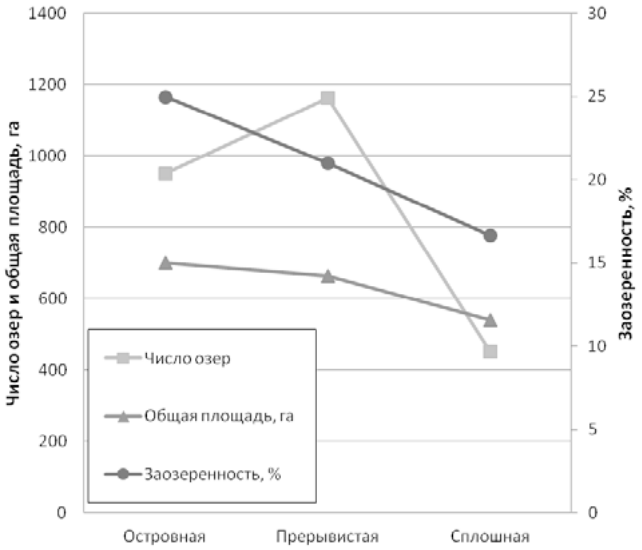


Рис. 2. График изменения характеристик полей озер в зависимости от зон мерзлоты

ты среднегодовая температура воздуха и амплитуды колебания температуры почвы выше, чем в других зонах. Можно предположить, что в данной зоне быстрее протекают термокарстовые процессы, что сопровождается возникновением большого количества новых озер и увеличением площадей уже существующих.

Важным показателем динамики развития термокарстовых процессов является форма береговых границ озер. Для оценки формы береговых границ используется коэффициент извилистости, который определяется по формуле:

$$f = \frac{p^2}{4\pi S}, \quad (1)$$

где p – периметр озера, S – площадь озера.

Коэффициент извилистости показывает, насколько сильно форма озера отличается от окружности, принятой как базовая форма для большинства озер. Значение коэффициента извилистости для окружности равняется единице. Соответственно, чем больше рассчитанное значение извилистости озера отличается от единицы, тем больше озеро отличается от окружности и тем больше извилистость его берегов.

Периметр и площадь озер вычислялись по космическим снимкам путем векторизации полей озер. В результате работы алгоритма бинарной классификации получаются контуры озер со «ступенчатыми» границами, так как классификация проводится по границам пикселей. Использование периметра озер, рассчитанного по таким контурам, приведет к завышению коэффициента извилистости. Для избавления от «ступенчатости» векторные контуры озер подвергались процедуре сглаживания по методу полиномиальной аппроксимации с использованием экспоненциального ядра (PAEK). После сглаживания рассчитывались основные характеристики полей озер и коэффициент извилистости. Как видно из таблицы 2, извилистость термокарстовых озер в разных зонах мерзлоты проявляет ту же зависимость, что и другие характеристики – при увеличении площади мерзлоты уменьшается извилистость озер. Это может быть также связано с активностью термокарстовых процессов – в зоне мерзлоты, где процессы протекают быстрее, озера имеют более извилистые контуры.

Выводы

В статье рассмотрены вопросы проведения дистанционных исследований основных характеристик полей малых термокарстовых озер в сплошной, прерывистой и островной криолитозонах Западной Сибири. Исследования проведены с использованием космических снимков высокого пространственного разрешения Канопус-В, позволивших определить число и измерить площади термокарстовых озер на 78 тестовых участках. С использованием этих данных определены основные характеристики полей термокарстовых озер в разных криолитозонах: число, суммарную площадь и извилистость озер, а также заозеренность территории. На основе проведенного анализа установлено, что при увеличении площадей распространения многолетней мерзлоты происходит уменьшение значений характеристик полей озер. Выдвинуто предположение, что данная зависимость может быть вызвана значениями среднегодовой температуры воздуха и амплитудой колебания температуры почвы, которые имеют более высокие значения в зоне распространения островной мерзлоты. Исследование извилистости береговых границ озер также подтверждает, что термокарстовые процессы быстрее протекают в зоне островной мерзлоты.

Благодарности

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (договоры № 18-47-700001, № 18-45-860002 и № 18-45-703001), Администрации Томской области и Департамента образования и молодежной политики ХМАО – Югры.

Список литературы

1. Кароль, И.Л. Оценки характеристик относительного вклада парниковых газов в глобальное потепление климата [Текст] / И.Л. Кароль // Метеорология и гидрология. – 1996. – № 11. – С. 5–12.

2. Walter, K. M. Methane bubbling from northern lakes: present and future contributions to the global methane budget [Text] / K.M. Walter, L.C. Smith, F.S. Chapin // Phil. Trans. R. Soc. – 2007. – Vol. 365. – P. 1657–1676.

3. Fates of methane from different lake habitats: Connecting whole-lake budget and CH4 emissions [Text] / D. Bastviken, J. J. Cole, M. L. Pace [et al.] // J. of Geoph. Res. – 2008. – Vol. 113. – G02024.

4. Methane dynamics in different boreal lake types [Text] / S. Juutinen, M. Rantakari, P. Kortelainen [et al.] // Biogeosciences. – 2009. – №. 6. – P. 209–223.

5. Seasonal variability as a source of uncertainty in the West Siberian regional CH_4 flux upscaling [Text] / A. F. Sabrekov, B. R. K. Runkle, M. V. Glagolev [et al.] // Environ. Res. Lett. – 2014. – № 9. – P. 1–9.

6. Глобальные изменения климата: «метановая бомба» – наукообразный миф или потенциальный сценарий? [Текст] / А.И. Жилиба, Г.А. Вандышева, К.Г. Грибанов [и др.] // Динамика окружающей среды и глобальные изменения климата. – 2011. – Т. 2. – № 1. – С. 3–16

7. Pokrovsky, O.S. Effect of permafrost thawing on the organic carbon and metal speciation in thermokarst lakes of Western Siberia [Text] / O.S. Pokrovsky, L.S. Shirokova, S. N. Kirpotin // Biogeosciences. – 2011. – Vol. 8. – P. 565–583.

8. Holgerson, M.A. Large contribution to inland water CO2 and CH4 emissions from very small ponds [Text] / M.A. Holgerson, P.A. Raymond // Nature Geoscience. – 2016. – № 9. – P. 222–226.

9. Spatial Extent, Age, and Carbon Stoks in Drained Thaw Lake Basins on the Barrow Peninsula [Text] / K.M. Hinkel, W.R. Eisner, J.G. Bockheim [et al.] // Alaska. Arctic, Antarctic, and Alpine Research. – 2003. – Vol. 35. – № 3. – P. 291–300.

10. Luoto, M. Thermokarst ponds as indicator of the former distribution of palsas in Finnish Lapland [Text] / M. Luoto, M. Seppala // Permafrost and Periglacial Processes. – 2003. – Vol. 14. – P. 19–27.

11. Кравцова, В.И. Распространение термокарстовых озер в России [Текст] / В.И. Кравцова // Вестник Московского университета. – 2009. – № 3. – С. 33–42.

12. Анализ развития морфологической структуры озерно-термокарстовых равнин на основе математической модели [Текст] / А.С. Викторов, В.П. Капралова, Т.В. Орлов [и др.] // Геоморфология. – 2015. – № 3. – С. 3–13.

13. Кравцова, В.И. Изменения размеров термокарстовых озер в различных районах России за последние 30 лет [Текст] / В.И. Кравцова, А.Г. Быстрова // Криосфера Земли. – 2009. – Т. 13, № 2. – С. 16–26.

14. Брыксина, Н.А. Анализ изменения численности термокарстовых озер в зоне мерзлоты Западной Сибири на основе космических снимков [Текст] / Н.А. Брыксина, Ю.М. Полищук // Криосфера Земли. – 2015. – т. 19, № 2. – С. 114–120.

15. Polishchuk, Y.M. Remote analysis of changes in the number and distribution of small thermokarst lakes by sizes in Cryolithozone of Western Siberia [Text] / Y.M. Polishchuk, N.A. Bryksina, V.Y. Polishchuk // Izvestia. Atmospheric and Oceanic Physics. – 2015. – Vol. 51. – № 9. – P. 999–1006.

16. Downing, J.A. The global abundance and size distribution of lakes, ponds, and impoundments [Text] / J.A. Downing, Y.T. Prairie // Limnol. Oceanogr. – 2006. – Vol. 51. – P. 2388–2397.

17. Karlsson, J.M. Temporal behavior of lake size-distribution in a thawing permafrost landscape in Northwestern Siberia [Text] / J.M. Karlsson, S.W. Lyon, G. Destouni // Remote sensing. – 2014. – № 6. – P. 621–636.

18. Polishchuk, V.Y. Modeling of thermokarst lake dynamics in West-Siberian permafrost. Ch. 6 [Text] / V.Y. Polishchuk, Y.M. Polishchuk // Permafrost: Distribution, Composition and Impacts on Infrastructure and Ecosystems / Ed. by O. Pokrovsky. – NY: Nova Science Publishers. – 2014. – P. 205–234.

19. Викторов, А.С. Математическая модель морфологической структуры озерно-термокарстовых равнин в изменяющихся климатических условиях [Текст] / А.С. Викторов, В.П. Капралова, О.Н. Трапезникова // Криосфера Земли. – 2015. – Т. 19, № 2. – С. 26– 4.

20. Методические вопросы оценки запасов метана в малых термокарстовых озерах в зоне мерзлоты Западной Сибири [Текст] / Ю.М. Полищук, В.Ю. Полищук, Н.А. Брыксина [и др.] // Известия Томского политехнического университета. – 2015. – Т. 326, № 2. – С. 127–135.

21. Полищук, Ю.М. Методические вопросы построения обобщенных гистограмм распределения площадей озер в зоне мерзлоты на основе космических снимков среднего и высокого разрешения [Текст] / Ю.М Полищук, А.Н. Богданов, И.Н. Муратов // Современные проблемы дистанционного зондирования Земли из космоса. – 2016. – Т. 13, № 6. – С. 224–232.

22. Полищук, Ю.М. Зоны активного термокарста на территории многолетней мерзлоты и их выявление по космическим снимкам [Текст] / Ю.М. Полищук, А.Н. Богданов // Известия Томского политехнического университета. – 2015. – Т. 326, № 12. – С. 104–114.

УДК 519. 711.3

А.А. Семенов, М.А. Вылков

ИСПОЛЬЗОВАНИЕ ПОДСИСТЕМЫ ПРОСТРАНСТВЕННЫХ ДАННЫХ В АИС «ИННОВАЦИИ ДЛЯ ЮГРЫ»

Аннотация

Современные геоинформационные системы направлены на управление пространственными данными и проведение территориального анализа, в том числе анализа социально-экономического развития региона. В статье на примере реализуемого проекта по внедрению автоматизированной информационной системы поиска, отбора и реализации инновационных проектов «Инновации для Югры» рассмотрены возможности применения геоинформационной системы для повышения качества и доступности предоставления государственной услуги Автономного учреждения Ханты-Мансийского автономного округа – Югры «Технопарк высоких технологий» с помощью информационных технологий, что предусмотрено Концепцией региональной информатизации. В статье рассмотрены поставленные задачи и реализуемые возможности системы геопространственных данных, использующей базу данных автоматизированной информационной системы «Инновации для Югры». Введение в промышленную эксплуатацию системы геопространственных данных, как составной части автоматизированной информационной системы «Инновации для Югры», позволяет анализировать и визуализировать данные о инновационных проектах и ключевых показателях их реализации. Подсистема геопространственных данных автоматизированной информационной системы «Инновации для Югры» представляет особую важность как для резидентов Автономного учреждения Ханты-Мансийского автономного округа – Югры «Технопарк высоких технологий», так и для их действующих и потенциальных партнеров.

Ключевые слова: АИС «Инновации для Югры», ГИС, предоставление государственной услуги, анализ реализации инновационных проектов.

Введение

В настоящее время в Российской Федерации особую актуальность и социальную значимость имеют вопросы, связанные с качеством предоставляемых государством услуг. Организации, предоставляющие государственные услуги, входят в систему, которая служит для взаимодействия населения страны с государственными органами, и эта система должна гарантировать обеспечение максимального удовлетворения потребностей граждан в государственных услугах [1]. При этом, в соответствии с одним из основных направлений реализации Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы [2], предусмотрено повышение качества и оперативности предоставления государственных услуг, в том числе за счет их перевод в электронный вид [3].

Ханты-Мансийский автономный округ – Югра (далее – ХМАО – Югра, автономный округ, округ) является одним из регионов-лидеров в сфере информатизации. В рамках реализации положений Национальной технологической инициативы (далее – НТИ) в ХМАО – Югре успешно реализуются стратегические цели государственной политики в сфере информационных технологий, в том числе повышается эффективность государственного и муниципального управления [4]: в электронный вид переведены более 90 % региональных государственных услуг [5].

В округе внедряются автоматизированные информационные системы (далее – АИС), предназначенные для информационного и научно-аналитического обеспечения принятия управленческих решений и призванные решать задачи, поставленные перед региональными ведомствами и учреждениями. АИС являются неотъемлемой частью не только современного бизнеса [6], но и государственного управления. Данные системы дают возможность получать аналитическую информацию по предприятию или региону и применять сложные системы управления производством [7].

Методика

Стоит отметить, что у многих региональных АИС, в том числе использующих геоинформационные системы (далее – ГИС) имеются проблемы, связанные с их разработкой и внедрением в промышленную эксплуатацию.

ГИС – это система сбора, хранения, анализа, графической визуализации пространственных (географических) данных и связанной с ними информации об объектах [8]. Географические методы и ранее использовались в качестве вспомогательных в различных областях применения, однако ГИС выводят их применение на качественно иной уровень. Современные программные средства позволяют визуализировать информацию, что значительно повышает эффективность проводимого анализа [9]. В настоящее время ГИС – это интегрированные информационные системы, которые включают методы обработки данных существующих АИС, а также обладают определенной спецификой в организации и обработке данных [10]. Именно это определяет ГИС как многоцелевые и многоаспектные [11].

Важный аспект, влияющий на обработку информации – эффективность её использования зависит от

информационных баз [12]. При этом одной из наиболее актуальных проблем является то, что значительная часть ранее накопленной информации хранится как на бумажном носителе, так и в электронном виде, но в различных объемах и в различных подразделениях и организациях, зачастую – в различном подведомственном подчинении. Поэтому при внедрении ГИС нужно обращать внимание не только на анализ данных, но также и на их обработку и хранение.

Также большую проблему представляет тот факт, что информация, хранящаяся в базах данных различных ведомств, устаревает и перестаёт отражать ситуацию на текущий момент времени [13]. Для реализации поставленных перед ГИС задач и для её эффективного функционирования, она должна оперировать актуальными на текущий момент времени данными. Алгоритмы ввода, хранения, обработки и анализа картографической информации должны учитывать взаимное влияние объектов, что в свою очередь также влияет на принятие управленческих решений и прежде всего на их оправданность [14].

Основная часть

Распоряжением Правительства ХМАО – Югры от 15.11.2017 г. № 661-рп для АУ «Технопарк высоких технологий» [15] в соответствии со «Стратегией научно-технологического развития Российской Федерации» [16], «Стратегией социально-экономического развития Ханты-Мансийского автономного округа – Югры до 2030 года» [17] были определены следующие основные направления деятельности:

- реализация НТИ путем создания условий для обеспечения лидерства предприятий и организаций промышленного сектора ХМАО – Югры на новых высокотехнологичных рынках;

- формирование условий для перехода промышленности ХМАО – Югры к передовым цифровым, интеллектуальным производственным технологиям, созданию систем обработки больших объемов данных, машинного обучения и искусственного интеллекта;

- инновационная поддержка промышленного роста, в том числе за счет разработки и внедрения новых информационных технологий в интересах промышленности ХМАО – Югры пр.

Для реализации этих направлений проведены работы по созданию АИС поиска, отбора и реализации инновационных проектов «Инновации для Югры» (далее – АИС «Инновации для Югры») [18]. Создание АИС «Инновации для Югры» было запланировано с целью повышения качества и доступности предоставления государственных и муниципальных услуг с помощью информационных технологий, предусмотренных Концепцией региональной информатизации [19].

АИС «Инновации для Югры» предназначена для автоматизации и организации информационного обеспечения решения таких задач, как:

- организация поиска, включения в реестр и отбора инновационных проектов, удовлетворяющих критериям инновационности, реализуемости, конкурентоспособности и практической значимости для экономики округа;

- сопровождение реализации отобранных инновационных проектов;

- организация взаимодействия между субъектами инновационной деятельности.

Практическая значимость

Подсистема ГИС использует базу данных АИС «Инновации для Югры», поддерживающую возможность хранения и актуализации пространственных данных, содержит средства пространственного анализа и визуализации картографического материала. Введение в эксплуатацию ГИС, как составной части АИС «Инновации для Югры», позволит анализировать и визуализировать данные о проектах и ключевых показателях их реализации.

Основными задачами ГИС АИС «Инновации для Югры» являются:

- организация региональной картографической базы данных, содержащей информацию в сфере инновационной деятельности;

- создание технических условий для накопления пространственных данных, а также их систематического обновления;

- реализация единого информационного и геоинформационного пространства сфере инновационной деятельности региона;

- формирование интерактивного картографического сервиса;

- обеспечение свободного доступа к пространственным данным в сфере инновационной деятельности.

Для реализации указанных задач в модуле ГИС предусмотрены следующие возможности:

- Визуализация: представление / визуализация инновационных проектов и решений на географических картах;

- Актуализация: добавление пространственной информации к инновационным проектам и решениям для визуализации объектов на карте, редактирование местоположения объектов и др.;

- Анализ: географический метод работы позволяет пользователям получать ответ на различные вопросы: «Где находится…?», «Сколько…?», или «Что, если…?».

- Представление: результатом работы ГИС является карта распределения проектов и решений в сфере инновационной деятельности, помогающая решать вопросы развития с помощью инструментов пространственного анализа.

Подсистема ГИС АИС «Инновации для Югры» состоит из следующих основных элементов интерфейса (рис. 1).

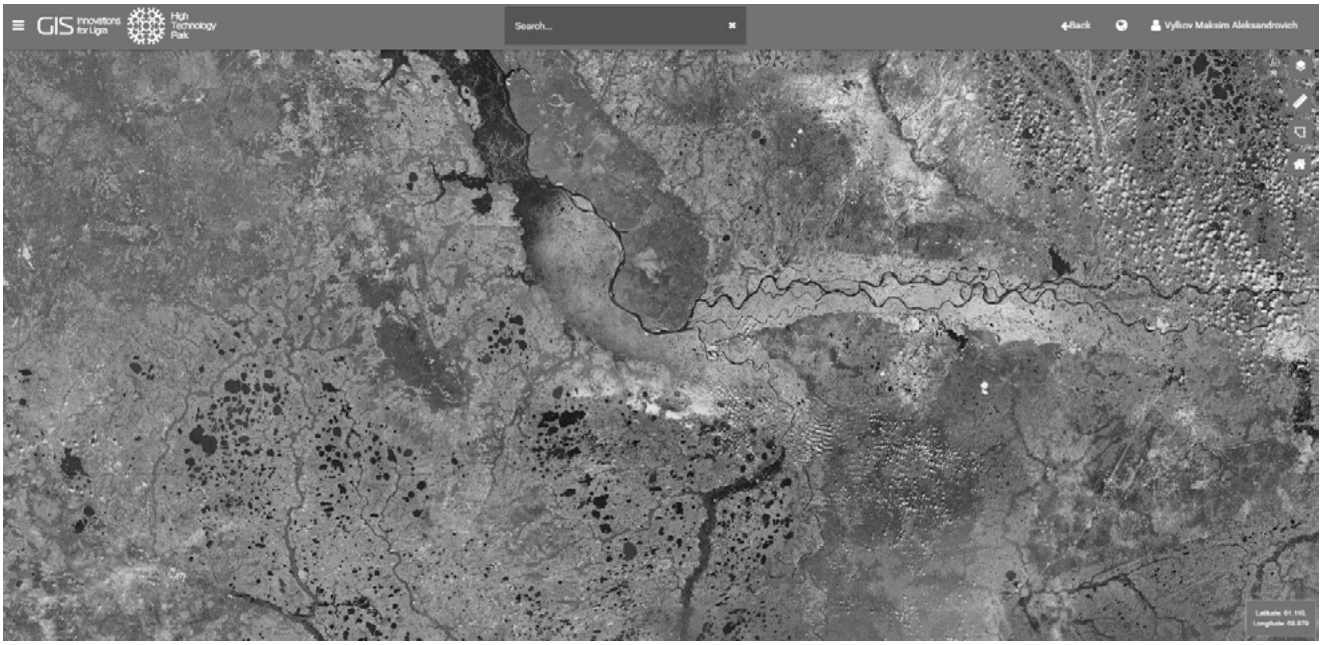


Рис. 1. Основные элементы интерфейса подсистемы ГИС АИС «Инновации для Югры»

Описание основных элементов интерфейса данной подсистемы:

1. Интерактивная карта – рабочая область для визуализации и работы с объектами карты;

2. «Шапка» ГИС – верхняя часть веб-ресурса, расположенная на всех страницах системы, которая содержит: главное меню, поле ввода для поиска объектов по базе данных, кнопку для изменения языковых настроек, кнопку для возвращения на главную страницу АИС и кнопкой для авторизации пользователя;

3. Панель инструментов для работы с картой – содержит инструменты для управления картой и пространственного анализа;

4. Панель для отображения географических координат.

Интерактивная карта ГИС представляет собой рабочую область для отображения электронных карт, на которых представлена картографическая информация различного содержания, привязанная к географическому контексту. Интерактивная карта ГИС позволяет отображать картографическую основу, необходимую для визуализации пространственных объектов в соответствии с географической привязкой к местности.

Картографическая основа представляет собой цифровые тайловые карты, получаемые из открытых тайловых картографических сервисов, на которых должны отображаться основные географические объекты (основная автодорожная и железнодорожная сеть, крупнейшие озёра и реки, границы субъектов Российской Федерации, административные центры и др.).

Основной задачей ГИС является работа с геоданными. Основными типами геоданных являются точки, линии и полигоны. Слои, использующиеся в ГИС, представлены типом объектов – точки. Для реализации отображения объектов на карте в ГИС используется поле в модели данных, содержащее информацию о геометрии объекта и координаты.

Выводы

Введение в промышленную эксплуатацию системы геопространственных данных, как составной части АИС «Инновации для Югры», позволяет анализировать и визуализировать данные о инновационных проектах и ключевых показателях их реализации. Заинтересованное в сотрудничестве лицо посредством ГИС имеет возможность ознакомиться не только с территориальным положением резидента на карте, но также может получить информацию о его деятельности в кратком сформированном виде (рис. 2).

Описание основных элементов интерфейса при работе со слоем «Инновации для Югры»:

1. Кнопка вызова меню выбора слоев – кнопка, вызывающая меню выбора слоев;

2. Счетчик инновационных проектов – отображает количество инновационных проектов на местности;

3. Описательная часть инновационного проекта – содержит полную информацию о реализуемом проекте, в том числе, информацию о результатах, достигнутых в процессе реализации проекта.

При просмотре карты в уменьшенном масштабе ГИС автоматически оптимизирует счетчик инновационных проектов, отображая в рамках небольшой отдельно взятой зоны один графический объект, содержащий в себе число, характеризующее количество инновационных проектов. При увеличении масштаба доступно более подробное расположение инновационных проектов на местности (рис. 2). Данная особенность при совместном использовании с картографической подложкой «Физическая карта» или «OSM» позволяет получить визуальную информацию о реа-

Рис. 2. Основные элементы интерфейса слоя «Инновации для Югры»

лизации инновационных проектов в муниципальных образованиях (районы, города и т.д.). OpenStreetMap (OSM) является одним из самых популярных примеров проекта добровольной географической информации (VGI). В последние годы он стал серьезным альтернативным источником геоданных [20].

Благодарности

Авторы выражают особую благодарность ведущему менеджеру Центра автоматизации и контроля АУ ХМАО – Югры «Технопарк высоких технологий» Ирине Владимировне Дейковой (Irina Vladimirovna Deykova) за неоценимую помощь при написании данной статьи.

Список литературы

1. Chernyakhovskaya L.R., Galiullina A.F., 2017. Development of requirements for a decision support system aimed at quality assessment of public services provided based on the ontological approach. Бизнес-информатика, 1 (39). Дата обращения 06.02.2019 https://elibrary.ru/item.asp?id=29143561.

2. Стратегии развития информационного общества в Российской Федерации на 2017–2030 гг.

3. Исаев Е.А., Коровкина Н.Л., Табакова М.С., 2018. Оценка готовности ИТ-подразделения компании к цифровой трансформации бизнеса. Бизнес-информатика. № 2 (44). Дата обращения 05.02.2018 https://elibrary.ru/item.asp?id=35233563.

4. Информация о достигнутых результатов деятельности Депинформтехнологий Югры за 2018 год. Дата обращения 09.02.2019 https://depit.admhmao.ru/deyatelnost/otchyety/2018/2021197/informatsiya-o-dostignutykh-rezultatakh-deyatelnosti-depinformtekhnologiy-yugry-za-2018-god.

5. Стратегия развития автономного учреждения Ханты-Мансийского автономного округа – Югры «Технопарк высоких технологий» до 2030 года.

6. Van Grembergen, W., De Haes, S., 2005. Measuring and Improving IT Governance Through the Balanced Scorecard // Information systems control journal, volume 2. Дата обращения 06.02.2019 http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.469.7541&rep=rep1&type=pdf.

7. Шутов А.А., Трушин Д.Н., 2009. Методы исследования экономической эффективности от внедрения информационных систем на предприятии. Бизнес-информатика. № 4 (10): 33–39.

8. Magnuson L. GIS and Geospatial Data Tools. M. : ACRL TechConnect, 2015. Дата обращения 02.02.2019 http://acrl.ala.org/techconnect/post/gis-and-geospatial-data-tools.

9. Grishentcev A., Korobeynikov A.G., Velichko E.P., Korikov S., 2015. On Interoperability in Distributed Geoinformational Systems. Internet of Things, Smart Spaces, and Next Generation Networks and Systems. S. Balandin et al. (Eds.): NEW2AN 2015, LNCS 924: 504–512.

10. Yang S., Wang F., 2015. Visualizing information science: Author direct citation analysis in China and around the world. Journal of Informetrics. № 9. Дата обращения 05.02.2019 https://www.sciencedirect.com/science/article/abs/pii/S1751157715000024?via%3Dihub.

11. Barron, C., Neis, P. and Zipf, A., 2014, A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. Transactions in GIS, 18. Дата обращения 09.02.2019 https://doi.org/10.1111/tgis.12073.

12. Цветков В.Я., 2016. Информационные модели и геоинформационные модели. Образовательные ресурсы и технологии. № 3 (15):114–120.

13. Васильев В.Н., 2016. Обзор существующих ГИС. Молодой учёный. № 14 (118). Дата обращения 06.02.2019 https://elibrary.ru/item.asp?id=26389027.

14. Грибкова И.С., Попова О.С., 2016. Муниципальные геоинформационные системы: проблемы и пути их решения. Научные труды Кубанского государственного технологического университета. № 11. Дата обращения 06.02.2019 https://elibrary.ru/item.asp?id=27677209.

15. Распоряжение Правительства ХМАО – Югры от 15.11.2017 № 661-рп «О внесении изменений в распоряжение Правительства Ханты-Мансийского автономного округа – Югры от 2 ноября 2017 года № 637-рп «О внесении изменений в некоторые распоряжения Правительства Ханты-Мансийского автономного округа – Югры».

16. Указ Президента РФ от 01.11.2016 № 642 «О Стратегии научно-технологического развития».

17. Распоряжение Правительства ХМАО – Югры от 22.03.2013 № 101-рп (ред. От 09.06.2017) «О Стратегии социально-экономического развития Ханты-Мансийского автономного округа – Югры до 2030 года».

18. Стратегия развития автономного учреждения Ханты-Мансийского автономного округа – Югры «Технопарк высоких технологий» до 2030 года.

19. Распоряжение Правительства РФ от 29.12.2014 № 2769-р (ред. от 18.10.2018) «Об утверждении Концепции региональной информатизации».

20. Ozhereleva T.A., 2016. Information Barriers. European Journal of Technology and Design. Vol. 11. Is. 1. Дата обращения 09.02.2019 http://oaji.net/articles/2016/16-1457035336.pdf.

УДК 681.518 (075.32)

С.П. Семенов, А.О. Ташкин

ОЦЕНКА ЭФФЕКТИВНОСТИ ГИС ДЛЯ МАЛОМОБИЛЬНЫХ ГРУПП НАСЕЛЕНИЯ

Аннотация

В статье описывается методика оценки эффективности геоинформационного ресурса, ориентированного на людей с ограниченными возможностями с помощью технологий имитационного моделирования. Представлена агентно-ориентированная имитационная модель ГИС для МГН geowheel.ru, определены элементы системы, установлены параметры входных и выходных переменных, введены ограничения, созданы структурно-функциональные и логические схемы. В качестве результатов исследования проведены экспериментальные исследования, предложены количественные оценки эффективности ГИС для МГН.

Ключевые слова: информация, информационные системы, эффективность, имитационный, модели, ГИС, МГН, надежность, ИС, отказ, пропускная способность.

Процесс создания любой информационной системы (ИС) неизбежно связан с проблемой оценки эффективности её работы, что и является целью настоящего исследования. Эффективность работы системы можно определить с помощью следующих выходных параметров:

- надежность (процент отказов системы);
- среднее время нахождения в системе;
- средняя пропускная способность системы;
- пороговые значения входных параметров, при которых наступает тотальный отказ системы.

Известно несколько методов оценки эффективности информационных систем, например, с помощью методов системного анализа, используя инструменты нагрузочного тестирования. В ходе применения методики нагрузочного тестирования на реальной информационной системе с помощью программных средств воспроизводятся действия пользователей [1], в ходе которых эмулируется подача заявок (запросов) к действующей информационной системе. В случае применения данной методики можно выявить пороговые значения входных параметров, при которых наступает тотальный отказ системы. В данной статье, ввиду принадлежности ИС к классу сложных систем, оценка эффективности ИС производится с помощью методов имитационного моделирования, что позволяет учитывать одновременно все уровни архитектуры для достижения большей точности и объективности результатов. Имитационная модель – это формальное описание логической структуры и динамики взаимодействия отдельных элементов реального объекта с учётом стохастических факторов, реализованное как программа для компьютера [2, 3]. С помощью технологий имитационного моделирования можно представить работу существующей системы в виде имитационной модели. Такой подход позволяет выявить не только пороговые значения входных параметров, но и произвести многоплановое комплексное тестирование работы ИС, позволяющее изменять характеристики ИС, алгоритм функционирования и иные параметры работы ИС.

Концептуальная модель ГИС для МГН geowheel.ru, представлена в статье [3]. В настоящем исследовании представлено продолжение работы по развитию гибридного подхода к построению имитационных моделей, включающего в себя характерные черты дискретно-событийного и агентно-ориентированного подходов [4, 5]. Смысл объединения двух подходов заключается в том, что с помощью дискретно-событийного подхода возможно построить логическую структурно-функциональную схему модели, а с помощью агентно-ориентированного подхода отобразить динамику взаимодействия заявки и элементов системы. Для построения и исследования компьютерных моделей различных процессов существует большое количество программных пакетов (AnyLogic, Arena, MATLAB Simulink, Simula и т.д.). Для реализации гибридной модели было выбрано ПО AnyLogic, которое позволяет создавать агентно-ориентированные модели, отражающие общее представление о поведении моделируемой системы посредством назначения перечня параметров и правил поведения ее элементов (агентов), не имея сведений о каких-либо общих законах поведения системы в целом.

Процесс создания имитационной модели начнем с описания реальной ГИС. Реальную действующую систему можно представить в виде виртуальной компьютерной модели, а объекты и функционал реальной системы представить в виде математических обозначений и функций. Геоинформационная система для удовлетворения информационных потребностей маломобильных групп населения geowheel.ru реализована в виде клиент-серверной системы. Серверная часть представляет собой набор сервисов для обработки заявки пользователя, включая базы данных и приложения. Клиентская часть представлена в виде интернет-сайта и наделена интерфейсом взаимодействия с серверной частью. Пользователь при обращении к серверу исходя из возможностей системы имеет возможность выбора характеристик заявки, что определяет её сложность. Таким образом, исследуемая система выполняет функции в соответствии с запросом пользователя. На рисунке 1 отражена логиче-

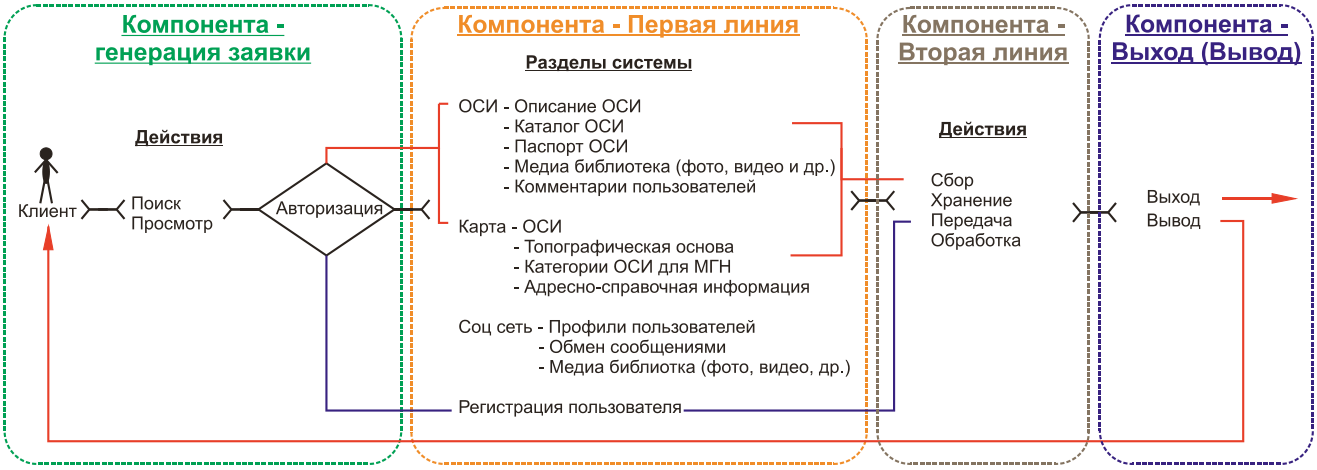


Рис. 1. Логическая схема модели реальной ГИС для МГН geohweel.ru

ская схема, описывающая компоненты ГИС для МГН geohweel.ru и их функциональный состав, выбор которого соответствует запросу пользователя [6, 7, 8].

Для построения модели введем количественные показатели. ГИС для МГН можно описать в виде терминов систем массового обслуживания [4, 5]. Рассматриваемая информационная система оснащена одним сервером, на который с определенной интенсивностью λ на обработку от пользователей поступают заявки с некоторой сложностью S, являющейся случайной величиной. Интенсивность поступления заявок λ определяется количеством обращений пользователей в единицу времени. Сложность заявки определяется количеством операций S, требуемых на ее выполнение. Если сервер занят, то заявка перенаправляется в очередь, который выполняет роль буфера заявок, вместимостью N. Очередь имеет определенный период ожидания, длительностью t_{out}. Производительность обслуживания заявок системой Q (операций/сек) считается известной величиной. Время обработки заявки системой может быть выражено как

$$t_{serv_delay} = \frac{S}{Q}. \quad (1)$$

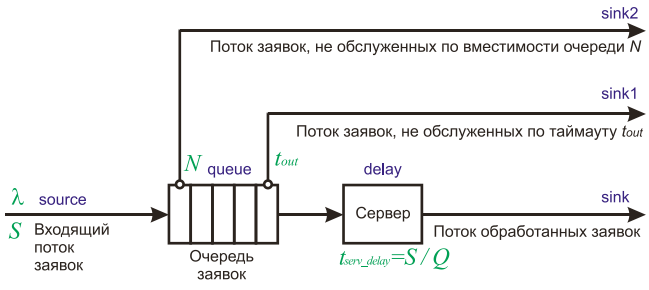


Рис. 2. Схема имитационной модели ГИС geohweel.ru в виде СМО

Наиболее значимыми входными параметрами являются:

λ – интенсивность поступления заявок (количество обращений пользователей в единицу времени).

S – сложность заявки (количество операций, требуемых на выполнение заявки).

Q – производительность системы (количество операций в секунду).

N – максимальная вместимость очереди заявок.

t_{out} – интервал критического времени ожидания ответа от сервера в очереди.

Последовательность обращений пользователей, происходящих одно за другим в случайные моменты времени принято называть входным потоком. Предполагается, что входной поток в рассматриваемой модели является пуассоновским, т.е. стационарным и без последействий [4, 5]. Соответственно, число обращений n за любой фиксированный интервал времени t, будет распределено по закону Пуассона:

$$\mathbb{P}(n,t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad (2)$$

где P(n,t) – вероятность поступления n заявок за некоторый интервал времени t.

Интервалы времени τ между обращениями пользователей распределены экспоненциально:

$$P(\tau < t) = 1 - e^{-\lambda t}, \quad (3)$$

где $P(\tau < t)$ – вероятность того, интервал времени между обращениями будет меньше .

Для работы с системой geowheel.ru пользователь использует интернет-браузер, который с помощью протокола TCP/IP устанавливает соединение с web-сервером. Если соединение неактивно в течение интервала времени t_{out}, браузер повторно устанавливает его. При поступлении на сервер запросов на открытие соединения, сервер рассылает на каждую заявку подтверждение и ожидает дальнейших действий, если же буфер очереди на входящие соединения заполнен, сервер становится недоступным и отклоняет заявки. Если время ожидания ответа от сервера превысило установленный интервал t_{out}, браузер прекращает попытки установить соединение и сообщает пользователю о недоступности сервера.

Характеристики серверной части включают в себя такие параметры, как производительность системы Q

и максимальная вместимость очереди заявок N. Производительность системы Q представлена в виде количественного показателя, консолидирующего характеристики системы (производительность процессора сервера, объем памяти, скорость чтения и записи, пропускная способность сети, и др.) и выраженного в виде количества выполняемых операций в секунду. Для ГИС geowheel.ru были проведены тесты производительности с помощью серверной утилиты sysbench. Под производительностью Q понимается итоговая совокупность параметров системы: производительность процессора, скорость выполнения системных, сетевых, файловых операций и прочие показатели. Вместимость очереди заявок N задается при настройке сервера баз данных.

В текущем исследовании сложность заявки определяется как случайная величина, которая подчиняется симметричному треугольному распределению, и записывается в виде: $[S_{moda} \pm Z]$, где S_{moda} – мода треугольного распределения, а $S_{min} = S_{moda} - Z$ и, соответственно, $S_{min} = S_{moda} + Z$.

Опираясь на формализацию логической схемы ГИС для МГН geowheel.ru, а также используя значения количественного показателя моды сложности S_{moda} была построена компьютерная имитационная модель, реализованная в среде имитационного моделирования Anylogic. В ходе исследования была протестирована работа модели, устойчивость, адекватность, выявлена чувствительность, произведена калибровка, проведен ряд экспериментов. В таблице представлены элементы модели, а на рисунке 3 представлен снимок экрана созданной модели в ПО Anylogic.

После калибровки модели была проведена серия экспериментов, с целью оценки эффективности ГИС для МГН, которые условно можно разделить на 3 группы. Для проведения экспериментов были установлены следующие параметры:

- Единица модельного времени – 1 мин.
- Время длительности эксперимента – 7 суток (10 080 мин.)
- Q – производительность системы, операций/сек. (реальное значение итоговой производительности системы, установленное путем экспертного анализа).
- N – максимальная вместимость очереди заявок, $N=10$.
- t_{out} – критическое время ожидания ответа от сервера в очереди, мин.

Кроме того, использовались параметры сложности S_{moda} и интенсивности λ.

Наименование	Объект Anylogic	Параметры
Входной поток заявок пользователя	source	Гипотеза о том, что входной поток в рассматриваемой модели является стационарным пуассоновским. λ – интенсивность S_{moda} – мода распределения сложности заявки
Очередь заявок	queue	N – максимальная вместимость очереди заявок. t_{out} – интервал критического времени ожидания ответа от сервера в очереди
Задержка системы на обработку заявки	delay	$t_{serv_delay} = \frac{S}{Q}$ Q – производительность системы.
Не обслуженные заявки из-за вместимости очереди	sink2	Количество не обслуженных заявок из-за вместимости очереди N
Не обслуженные заявки из-за времени ожидания	sink1	Количество не обслуженных заявок из-за времени ожидания t_{out}
Обслуженные заявки	sink	Количество обслуженных заявок

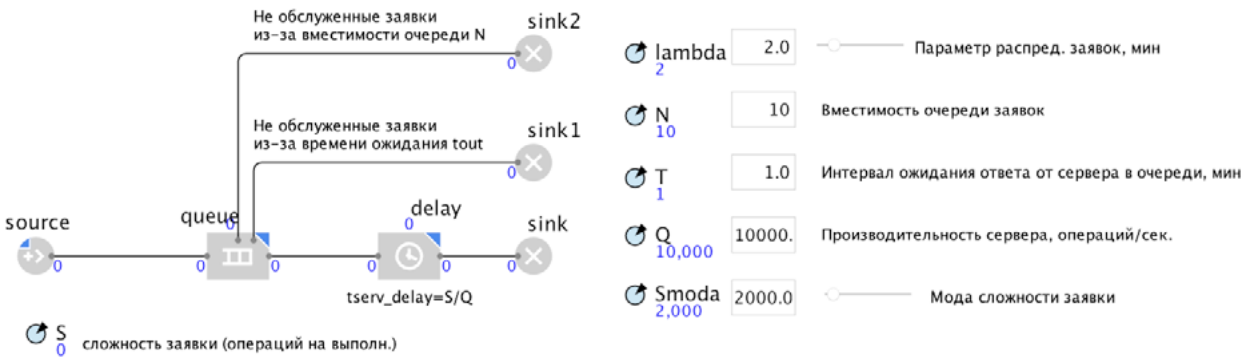


Рис. 3. Снимок экрана реализации модели ГИС для МГН geowheel.ru в ПО Anylogic

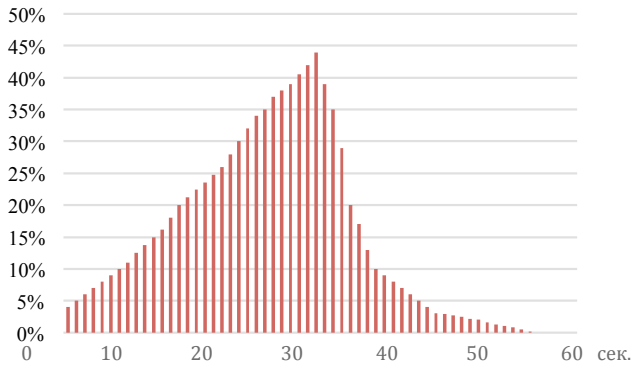


Рис. 4. Распределение времени обработки заявок

В первую группу вошли эксперименты с фиксированной сложностью S_{moda} = ±2000 опер., а интервалы времени между поступлениями заявок распределены по экспоненциальному закону с интенсивностью поступления заявок $\lambda = 2$. В результате экспериментов первой группы была определена средняя длина очереди, средняя загрузка системы, пропускная способность (количество обслуженных заявок), зависимость количества обработанных заявок от времени, распределение времени в системе. На рисунке 4 представлены результаты первой группы экспериментов.

Количество обслуженных заявок – 21 387 заявок (97 %);

Количество отказов по N – 534 заявки (2 %);

Количество отказов по t_{out} – 118 заявок (1 %);

Средняя длина очереди – 1,37;

Средняя загрузка системы – 0,43.

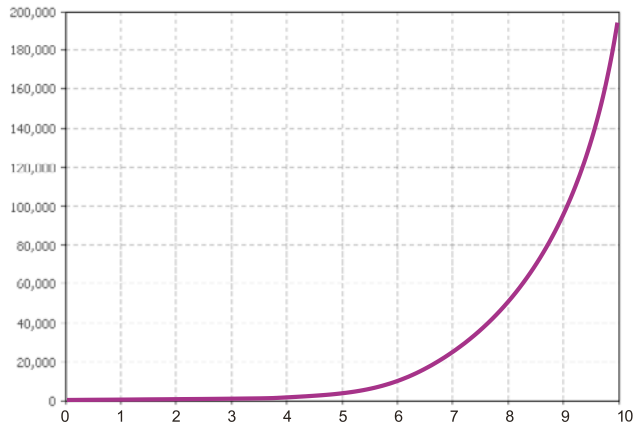
Вторая группа экспериментов – это исследование зависимостей между входными и выходными параметрами модели. В данную группу вошли графики зависимости среднего значения отказов в обслуживании от интенсивности поступления заявок, зависимости среднего значения отказов в обслуживании от сложности заявки.

В третью группу вошли эксперименты, в ходе которых выявлялись пороговые значения входных параметров. Катастрофическим случаем в работе информационной системы является ситуация, при которой процент отказов системы превышает установленный уровень для области применения информационной системы.

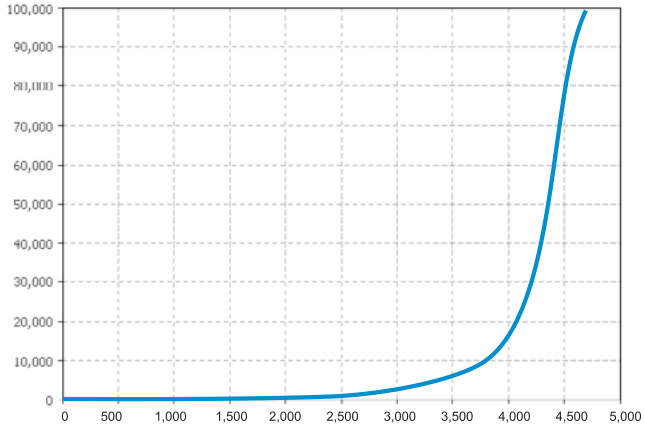
Предложено несколько уровней количества отказов системы L при котором при фиксированной сложности S_{moda} = ±2000 опер. определен критический показатель интенсивности поступления заявок λ. В таблице 2 приведены результаты экспериментов, найдены значения интенсивности для установленных уровней отказов.

При фиксированном показателе параметра интенсивности поступления заявок $\lambda = 2$ и распределённых по экспоненциальному закону интервалов времени между поступлениями заявок были определены критические показатели сложности заявки S для установленных уровней отказов (табл. 3).

Применение технологий имитационного моделирования позволяет комплексно и многопланово проводить оценку эффективности социально-ориентированных ГИС для МГН, при этом под эффективностью системы понимаются такие показатели как средняя длина очереди, средняя загрузка системы, пропускная способность (количество обслуженных заявок), распределение времени в системе, зависимость количества обработанных заявок от времени, зависимость среднего значения отказов в обслуживании от интенсивности поступления заявок, зависимость среднего значения отказов в обслуживании от сложности заявки.



Зависимость среднего значения отказов в обслуживании от интенсивности λ, при заданной сложности S_{moda} = 2000 опер.



Зависимость среднего значения отказов в обслуживании от сложности заявки S_{moda} при заданной интенсивности поступления заявок $\lambda = 2$

Рис. 5. Графики второй группы экспериментов

Таблица 2

Пороговые значения интенсивности λ для уровней отказов

Кол-во отказов, L	*L* = 10 %	*L* = 30 %	*L* = 50 %	*L* = 70 %	*L* = 90 %	*L* = 100 %
Интенсивность, λ	4,45	6,89	8,33	8,93	9,53	≥ 10,68

Таблица 3

Пороговые значения сложности *S* для уровней отказов

Кол-во отказов, L	*L* = 10 %	*L* = 30 %	*L* = 50 %	*L* = 70 %	*L* = 90 %	*L* = 100 %
Мода сложности, S_{moda}	2463	4078	4201	4352	4576	4670

Список литературы

1. Маслова Н.А. Методы оценки эффективности систем защиты информационных систем. – Донецк, 2008.
2. Белых А.А. Основы методологии прогнозирования и оценки эффективности информационных систем. Научный журнал КубГАУ, № 71(07), 2011.
3. С.П. Семенов, В.В. Славский, А.О. Ташкин. Разработка имитационной модели геоинформационной системы для маломобильных групп населения. Вестник Югорского государственного университета, выпуск 3 (46), 2017. – С. 78–85
4. Шеннон, Р. Имитационное моделирование систем – искусство и наука. – М.: Мир, 1978.
5. Кельтон В. Имитационное моделирование. Классика CS [Текст] / В. Кельтон, А. Лоу. 3-е изд. Пер. с англ. – СПб.: Издательская группа BNV, 2004. – 847 с.
6. Семенов С.П., Славский В.В., Ташкин А.О. Анализ информационных ресурсов, направленных на удовлетворение информационных потребностей людей с ограниченными возможностями. Вестник НГУ Серия: Информационные технологии. – 2016. – Том 14, Выпуск № 1. – С. 115. – ISSN 1818-7900.
7. [Электронный ресурс] // Семенов С.П., Ташкин А.О. Методика разработки геоинформационной системы для маломобильных граждан. Современные проблемы науки и образования. – 2014. – № 1; [Офиц. сайт]. URL: http://www.science-education.ru/115-12206 (дата обращения: 15.12.2016).
8. Семенов С.П., Ташкин А.О. Интерактивная геоинформационная система для маломобильных граждан. Сборник научных статей международной конференции «Ломоносовские чтения на Алтае: фундаментальные проблемы науки и образования», Барнаул, 20–24 октября, 2015. – Барнаул: Изд-во Алт. ун-та, 2015. – 3797 с. С. 1007–1010.

УДК 004.62 519.711.3

Н.И. Юсупова, О.Н. Сметанина, Е.Ю. Сазонова, А.И. Агадуллина

DATA MINING ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПО СОВЕРШЕНСТВОВАНИЮ ПСИХОФИЗИЧЕСКОЙ ГОТОВНОСТИ ЧЕЛОВЕКА К УСПЕШНОЙ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ

Аннотация

В статье рассматриваются проблемы организации поддержки принятия решений по совершенствованию психофизической готовности человека к профессиональной деятельности и современное состояние вопроса. Дается характеристика данных, используемых для анализа, и отмечаются их особенности. Приведены экспертные знания в области требований, предъявляемых к психофизической готовности человека к определенной профессиональной деятельности. Рассмотрена постановка задачи, предложена методика по организации информационной поддержки принятия решений в виде рекомендаций по совершенствованию психофизической готовности к предстоящей профессиональной деятельности для обучающихся вузов и их реализации на занятиях по физической культуре. Предложенная методика включает подготовку данных к анализу, выявление новых знаний на основе сходства объектов с использованием кластеризации, их интеграцию с экспертными знаниями, формализацию знаний и формирование базы знаний, получение решения с использованием знаний и механизма логического вывода. Показаны средства Data mining (алгоритмы очистки данных (обнаружение аномалий, заполнение пропусков, выявление дублей и противоречий), аппарат нейронных сетей, в частности, сеть Кохонена)), с использованием которых выявляются неявные знания. Показаны формализация и выбор модели представления знаний с участием эксперта и когнитолога. Приведены результаты экспериментальных исследований на основе предложенной методики. Рассмотрены рекомендации по усовершенствованию психофизической готовности человека к профессиональной деятельности с учетом результатов кластеризации.

Ключевые слова: Data mining, продукционные правила, карты Кохонена, механизм логического вывода, поддержка принятия решений.

Введение

Современное развитие техники и технологий, все возрастающие объемы данных и знаний, возможности получения новых знаний на основе Data mining, возможности использования экспертных знаний, приводят к необходимости использования знаний для принятия решений.

Вопросами формализации знаний занимаются многие специалисты, как в России, так и за рубежом, в частности, Васильев С.Н, Райков А.Н., Gressgard L.J., Nesheim T., Medina Moya J.L., Jarauta Borrasc B. Menegaz J. и др. Вопросы получения неявных знаний на основе Data mining рассмотрены в работах, Евсюкова В.В., Asci F.H., Wei C., Xiaodong L., Lihong M., Liang C., Yu L, Bi Y., Jia L., Wang, QB, Li M. и др. Однако, специфика предметных и проблемных областей требует дополнительных исследований, формализации знаний.

Так проблемы управления психофизической готовностью человека к успешной деятельности рассмотрены многими специалистами, такими как Егорычев Е.А., Шаропин К.А., Pichurin V.V. и др. Как правило, в работах отражены математические модели и методы.

В данной статье авторы предлагают для совершенствования психофизической готовности человека к профессиональной деятельности создать систему продукционных правил. Консеквент правила представлен комплексом специальных упражнений для подготовки, антецедент – номером кластера и граничными значениями одной или нескольких характеристик. Кластер определяется путем проведения кластерного анализа, для проведения которого предполагается использовать аппарат нейронных сетей.

Современное состояние проблемы

Программа «Цифровая экономика», концепция «Индустрия 4.0» предполагают, что все большая часть процессов будет использовать современные IT технологии и средства для автоматизации. Такие системы включают датчики, контроллеры, устройства ввода/вывода, серверы и пр. В более сложных системах с участием человека возникает необходимость принимать решения. Для выполнения своей работы человек должен обладать рядом физических качеств и психических свойств в зависимости от системы. Чем более совершенными будут значения характеристик, тем более эффективно человек будет решать поставленные перед ним задачи. Поэтому проблема организации поддержки принятия решений по совершенствованию психофизической готовностью человека к успешной профессиональной деятельности является актуальной, несмотря на то, что ей занимаются многие специалисты.

Следует отметить результаты исследований в направлениях формализации знаний, использования Data mining для выявления неявных знаний и поддержки принятия решений, описания экспертных знаний в

области совершенствования психофизической готовности человека к профессиональной деятельности.

Исследованиями формализации знаний занимаются многие специалисты, как в России, так и за рубежом, результаты исследований представлены в работах [1–4]. Однако, специфика предметных и проблемных областей требует дополнительных исследований, формализации знаний. Вопросы получения неявных знаний на основе интеллектуального анализа данных рассмотрены в работах [5–13].

Современные тенденции профессиональной подготовки свидетельствует, что целенаправленная психофизическая подготовка специалистов возможна с использованием модели или профессиограммы, которая включает подробное описание условий труда, его характера и специфики [14]. Для профессий с повышенными требованиями к уровню психофизической готовности требуется разработка математической модели, учитывающей связь квалификации, профессионально важных качеств и их взаимовлияние [15]. Егорычев Е.А. также отмечает, что применение математического моделирования структуры высших психических функций позволяет целенаправленно выбирать средства физической культуры и спорта для формирования психологической готовности будущих специалистов к экстремальным условиям труда.

Шаропиным К.А. [16] разработана информационная система оценки психофизической готовности студентов к профессиональной деятельности, отличительной особенностью которой является возможность получения интегральной оценки профессиональной прикладной физической готовности и использования в качестве инструментария тестовых технологий и методов экспертного оценивания. Также автор оценивает необходимость количественного определения уровня обладания специалистом требуемыми качествами [17]. Pichurin V.V. [18] описывает роль физического воспитания в развитии психологической и психофизической подготовки студентов к профессиональной работе. Результаты анализа работ позволили сделать вывод о возможности использования средств Data mining и результатов анализа для поддержки принятия решений в рассматриваемой предметной области.

Характеристика используемых для анализа данных

Основу концепции Индустрии 4.0 составил ряд принципов, в частности, функциональная совместимость человека и машины. Для решения задач данной концепции (в частности, программистом) и ее реализации от человека потребуется определенная психофизическая готовность. Методы оценки профессионально-важных физических качеств и психических свойств рассмотрены авторами ранее [19, 20]. Фрагмент профессиограммы программиста описывает наиболее значимые характеристики профессии (рис. 1) и могут быть названы экспертными знаниями.

Оценить те или иные профессиональные характеристики позволяют специальные тесты. Некоторые

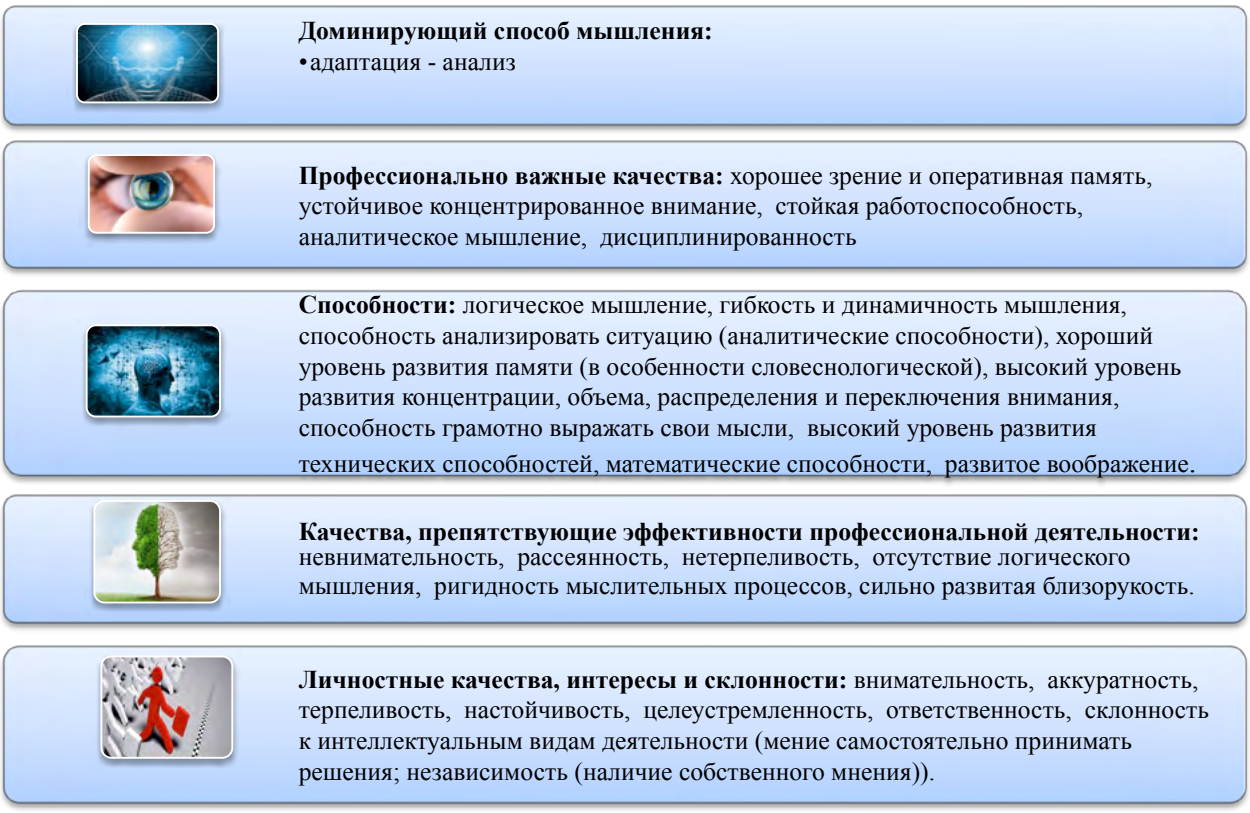


Рис. 1. Фрагмент профессиограммы программиста

Таблица 1

Тесты для оценки профессиональных характеристик (фрагмент)

Наименование теста	Описание метода	Шкала (оценка)/границы
Тест Шульте	Оценка внимания: быстрое отыскание/показ ряда чисел от 1 до 25 в квадрате 5х5клеток при их беспорядочном расположении. Оценка устойчивости внимания: тест выполняется несколько раз, каждый раз с новым вариантом чисел; используется и обратный счет	Отлично – <34 сек; хорошо – 34–42 сек; удовлетворительно – 43–58 сек; плохо – > 58 сек.
Тест Риссу	Определяет устойчивость внимания при его сосредоточении и влияние длительной работы на концентрацию внимания	Отлично – 3 мин. 30 сек., средне – 6–7 мин., плохо – от 13 мин. и выше

из них, в частности, для оценки внимания приведены в таблице 1. Для развития и совершенствования профессионально важных физических качеств и психических свойств имеется определенный состав средств (табл. 2). Приведенный фрагмент средств позволяет улучшить те или иные характеристики, и могут выступать в качестве рекомендации при неудовлетворительной оценке. Знание об использовании указанных средств можно также назвать экспертными знаниями по совершенствованию психофизических качеств.

Вариантом рекомендаций может быть комплекс упражнений. Так, для развития и совершенствования координационных способностей предлагается выполнение следующих групп физических упражнений: упражнения на согласованность движений; упражнения на точность движений; упражнения в прыжках и поворотах.

Упражнения на координацию способствуют развитию и совершенствованию таких психологических качеств, как внимание, мышление и память, поэтому их необходимо развивать совместно.

Таким образом, в качестве данных для анализа выступают результаты тестирования. Рекомендации могут быть как индивидуальными для конкретного человека, так и для группы лиц.

Если ориентироваться на профессионально прикладную физическую подготовку студента, то, как правило, рекомендации даются некоторой группе студентов с близкими значениями показателей.

Группы определяются путем кластеризации. В статье рассматриваются вопросы формирования рекомендаций для групп.

Постановки задачи исследования и методика ее решения

Формальная постановка задачи заключается в выявлении групп студентов с близкими значениями показателей (результатов тестирования) с целью выработки общих рекомендаций по совершенствованию психофизических свойств.

Методика решения задачи (рис. 2) включает четыре этапа. Первый этап направлен на подготовку данных к анализу. Средствами подготовки являются алгоритмы очистки данных (обнаружение аномалий, заполнение пропусков, выявление дублей и противоречий). На втором этапе для выявления сходства объ-

Таблица 2

Состав средств для развития и совершенствования профессионально важных физических качеств и психофизиологических свойств

Качества и свойства беговые / Состав средств		Легкая атлетика (виды)		Виды упражнений								
		прыжковые	координацион-ные	вращения, кувырки	на выносливость	эстафеты		настольный теннис	баскетбол	волейбол	футбол	подвижные игры
Память	оперативная			+			+		+	+		+
	долговременная			+				+		+	+	
Мышление				+			+		+			+
Внимание	концентрация		+	+				+		+		+
	распределение, переключение			+			+	+	+		+	+

Рис. 2. Методика решения задачи

ектов путем кластеризации применяется аппарат нейронных сетей, в частности сеть Кохонена. На третьем этапе выявленные новые знания в виде характеристик схожих объектов позволяют сформулировать рекомендации по совершенствованию психофизических свойств в виде комплекса упражнений. На данном этапе, решения задачи в дополнении к результатам кластерного анализа предлагается использовать знания экспертов, в том числе, и представленные таблицей 2. На завершающем этапе предложенной методики используется сформированная продукционная база знаний. Для реализации методики используются комплексные аналитическая платформа Deductor Studio и оболочка, позволяющая реализовать продукционную систему.

Кластерный анализ данных и результаты эксперимента

Для анализа данных используется аналитическая платформа Deductor Studio, в частности, для кластеризации объектов – нейронная сеть Кохонена. Объекты (их характеристик, т.е. данные о каждом студенте) описаны некоторым вектором признаков. Целью анализа является выявление схожих объектов, для формирования комплекса упражнений для группы студентов. Сеть Кохонена включает входной и выходной слои нейронов. Процедура кластеризация состоит из обучения сети и объединения ячеек в кластеры.

Результаты кластерного анализа (рис. 3) и их интерпретация позволяют сформулировать рекомендации для выявленных групп. Анализ полученных резуль-

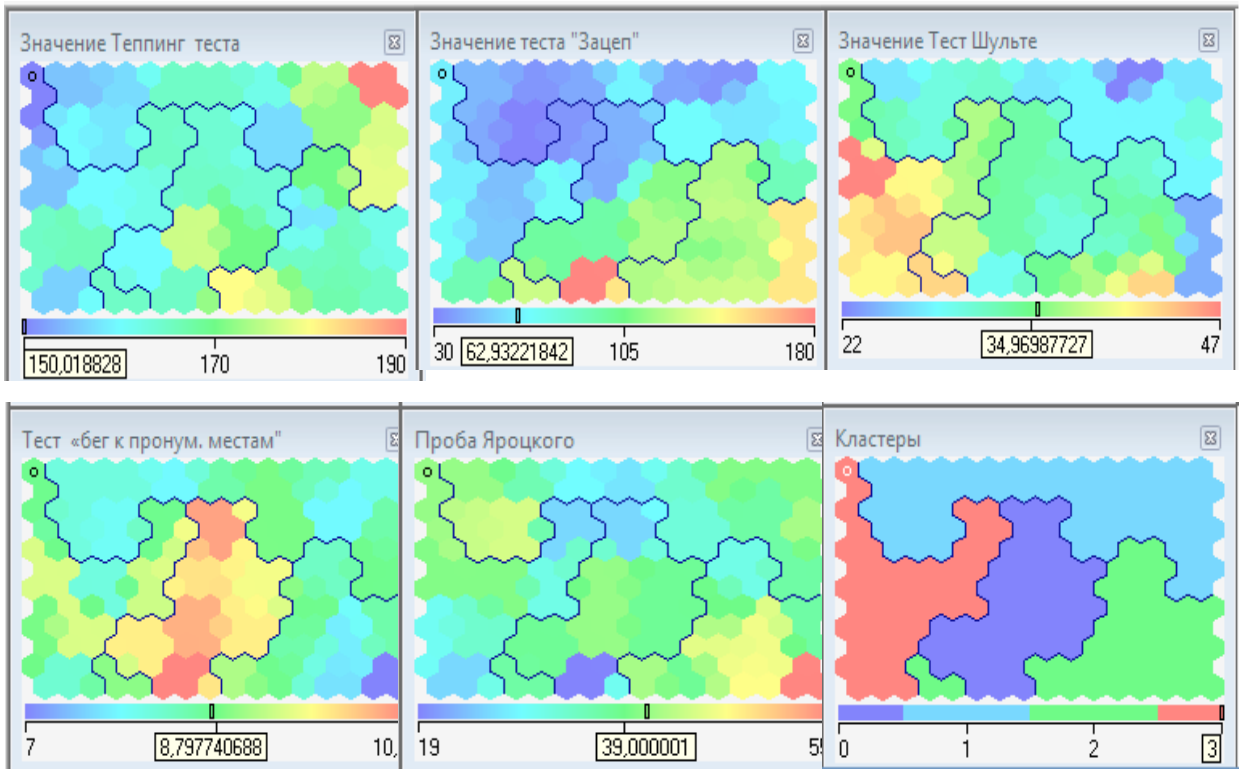


Рис. 3. Результаты кластерного анализа

Таблица 3

Анализ результатов

# кластера	Теппинг тест (характеристика нервной системы)	тест «Зацеп (статическая силовая выносливость мышц кистей рук)»	тест Шульте (характеристика объема, распределения и переключения внимания)	тест «бег к пронумерованным местам (пространственная ориентация и память)»	тест «проба Яроцкого (общее равновесие)»
0	160-180	40-180	28-40	9,75-10,7	19-39
1	165-190	30-63	22-30	7,8-8,8	30-43
2	155-180	105-145	22-40	7-8,8	39-50
3	150-170	45-105	34-47	7,9-9,7	23-39

татов показывает (табл. 3), что студенты, вошедшие в кластер 2, могут выполнять базовый комплекс упражнений. Для студентов, составляющих кластер 1, следует выполнять упражнения, которые позволят развить статическую силовую выносливость мышц кистей рук. Для незначительной части (значение показателя теста «проба Яроцкого» которых менее 39) группы также добавляются упражнения на равновесие. Для студентов, вошедших в кластер 0, характерен комплекс, аналогичный для предыдущего кластера. В дополнение к этому комплексу, необходимы упражнения на пространственную ориентацию и память. Самой «сложной» группой оказался кластер 3. В данном случае необходимо составлять комплекс, способствующий совершенствовать все «характеристики».

Выводы

Несмотря на практически повсеместное использование автоматизации деятельности роль человека остается значительной. Для эффективной и успешной деятельности человек должен обладать рядом психофизических свойств, указанных в профессиограмме. Оценка этих свойств осуществляется на основе периодически проводимого тестирования.

Результаты тестирования являются исходными данными для анализа. При организации подготовки к профессиональной деятельности на занятиях по физическому воспитанию студенты могут быть разбиты на группы с близкими значениями оценок (характеристик) с целью получения общих рекомендаций в виде комплекса упражнений.

Для анализа данных используются средства Data mining, рекомендации формализуются в виде системы продукционных правил. Консеквент правила представлен комплексом специальных упражнений для подготовки, антецедент – номером кластера и граничными значениями одной или нескольких характеристик.

Благодарности

Результаты исследований, приведенные в статье, частично поддержаны грантами РФФИ 19-07-00709, 18-07-00193.

Список литературы

1. Васильев С.Н., 2008. Формализация знаний и управление на основе позитивно образованных языков. Информационные технологии и вычислительные системы, 1: 3–17.

2. Райков А.Н., 2018. Автоматизированный синтез когнитивной модели на основе анализа больших данных и глубокого обучения. Труды 21-й Международной объединенной конференции Интернет и современное обществ, Санкт-Петербург, 2018. Т. 2. С. 103–111.

3. Gressgård L.J. and Nesheim T., 2018. Knowledge Management Systems and Work Improvements: The Moderating Effects of Work Characteristics. Journal of Information & Knowledge Management (JIKM), World Scientific Publishing Co. Pte. Ltd., 17(04): 1–18.

4. Medina Moya J.L., Jarauta Borrasc B. and Menegaz J., 2018. The formalization of professional knowledge in the curriculum. Revistaibero-american adeestudosemeducacao, 13 (2): 588–603.

5. Евсюков В.В., 2014. Интеллектуальный анализ данных как инструмент поддержки принятия решений в системе банковского финансового менеджмента. Известия Тульского государственного университета, 4(1): 374–384.

6. Asci F.H., 2003. The effects of physical fitness training on trait anxiety and physical self-concept of female university students. Psychology of Sport and Exercise, 4: 255–264. http://dx.doi.org/10.1016/S1469-0292(02)00009-2.

7. Wei C., Xiaodong L., Lihong M. and Liang C., 2017. Comparison of the promoting effect of different training programs on the physical and psychological quality of college students based on data mining algorithm. Boletin Tecnico/Technical Bulletin, 55 (19): 583–588.

8. Yu L, 2010. Applying clustering to data analysis of Physical Healthy Standard. In the Proceedings of the 2010 7th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2010, 6, pp. 2766–2768. 10.1109/FSKD.2010.5569224.

9. Yu L., 2009. Association Rules Based Data Mining on Test Data of Physical Health Standard. In the Proceedings of the 2009 International Joint Conference

on Computational Sciences and Optimization (CSO), Sanya, Hainan, China, 2009, pp. 322–324. doi:10.1109/CSO.2009.428.

10. Yu L., 2015. Occupation Oriented and Data Mining Based Personalized Physical Quality Promotion. In the Proceedings of the 2nd International Conference on Information and Communication Technology for Education (ICTE 2015), 2015, pp. 234–237.

11. Bi, Y. Jia, L. and Wang, QB, (2016). Formulation of Scheme for Sports Training Management Based on System Dynamics. In the Proceedings of the 4th International Conference on Machinery, Materials and Information Technology Applications (ICMMITA), 2016, pp. 1500–1503.

12. Li M., 2018. Investigation of Action Decomposition of Track and Field Sports Based on Data Mining. CHIMICA OGGI-CHEMISTRY TODAY, 36(6): 746–749.

13. Pichurin V.V., 2015. Psychological and psycho-physical training as a factor of personal anxiety at students. Pedagogics, psychology, medical-biological problems of physical training and sports, 19: 46–51. 10.15561/18189172.2015.0307.

14. Егорычев Е.А., 2005. Выбор критериев при управлении психофизической готовностью студентов к профессиональной деятельности. Физическая культура. Научно-методический журнал.1. Дата обращения 01.02.2019http://sportlib.info/Press/FKVOT/2005N1/p51-55.htm.

15. Егорычев Е.А., 2005. Теория и технология управления психофизической подготовкой студентов к профессиональной деятельности, Диссертации на соискание ученой степени доктора педагогических наук, Российский государственный университет нефти и газа имени И.М. Губкина, Ярославль.

16. Шаропин К.А., 2007. Информационная система оценки психофизической готовности студентов к профессиональной деятельности, Автореферат диссертации на соискание учетной степени кандидата технических наук, Томский Политехнический Университет, Томск.

17. Шаропин К.А., 2011. Психофизическая готовность студентов к профессиональной деятельности. Информационная система оценки. LAP LAMBERT Academic Publishing,pp.196.

18. Pichurin V.V., 2014.Psychological and psycho-physical training as a part of physical education of students in higher educational establishments. Pedagogigs. Psichology, 11: 44–48.

19. Сметанина, О.Н., М.М. Гаянова, Т.В. Наумова, Р.Ч. Гаянов, 2016. Информационные аспекты профессиональной прикладной физической подготовки студентов. Труды Международной конференции Информационные технологии интеллектуальной поддержки решений, Уфа, май 2016, Т.1. С. 186–191.

20. Сметанина, О.Н., Т.В. Наумова, А.Ю. Адельметова, К.В. Назмиева, 2018.Формализация знаний при поддержке управленческих решений. Труды Международной конференции Информационные технологии интеллектуальной поддержки решений, Уфа, май 2018, Т.3. С. 7–16.

УДК 004.852

Н.И. Юсупова, О.С. Нургаянова

ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ ПРОГНОЗИРОВАНИЯ СВОЙСТВ НИКЕЛЕВЫХ СПЛАВОВ НА ОСНОВЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА

Аннотация

В статье рассматривается вопросы прогнозирования свойств жаропрочных никелевых сплавов с монокристаллической структурой, применяемых в авиационной промышленности для изготовления лопаток газотурбинных двигателей. Приводятся постановка задачи, модели, методы и алгоритмы решения с помощью технологии искусственных нейронных сетей. Даны подробные описания алгоритмов для решения задачи прогнозирования свойств сплавов с помощью нейронных сетей – алгоритм обратного распространения ошибки и генетический алгоритм обучения; сравнение результатов прогнозирования свойств сплавов с помощью нейронных сетей, обученных методом обратного распространения ошибки и генетическим алгоритмом; особенности программной реализации и обсуждение полученных результатов.

Нейронные сети, жаропрочные никелевые сплавы, обратное распространение ошибки, генетический алгоритм обучения.

Введение

Известно, что процесс разработки новых материалов достаточно трудоемкий и включает в себя множество последующих испытаний на исследование различных характеристик. Использование современных методов математического моделирования и прогнозирования позволит ускорить процесс разработки моделей «состав-свойство» и предусмотреть возможные диапазоны попадания исследуемых характеристик.

В задаче прогнозирования свойств никелевых сплавов на основе использования технологии искусственных нейронных сетей важным является выбор метода обучения сети. Основной алгоритм обучения искусственных нейронных сетей – метод обратного распространения ошибки – известен уже достаточно давно и хорошо исследован, но, тем не менее, и сегодня обучение нейронных сетей по-прежнему является нетривиальной и в значительной степени творческой задачей [1, 2]. При решении реальных задач создаются нейронные сети со сложной структурой, в связи с чем, их обучение требует гораздо большего времени [4, 8].

Ранее [7], задача, связанная с прогнозированием свойств сплавов, решалась с помощью многослойного персептрона с количеством входных нейронов равных количеству химических элементов, с обучением на основе метода обратного распространения ошибки.

В данной статье рассматривается генетический алгоритм обучения нейронной сети для решения задачи прогнозирования комплексного свойства – жаропрочности никелевых сплавов и требования к программному модулю, реализующему выбранную архитектуру сети и метод обучения на основе генетического алгоритма. Выбор данного метода обусловлен возможностью поиска наилучшего значения жаропрочности среди большого числа комбинаций химических элементов при наличии внутренних латентных зависимостей в переменных.

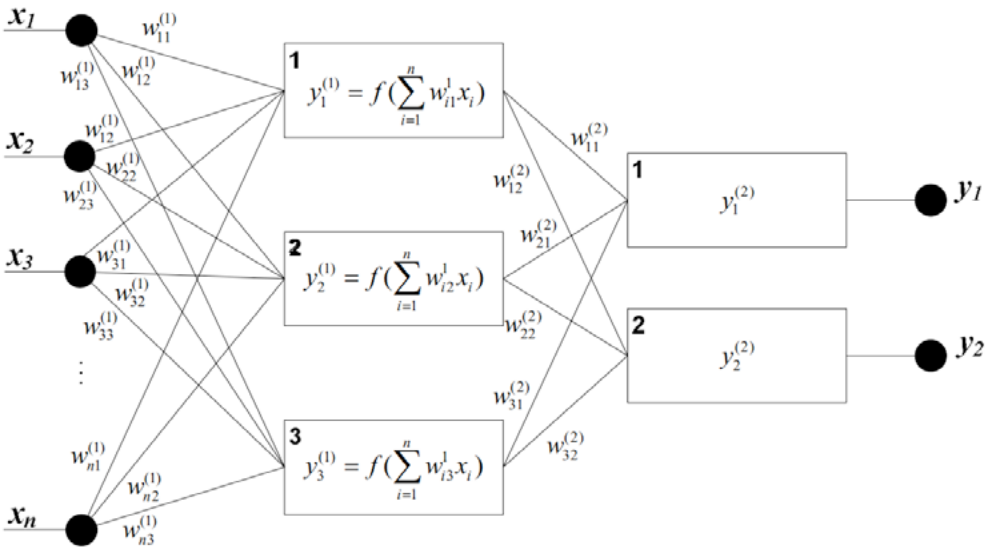


Рис. 1. Двуслойная нейронная сеть с весовыми коэффициентами факторов

Постановка задачи

Дано множество объектов – сплавов, принадлежащих одному классу, которое описывается набором векторов $<x_1,x_2,...,x_n>=y$, где $<x_1,x_2,...,x_n>$ – вектор признаков – состав сплава, представленный процентным содержанием химических элементов, y – комплексная характеристика объекта – жаропрочность, зависящая от вектора $<x_1,x_2,...,x_n>$ – состава сплава. Вид зависимости между входными значениями и выходными не известен. Таким образом, имеем оптимизационную задачу, в которой требуется подобрать такой вектор признаков, которых бы максимизировал функцию жаропрочности – y. На рисунке 1 показана выбранная для решения задачи прогнозирования архитектура нейронной сети:

Анализ известных работ

Разработке и обучению нейронной сети методом обратного распространения ошибки для анализа и интерпретации данных посвящен ряд работ [12–15, 2], в которых так же говорится, что вычислительные эксперименты выявили следующие недостатки: возможность преждевременной остановки из-за попадания в область локального минимума; необходимость многократного предъявления всего обучающего множества для получения заданного качества распознавания; отсутствие приемлемых оценок времени обучения. Исследования показали, что исключить данные недостатки и повысить скорость обработки данных позволяет генетический алгоритм, используемый на этапе обучения нейронной сети. Разработке методики его применения для прогнозирования свойств никелевых сплавов и посвящена данная статья.

Интерпретация терминов генетического алгоритма для задачи обучения нейронной сети

Генетический алгоритм является самым известным на данный момент представителем эволюционных алгоритмов и по своей сути является алгоритмом для нахождения глобального экстремума многоэкстремальной функции. Суть его заключается в параллельной обработке множества альтернативных решений, при этом поиск концентрируется на наиболее перспективных из них [1].

При описании генетических алгоритмов используются определения, заимствованные из генетики. Говоря о популяции особей, в качестве базовых понятий применяются термины: ген, хромосома, генотип, фенотип, аллель.

В терминах обучения нейронной сети основные понятия генетического алгоритма интерпретируются следующим образом:

• ген – весовой коэффициент нейронной сети;

• хромосома – набор генов (т.е. весовых коэффициентов нейронной сети, считываемых в определенном порядке сверху вниз, справа налево). Каждая хромосома является возможным решением (т.е. таким набором весовых коэффициентов, которые лучше подходят для установления зависимости «состав-свойство»);

• популяция – множество хромосом, вариантов наборов весовых коэффициентов;

• эпоха – итерация, соответствующая созданию нового поколения хромосом.

Хромосомы являются основными сущностями, над которыми в определенном порядке в пределах одной эпохи проводятся следующие операции:

• скрещивание – создание с определенной степенью вероятности новой хромосомы из генов двух других и добавление ее в популяцию;

• мутация – изменение с определенной степенью вероятности значения произвольного гена любой хромосомы и добавление ее в популяцию;

• приспособление – удаление из популяции хромосом (т.е. наборов весовых коэффициентов), показавших худший результат при распознавании.

На рисунке 1 показана выбранная архитектура нейронная сеть, весовыми коэффициентами которой (производя обход сверху вниз, справа налево) заполняются хромосомы.

Разработка и реализация алгоритма

Перед тем как приступить к обучению нейронной сети генетическим алгоритмом, необходимо сформировать начальную популяцию, то есть какое-то количество матриц весов $W^{нач}$. Под популяцией будем понимать набор хромосом (решений), которые формируются случайным образом, но с учетом условия $-1 \le w_{ij} \le 1$. Данная матрица имеет размер $[NxN]$. В процессе обучения нейроны могут быть удалены из сети вследствие их малого влияния на работу всей сети. Следовательно, размер матрицы весов может уменьшаться. В случае удаления j-ого нейрона из сети, в матрице весов вычеркивается j-ый столбец и j-ая строка. Далее работа ведется уже с новой матрицей. Рассмотрим основные операторы генетического алгоритма.

Оператор скрещивания, иногда называемый кроссинговером, является основным генетическим оператором (рис. 2), за счет которого производится обмен частями хромосом между двумя (может быть и больше) хромосомами в популяции. Оператор скрещивания может быть одноточечным или многоточечным. Одноточечным называется скрещивание, если родительские хромосомы разрезаются только в одной случайной точке, многоточечным – в случае двух и более точек разреза. Реализация многоточечного скрещивания может быть происходит двумя способами: либо точек разрыва меньше, чем генов в хромосоме, либо – если длина хромосомы L бит, то число точек разрыва равно $(L–1)$ и потомки наследуют биты следующим образом: первому потомку достаются нечетные биты первого родителя и четные биты второго; у второго же потомка все наоборот.

Кроме описанных способов скрещивания существует ещё однородное скрещивание. Его особенность заключается в том, что значение каждого бита в хромосоме потомка определяется случайным образом из соответствующих битов родителей. Для этого вводится некоторая величина $-1 \le P_0 \le 1$, и если случайное число больше P_0, то на n-ю позицию первого потомка попадает n-й бит первого родителя, а на n-ю позицию второго – n-й бит второго родителя. В противном случае к первому потомку попадает бит второго родителя, а ко второму – первого. Такая операция проводится для всех битов хромосомы. Вероятность скрещивания самая высокая среди генетических операторов и равна обычно 60 % и более.

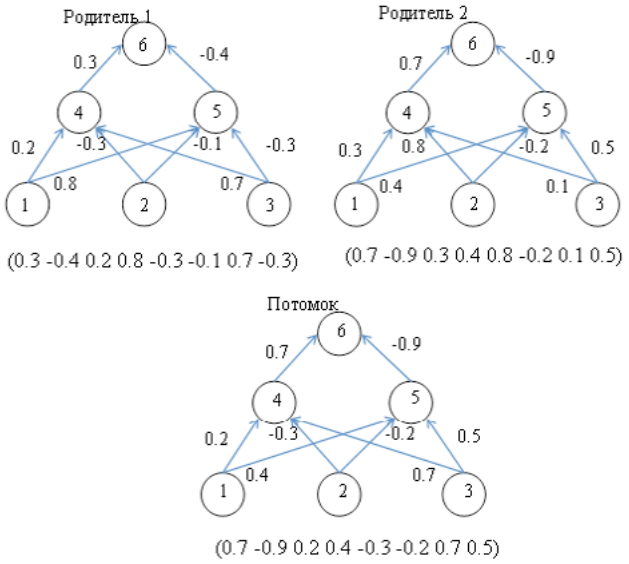


Рис. 2. Работа оператора скрещивания

Работа оператора мутации представляет собой стохастическое изменение части хромосом. Данная процедура необходима для «выбивания» популяции из локального экстремума и способствует защите алгоритма от преждевременной сходимости. Достигается это за счет того, что каждый ген строки, которая подвергается мутации, с малой вероятностью меняется на другой ген (к весу добавляется случайная величина между *–1.0* и *1.0*).

Так же как и скрещивание, мутация проводится не только по одной случайной точке. Можно выбирать некоторое количество точек в хромосоме для изменения, причем их число также может быть случайным.

Также можно изменять сразу некоторую группу подряд идущих точек. Вероятность мутации значительно меньше вероятности скрещивания и редко превышает 1 %. Выше на рисунке 3 показана работа оператора мутации: a) состояние объекта до начала мутации; b) состояние после мутации.

Селекция хромосом заключается в выборе (по рассчитанным на втором этапе значениям функции приспособленности) тех хромосом, которые будут участвовать в создании потомков для следующей популяции, т.е. для очередного поколения. Такой выбор производится согласно принципу естественного отбора, по которому наибольшие шансы на участие в создании новых особей имеют хромосомы с наибольшими значениями функции приспособленности. Существуют различные методы селекции. Наиболее популярным считается так называемый метод рулетки, который свое название получил по аналогии с известной азартной игрой. Каждой хромосоме может быть сопоставлен сектор колеса рулетки, величина которого устанавливается пропорциональной значению функции приспособленности данной хромосомы. Поэтому чем больше значение функции приспособленности, тем больше сектор на колесе рулетки. Все колесо рулетки соответствует сумме значений функции приспособленности всех хромосом рассматриваемой популяции.

Каждой хромосоме, обозначаемой Ch_i для i=1, 2, ...,N (где N – численность популяции) соответствует сектор колеса $v(Ch_i)$, выраженный в процентах согласно формуле:

$$v(Ch_i) = P_s(Ch_i)\ 100\%$$
$$P_s(Ch_i) = F(Ch_i) / \sum_{i=1}^{N} F(Ch_i) \qquad (1)$$

причем $F(Ch_i)$ – значение функции приспособленности хромосомы, а $P_s(Ch_i)$ – вероятность селекции хромосомы (Ch_i). Селекция хромосомы может быть представлена как результат поворота колеса рулетки, поскольку «выигравшая» (т.е. выбранная) хромосома относится к выпавшему сектору этого колеса. Очевидно, что чем больше сектор, тем больше ве-

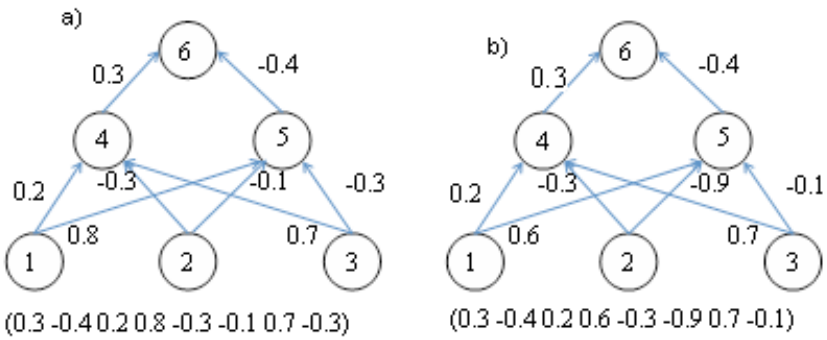


Рис. 3. Работа оператора мутации

роятность «победы» соответствующей хромосомы. Поэтому вероятность выбора данной хромосомы оказывается пропорциональной значению ее функции приспособленности. Если всю окружность колеса рулетки представить в виде цифрового интервала [0, 100], то выбор хромосомы можно отождествить с выбором числа из интервала [A, B], где A и B обозначают соответственно начало и окончание фрагмента окружности, соответствующего этому сектору колеса; очевидно, что $1 \le A < B \le 100$. В этом случае выбор с помощью колеса рулетки сводится к выбору числа из интервала [0, 100], которое соответствует конкретной точке на окружности колеса. Существуют также другие методы селекции. В результате процесса селекции создается родительская популяция, также называемая родительским пулом с численностью N, равной численности текущей популяции.

В процессе работы алгоритма все указанные выше операторы применяются многократно и ведут к постепенному изменению исходной популяции. Поскольку операторы отбора, скрещивания, мутации и редукции по своей сути направлены на улучшение каждой отдельной особи, то результатом их работы является постепенное улучшение популяции в целом. В этом и заключается основной смысл работы генетического алгоритма – улучшить популяцию решений по сравнению с исходной.

После завершения работы генетического алгоритма из конечной популяции выбирается та особь (сплав), которая дает экстремальное (максимальное или минимальное) значение целевой функции и является, таким образом, результатом работы генетического алгоритма. За счет того, что конечная популяция лучше исходной, полученный результат представляет собой улучшенное решение.

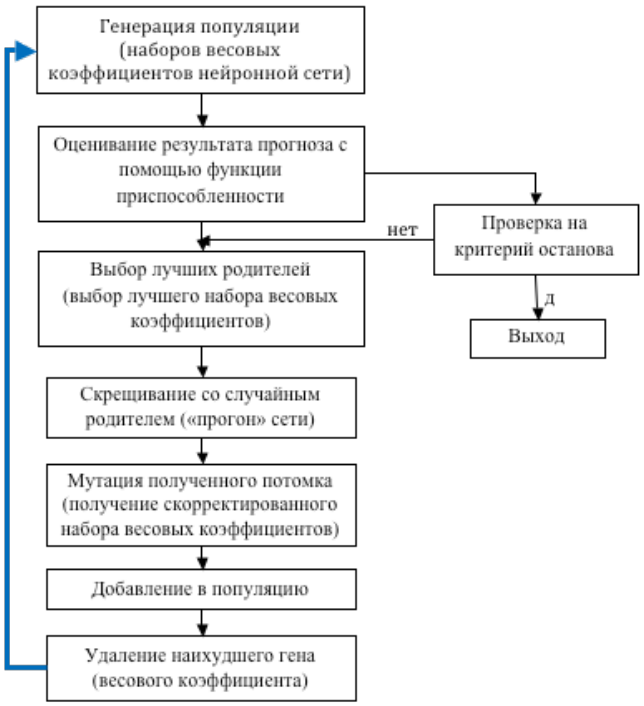


Рис. 4. Схема обучения нейронных сетей с помощью генетического алгоритма

Порядок обучения нейронной сети с применением генетического алгоритма представлен на рисунке 4.

Критерием останова работы алгоритма является достижение целевой функции оптимизации:

$$W=\Sigma\Delta_j \rightarrow max, j \in [0,...,K]. \quad (2)$$

Рассмотрим квадратичную сумму для одного примера из выборки:

$$\Delta_j=\Sigma(Y_i^{id} - Y_i^{real})^2, j \in [1,...,m]. \quad (3)$$

где Y_i^{id} – значение i-ого выхода нейронной сети на обучающем примере данной выборки, Y_i^{real} – значение i-ого выхода, полученное в результате реальной работы нейронной сети. На следующем шаге рассчитывается квадратичная сумма ошибки для каждого из K примеров выборки.

Анализ результатов

Сравнивалась работа нейронной сети, обученной с помощью генетического алгоритма с сетью той же архитектуры, обученной алгоритмом обратного распространения ошибки. На вход двуслойной нейронной сети, состоящей из 256 входных элементов (по 3 нейрона в первом и во втором слое и 1 нейрон в выходном) подавался обучающий набор векторов; обучение считалось законченным, когда во всем обучающем наборе результирующий выходной сигнал совпадал с требуемым эталонным выходным сигналом.

Вычислительный эксперимент показал:

Для обучения нейронной сети с помощью генетического алгоритма достаточно выполнить 100 генераций, в то время при обучении алгоритмом обратного распространения требуется более 500 итераций (где одна итерация – это полный пересчет всех обучающих данных: весовых коэффициентов, погрешности и выходных значений).

Две генерации в генетическом алгоритме эквивалентны одной итерации алгоритма обратного распространения, так как обратное распространение состоит из двух частей прямого прохода (вычисления выхода сети и ошибки) и обратного прохода (изменения весов). Генетические алгоритмы выполняют только первую часть. Таким образом, две генерации занимают меньше времени, чем вычисления единственной итерации алгоритма обратного распространения.

В целом, генетические алгоритмы значительно выигрывают по сравнению со схемой обратного распространения, позволяя получать лучшие векторы весовых коэффициентов за существенно меньшее время.

Выводы

Дальнейшие направления исследований связаны с разработкой более эффективных алгоритмов обучения нейронных сетей, введения в архитектуру сети

нейронов по «переменным протекции» и нейрона смещения [6]. Отдельным направлением исследований можно выделить изучение фазового состава сплавов с помощью методов Data Mining и методов хемоинформатики.

Благодарности

Исследование частично поддержано грантом РФФИ 18-07-00193-а.

Список литературы

1. Д. Рутковская, М. Пилинский, Л. Рутковский, «Нейронные сети, генетические алгоритмы и нечеткие системы», М.: Горячая линия – Телеком, 2006.

2. С. Хайкин, «Нейронные сети: полный курс», Пер. с англ. – М.: Издательский дом «Вильямс», 2006.

3. В. Дюк, А. Самойленко, «Data Mining: учебный курс (+CD)», СПб: Питер, 2001.

4. В. Круглов, В. Борисов, «Искусственные нейронные сети. Теория и практика», М.: Горячая линия–Телеком, 2001.

5. Superalloys II: High-Temperature Materials for Aerospace and Industrial Power 2nd Edition Chester T. Sims, Norman S. Stoloff, William C. Hagel © 1987 by John Willey & Sons Inc.

6. Рахманкулов М.М., Паращенко В.М. Технология литья жаропрочных сплавов. М: Интермет-Инжиниринг, 2000 – 464 с.

7. Нургаянова О.С. Автоматизированное проектирование литейных жаропрочных никелевых сплавов на основе методов искусственного интеллекта: дис. канд. техн. наук. – Уфа, 2006.

8. Нургаянова О.С. Применение искусственных нейронных сетей в задачах классификации многокомпонентных сплавов // Proceedings of the 6th All-Russian Scientific Conference «Information Technologies for Intelligent Decision Making Support», Volume 3, May 28–31, Ufa- Stavropol, Russia, 2018

9. Айвазян С.А., Анализ данных, прикладная статистика и построение общей теории автоматической классификации// Методы анализа данных/ Пер. с фр. – М.: Финансы и статистика, 1985.

10. Mohammed J. Zaki, Wagner Meira Jr. Data Mining and Analysis. Fundamental Concepts and Algorithms. – Cambridge University Press, 2014. 593 с.

11. Rashid T. Make Your Own Neural Network. CreateSpace, 2016. – 222 p.

12. Groumpos, P.P. (2016). Deep Learning vs. Wise Learning: A Critical and Challenging Overview. In IFAC Intern. Conf. on International Stability, Technology and Culture, TECIS 2016, Durres, Albania.

13. Юсупова Н.И., Ризванов Д.А., Сметанина О.Н., Еникеева К.Р. Модели представления знаний для поддержки принятия решений при управлении сложными системами в условиях неопределенности и ресурсных ограничений / Information technologies for intelligent decision making support (ITIDS'2016) Proceedings of the 4th International Conference. 2016

14. Обработка слабоструктурированной информации на основе методов искусственного интеллекта: монография / Н.И. Юсупова, Д.Р. Богданова, М.В. Бойко. – М.: Инновационное машиностроение, 2016. – 250 с.

15. Гузаиров М.Б., Юсупова Н.И., Сметанина О.Н. и др. Методические аспекты искусственного интеллекта (монография) / Уфимский государственный авиационный технический университет, Научный совет РАН по методологии искусственного интеллекта, Рецензенты: профессор В.М. Картак, профессор Д.Е. Сорокин. Москва, 2014.

УДК 004.912 004.822

М.-М. Аббаси, А.П. Бельтюков

АНАЛИЗ ЭМОЦИЙ ИЗ ТЕКСТА НА РУССКОМ ЯЗЫКЕ С ИСПОЛЬЗОВАНИЕМ СИНТАКСИЧЕСКИХ МЕТОДОВ

Аннотация

Эмоции играют важную роль в жизни, и их анализ с использованием синтаксиса представляет собой особый интерес. Для идентификации и анализа текста используются два типа подхода. Один имеет дело с его синтаксисом, а другой – с семантикой. Синтаксис обычно определяет грамматическую структуру текста, тогда как семантика имеет дело со значением текста. Интересно изучить роль семантического и синтаксического анализа в анализе эмоций. Синтаксическая модель основана на структуре аргументов предложений по отношению к глаголам. Синтаксические подходы используются для определения эмоционального держателя в предложениях. В этой статье мы рассматриваем различные типы методов, используемых для синтаксического анализа в целом и, в частности, в тексте на русском языке при выявлении выраженных в тексте эмоций. Мы выявим факторы, которые приводят к особенностям синтаксического анализа в этом случае. В заключении мы предложим технику и механизм, которые могут эффективно использоваться для анализа особенностей синтаксиса с целью выявления эмоций, выраженных в тексте, написанном на русском языке.

Ключевые слова: эмоции, текст, синтаксис, носителями эмоций, золотого стандартного корпуса, обработки естественного языка.

Введение

Эмоциональный анализ из текста из-за большого объема данных, доступных онлайн и офлайн в письменной форме – постоянно увеличивающаяся область исследований. Как правило, для идентификации и анализа текста используются два типа подхода. Один имеет дело с его синтаксисом, а другой – с семантикой. Синтаксис обычно определяет грамматическую структуру текста, тогда как семантика имеет дело со значением текста. Интересно изучить роль семантического и синтаксического анализа в анализе эмоций. Синтаксическая модель основана на структуре аргументов предложений по отношению к глаголам. Синтаксические подходы используются для определения эмоционального содержания в предложениях.

Было замечено, что при переводе текста с одного языка на другой. предложения теряют свою оригинальность. В 2015 году Juan Soler-Company и соавт. разработал глубокую синтаксическую структуру зависимостей, чтобы уменьшить разрыв между поверхностными синтаксическими структурами, создаваемыми анализаторами зависимостей, и семантическими логическими формами текста. Это также сохраняет оригинальную языковую структуру предложения [1].

Одним из основных источников текста в Интернете на разных языках является Википедия. Arazy et al. В 2015 году наблюдалась функциональная роль разных участников в отношениях с текстом в Википедии. Они применяли эмпирические методы для проверки своих теоретических основ [2].

Исследования по извлечению выраженных эмоций важны для того, чтобы различать настроения с разных точек зрения (Seki, 2007). Джесси Аллама и др. разработали предпосылку, основанную на обучении, используя двухфазный подход, сочетающий точный анализ доказательств с алгоритмами машинного обучения для выбора помещения для анализа эмоций [3]. Существует необходимость и возможность использования параллельных и распределенных вычислений в сфере обработки естественных языков (NLP – Natural Language Processing) [4]. Выражение эмоций в записанных текстах на естественном языке, играет важную роль в отношении читателя или писателя. Широкий спектр задач NLP, таких как отслеживание эмоций пользователей в отношении обзоров, событий или политики, выраженных в онлайн-форумах или новостях, управление взаимоотношениями с клиентами – все это использует эмоциональную информацию. Анализ текстового корпуса на основе эмоций может быть использован для обобщения текста [5].

Мы можем определить важные эмоции, которые чаще встречаются в тексте и более похожи на слова, упомянутые в заголовке текста. Для этого рассматриваются два свойства слов. Локальные и глобальные свойства. Локальное свойство – это частота встречаемости слова в тексте, а глобальное свойство – максимальное семантическое сходство между словом и заголовком текста [6].

Предшествующие работы по идентификации выраженных мнений иногда выявляли только одно мнение в предложении (Bethard et al., 2004), а иногда и несколько (Choi et al., 2005). Первый вариант соответствует случаю, когда, например, в блоге есть аннотация для указания эмоционального характера предложения. При этом надо отметить, что длинные предложения могут выражать более одной эмоции, и, вследствие этого, при обработке предложения связываются с несколькими так называемыми *носителями эмоций* (emotion holders – EH – «держатели эмоций»). Все вероятные *носители эмоций* предложения хранятся в специальном *векторе привязки.* Если существуют несколько носителей эмоций, последо-

вательные носители аннотируются и помещаются в вектор в соответствии с их порядком появления. Математическое знание и его представление в математической формуле используется для задач поиска и обучения [7]. Аннотация носителя эмоций на уровне предложения требует знания одного из двух его основных видов связей: явных или неявных. Явная связь квалифицирует один видимый эмоциональный элемент, который непосредственно связан с эмоциональным выражением, тогда как неявные связи квалифицируют все опосредованно прямо или косвенно связанные с ним носители эмоций.

Для синтаксического анализа с целью эмоциональной оценки может быть использована обычная иерархическая синтаксическая структура, хотя она и слабо мотивирована теорией оценки эмоций. Уровни в этой иерархической структуре тщательно разрабатываются, чтобы упростить задачу классификации на верхнем уровне и задержать её решение между очень неоднозначными классами до конца анализа. Частота, с которой мы используем разные слова, постоянно меняется, это происходит всякий раз, когда появляется новый лексический предмет или изымается из употребления другой [8]. Классификация текста на уровне документа или уровне предложения часто не дает необходимой информации. Синтаксический анализ использует концепцию эмоционального анализа уровня *аспекта.* Он направлен на классификацию настроений в отношении конкретных аспектов сущностей. Первый шаг – определить сущности и их аспекты. Владельцы мнений могут высказывать разные мнения по различным аспектам одного итого же предмета в предложении. «Качество звука этого телефона не очень хорошо, но срок службы батареи длинный».

Эмоциональные выражения не существуют изолированно. Они являются функцией настроения, личности, окружающей среды и рассматриваемой задачи [9]. Поэтому необходимо создать специальные модели эмоций для снижения присущей аффективной изменчивости. Это применяется, например, в растущем числе интерактивных образовательных систем, коммерчески доступных и находящихся в разработке [10, 11].

Базовая Система

Базовая система идентификации носителей эмоций на английском языке была создана для носителей эмоций на аннотированном корпусе VerbNet, построенном для английского языка. Роли, связанные с эмоциями, присутствующие в VerbNet (например, Agent, Experiencer, Beneficiary и Actor), были использованы для пометки элементов в полученных предложениях для подготовки *золотого стандартного корпуса.*

Для идентификации носителей эмоций были разработаны два типа базовых систем; первая была разработана на основе лексических паттернов, основанных на частях речи (POS – part of speech), а вторая была основана на синтаксических отношениях, которые были извлечены либо из анализатора зависимостей, либо из неглубокого синтаксического анализатора.

Модель, основанная на лексике, была разработана на основе фокальных шаблонов лексического уровня. Лексические шаблоны группируются в соответствии со своими частями речи (POS). В случае английского языка базовая система была разработана только на основе существительных. Существительные были рассмотрены для идентификации носителей эмоций. Предварительное наблюдение показало, что теги POS действительно содержат похожие шаблоны на лексическом уровне для носителей эмоций.

Рассматривается лексическая картина, состоящая из двух основных членов предложения: подлежащего и сказуемого. Дополнительные члены предложения рассматриваются без какого-либо лексического упорядочения.

Синтаксическая система

Синтаксис предложения является важным ключом для определения носителей эмоций, присутствующих в тексте. Более конкретно, структура аргументов глагола или информация о дополнительной категоризации играют важную роль для идентификации носителей эмоций в предложении. Кадр субкатегоризации представляет собой утверждение о том, какие типы синтаксических аргументов принимает глагол (или прилагательное), такие как дополнения, причастные обороты и т.д. Английская версия VerbNet (Kipperschuller, 2005) была идентифицирована как лексика субкатегоризации с открытым исходным кодом.

Важность

Самый большой источник данных, доступных в Интернете, находится в форме обзоров. Эти обзоры могут касаться продуктов, социальных проблем, экономических проблем и т.д. Обзоры важны для владельцев бизнеса, так как они могут принимать бизнес-решения в соответствии с результатами анализа мнения пользователей о своих продуктах. Эмоциональный анализ может также применяться на фондовых рынках [12, 13], в новостных статьях, [14] или политических дебатах [15]. Например, в политических дискуссиях мы могли бы определить мнение людей о некоторых кандидатах на выборах или политических партиях. Результаты выборов также могут быть предсказаны по политическим постам. Сайты социальных сетей и сайты с микро-блогами считаются очень хорошим источником информации, потому что люди представляют и свободно обсуждают свои мнения по определенной теме. Используя синтаксис, мы можем идентифицировать и конкретного автора текста.

Синтаксические особенности, характерные для произведений конкретного писателя, например, син-

таксис Гоголя, синтаксис Достоевского и т.д. Важно отметить, что стиль писателя во многом определяется синтаксисом. Есть многочисленные исследования по языку и стилю писателя, в частности, его синтаксис (работы академика В.В. Виноградова, например, «Стиль пиковой дамы», великая работа Е.А. Иванчиковой «Синтаксис художественной прозы Достоевского», Г.О. Винокур, Б. Эйхенбаум, Ю.М. Лотман и другие ученые).

Связанные работы

Работа по маркировке аргументов глаголов с их семантическими ролями с использованием новой техники сопоставления кадров упоминается в (Swier and Stevenson, 2005 [16], [17]). Идентификация предложений мнений и их держателей описана в (Bethard et al., 2004 [18]) в основном для глаголов. Идентификация носителей эмоций для ответа на вопрос с поддержкой аннотационной задачи была предпринята с самого начала (Wiebe et al., 2005 [19]). В работе [20] (Choi et al., 2005) использовали названные объекты (NE) для идентификации носителей эмоций с помощью машинного обучения и методов на основе шаблонов. Основываясь на традиционных перспективах, другая работа, обсуждаемая в (Hu et al., 2004 [21]), использует базу знаний эмоций для извлечения держателя эмоций. Задача классификации на основе машинного обучения «не держатель», «слабый держатель», «средний держатель» или «сильный держатель» выполняется в (Evans, 2007 [22]). Ким и Хови (Kim and Hovy, 2006 [23]) определили держателя мнения с темой из медиатекста с использованием семантической роли. Метод идентификации носителей эмоций, основанный на анафоре, использующий лексическую и синтаксическую информацию из онлайн-новостных документов, выполняется в (Kim et al., 2007 [24]). Синтаксические модели идентификации носителей эмоций для английских эмоциональных глаголов обсуждаются в (Das and Bandyopadhyay, 2010 [25]).

Работы на русском языке

В случае русского языка семантический и синтаксический язык сильно отличается от других языков. Лаура А. Джанда и Валерий Д. Соловьев в 2009 году определили конструкционный профиль как частотное распределение слов для русских существительных, обозначающих грусть и счастье, на основе данных о корпусе и проанализировали его количественно [26]. На сегодняшний день существует несколько работ ученых по анализу эмоций из синтаксиса текста. Четтьеркин И.И., Браславский П.И., Лукачевич Н.В. (2013) [27] и Пазельская А.Г., Соловьев А.Н. (2011) [28] внесли большой вклад в изучение данной проблемы. Благодаря богатой морфологии и свободному порядку слов синтаксис русского языка более сложный. Формула, построенная для определения потенциальных членов в качестве носителей эмоций:

$$PMI(potential\ term, text\ class) = -\log_2(P(term, text\ class))P(term) \times P(text\ class) \quad (1)$$

В этой формуле класс текста (text class) представляет текст, который может быть предложением или абзацем. Термин (term) или слово является членом одного из классов и является потенциальным термином (potential term), который может быть общим существительным или носителем эмоций. Функция PMI представляет собой точечную взаимную информацию (Pointwise Mutual Information) между потенциальным термином и текстовым классом, где он был найден, а P – вероятность.

Здесь носитель эмоций может быть больше, чем просто существительное. Это могут быть общие существительные, которые обозначают объекты класса. Например, для цифровых камер такими терминами являются «камера», «фототехника», «аппарат», «фотоаппарат» и т.д. Это могут быть правильные имена объектов класса. Имена моделей камер, или это могут быть общие существительные, которые обозначают параметры, свойства и части объектов класса. В качестве иллюстрации для цифровых камер параметры «формфактор», «качество фото», «разрешение», «матрица», «объектив», «вспышка» и т.д. Для фильмов и книг: «автор», «режиссер», «игра актеров», «атмосфера», «дубляж» и т.д.

Мавлютов Р.Р. и Остапук Н.А. [29] в своей работе над корпусом русского языка отмечали, что существует набор специфических синтаксических отношений, обычно используемых для выражения субъективности. Они отметили, что 80 % субъективных оценок являются грамматическими модификаторами, выраженными прилагательными, например, «Громкая музыка», «плохое обслуживание», 7 % предикатов выражались по-разному: «бармен кричал», «обслуживание было плохим», «обслуживание оставляло желать лучшего», «отель чудовищен», «отношение к клиентам просто ужас». 4 % наречия, выраженные наречиями и пропозициональными фразами, связанными с предикатом, грамматическим модификатором или непосредственно с объектом: «кран работал плохо», «плохо работающий кран», «связь на троечку» и 9 % другими способами. Эти способы включают выражение субъективности с междометиями («брр», «фууу» и т. д.). Сравнение объектов («А лучше Б», «А понравился меньше, чем Б»), ссылка на себя («мне стало плохо», «Я замучался его смотреть») и выражения, в которых объект и фраза мнения не связаны синтаксически («Вчера посмотрел этот фильм. До сих пор противно»)

Как правило, для синтаксического анализа каждое предложение преобразуется в факт, имеющий четыре разных значения.

- Объект из тезауруса
- Тип синтаксической связи между объектом и другой частью предложения

- Связанная часть предложения
- Отрицание

Например, исходная фраза – «Неделю назад я купил водонепроницаемую камеру от Nikon». В этом предложении объектом является «камера». Из всех синтаксических связей объекта только один может потенциально выражать субъективность (грамматический модификатор), поэтому будет извлечен один факт:

1. Объект: «камера»
2. Отношение: грамматический модификатор
3. Родственная часть: «водонепроницаемый»
4. Отрицание: ложь

Кроме того, некоторые суждения не основываются на мнениях. Например, давайте рассмотрим фразу «Брюс уже не тот». Поля факта:

1. Объект: «Брюс»
2. Отношение: предикат
3. Родственная часть: «тот»
4. Отрицание: правда

В русском языке отрицание выражается по-разному для разных частей речи. Поэтому для каждого типа синтаксических отношений в фактах мы написали другой набор правил для извлечения отрицаний.

Примеры:

1. «нет», «без», «отсутствие», «лишенный», «лишивший», «мало», «никакой», «ни» + noun in genitive case
2. «не», «мало» + verb in a finite form
3. «нельзя», «невозможно» + verb in an infinite form
4. «не» , «мало», «ничего» + adjective
5. «не» + adverb, preposition phrase

Наличие «не» (частица отрицания) не обязательно выражает отрицание. Например, выражение «не только мерзкий» не меняет знак «мерзкий». Поэтому мы также описали класс выражений, где слова отрицания не выражали отрицания.

Смыслы после анализа могут быть абстрактными чувствами, которые имеют отношение ко всему тексту, или они относятся к некоторому конкретному объекту в тексте. Отмечается, что классический метод обработки текста как синтаксический анализ может значительно повысить эффективность анализа настроений.

Выводы

Было замечено, что подходы, которые включают синтаксические отношения, в основном разработаны для английского языка [30, 31]. Для русского языка задача построения синтаксического дерева намного сложнее, учитывая богатую морфологию и свободный порядок слов.

Реализован синтаксический подход к анализу чувств для русского языка [32]. Эта система нацелена на определение тона текста новостей. Он извлекает объект оценки, а также синтаксические группы с мнением слов и в соответствии с некоторым набором правил объединяет их.

Методы, используемые для синтаксического анализа текста с других языков, не могут быть просто использованы для синтаксического анализа текста с русского языка. Это связано с природой русского языка.

В этом обзоре и исследовании представлены подробные механизмы синтаксического анализа текста, написанного на русском языке. Компоненты, которые требуют особого внимания, такие как характер отрицаний, окончание слов, идентификация носителя эмоций, подробно объясняются в этом исследовании.

В основном в теориях и методологиях синтаксического анализа, как упоминалось выше, используются собственные и общие имена и их параметры. При этом либо пропускаются другие типы существительных, либо они рассматриваются как часть общего класса существительных. Это усложняет процесс анализа и влияет на точность его результатов, так как не учитываются категории этих существительных. Например, это может быть абстрактное существительное, которое является названием действия, идеи, физического состояния, качества или состояния ума. Другая категория – конкретные существительные, относящиеся к людям и к вещам, которые существуют физически и которые можно увидеть, потрогать, понюхать, услышать или попробовать. Есть также и собирательные существительные, которые относятся к группам людей или вещей, например, аудитория, семья, правительство, команда, жюри. Использование всех этих категорий с их собственной идентификацией будет полезно для улучшения алгоритма.

Список литературы

1. Juan S.C., Miguel B., Bernd B., Simon M and Leo W, 2017. Visualizing Deep-Syntactic Parser Output. In the Proceedings of the 2015 NAACL-HLT , Denver, Colorado, May 31 – June 5, 2015, pp. 56–60. SJR-2017 0.36 https://www.aclweb.org/anthology/N15-3000.

2. Arazy O., Felipe N., Oded Y., Melanie B and Adam, 2015. Functional Roles and Career Paths in Wikipedia. In the Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work & Social Computing, 2015. Vancouver, Canada March 14–18 2015, pp.1092–1105. SJR 0.29.

3. Jesse A., T. Heskes, D. Kühlwein, E. Tsivtsivadze and J. Urban, 2014. Premise Selection for Mathematics by Corpus Analysis and Kernel Methods. Journal of Automated Reasoning, 52 (2):191–213. https://link.springer.com/article/10.1007/s10817-013-9286-5. Q1, SJR 0.65.

4. Solovyev V., V. Polyakov, V. Ivanov, I. Anisimov and A. Ponomarev, 2013. An approach to semantic natural language processing of Russian texts. Research in Computing Science, 65: 65–73. https://pdfs.semanticscholar.org/a4d7/d89ca769a447bcac8db006a008b6666fd8bb.pdf#page=65.

5. Ashmita S and B. Ruhi, 2015. Auto Text Summarization with Categorization and Sentiment Analysis. International Journal of Computer Applications, 130: 07–12. https://pdfs.semanticscholar.org/560b/3aa8667c8b9f65e45d38c00d90f964ea6eea.pdf.

6. Maryam K., N.D Mohammad, M.S Sayed, 2012. A Novel Method of Significant Words Identification in Text Summarization. Journal of Emerging Technologies in Web Intelligence, 4 (3): 252–258. Q4 SJR 0.13.

7. Elizarov A.M., A.V. Kirillovich, E.K. Lipachev, O.A. Nevzorova, V.D. Solovyev and N.G. Zhiltsov, 2014. Mathematical knowledge representation: semantic models and formalisms. Lobachevskii Journal of Mathematics, 35 (4): 348–354. https://link.springer.com/article/10.1134/S1995080214040143 Q3, SJR 0.23

8. Vladimir B., V. Solovyev and S. Wichmann, 2014. Universals versus historical contingencies in lexical evolution. Journal of the Royal Society Interface, 11(101): 08–41. https://doi.org/10.1098/rsif.2014.0841 Q1, SJR 1.72

9. Leslie R.B and J. Hall, 2009. Gender, Emotion and Socialization. In: Chrisler J., McCreary D. (edition) Handbook of Gender Research in Psychology. Springer, New York, NY, pp: 249–454.

10. Takayuki K., T. Hirano, D. Eaton and H. Ishiguro, 2004. Interactive robots as social partners and peer tutors for children: A field trial. Human-Computer Interaction, 19 (1): 61–84.https://www.tandfonline.com/doi/abs/10.1080/07370024.2004.9667340 Q2. SJR-2017 0.52

11. Ashish K., and R. Picard, 2005. Multimodal affect recognition in learning environments. In Proceedings of the13th annual ACM international conference on Multimedia. ACM New York, NY, USA, 677–682. https://dl.acm.org/citation.cfm?id=1101300 (Scopus).

12. Liang C.Y., W.J. Long, C.P. Chann and C.H. Shou, 2013. Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news. Knowledge Based System, 41:89–97. https://doi.org/10.1016/j.knosys.2013.01.001 Q 1 SJR-2017 1.38.

13. Michael H., M. Liebmann and D. Neumann, 2013. Automated news reading: Stock price prediction based on financial news using context-capturing features. Decision Support Systems, 55 (3): 685-697. https://doi.org/10.1016/j.dss.2013.02.006 Q1 SJR-2017 1.66.

14. Tao X., P. Qinke and C. Yinzhao, 2012. Identifying the semantic orientation of terms using SHAL for sentiment analysis. Knowledge Based Systems, 35: 279–89. DOI: 10.1016/j.knosys.2012.04.011 Q 1 SJR-2017 1.38.

15. Isa M. and V. Piek, 2012. A lexicon model for deep sentiment analysis and opinion mining applications. Decision Support Systems, 53(4):680–688. https://doi.org/10.1016/j.dss.2012.05.025 Q1 SJR-2017 1.66

16. Robert S.S and S. Stevenson, 2005. Exploiting a Verb Lexicon in Automatic Semantic Role Labeling. In Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural (HLT/EMNLP), Vancouver, 883–890. http://aclweb.org/anthology/H05-1111 (Scopus).

17. Steven B., H. Yu, A. Thornton, V. Hatzivassiloglou, and D. Jurafsky, 2004. Automatic Extraction of Opinion propositions and their Holders. AAAI Spring Symposium on Exploring Attitude and Affect in Text, 224.

18. Bethard S., H. Yu, A. Thornton, V. Hatzivassiloglou, and D. Jurafsky, 2004. Automatic Extraction of His Holders In: Exploring Attitude and Affects (AAAI-EAAT 2004), Stanford University.

19. Wiebe J., T. Wilson, and C. Cardie, 2005. Annotating expressions of opinions and emotions in language. Language Resources and Evaluation, 39 (2): 165–210. https://link.springer.com/article/10.1007/s10579-005-7880-9 Q2, SJR 2017-0.12

20. Breck E., Y. Choi, and C. Cardie, 2007. Identifying expressions of opinion in context. In Proceedings of the Twentieth International Joint Conference on Artificial Intelligence, 7: 2683–2688. http://www.aaai.org/Papers/IJCAI/2007/IJCAI07-431.pdf SJR-2017 0.5.

21. Ye in, C. Cardie, E. Riloff and S. Patwardhan, 2005. Identifying Sources of Opinions with Conditional Random Fields and Extraction Patterns. In Proceedings of the 5th Conference on Human Language Technology and Empirical Methods in Natural Language Processing HLT/ EMNLP, Vancouver, Canada, 355–362. doi>10.3115/1220575.1220620 (Scopus).

22. Matthew H and K. Nigam, 2004. Retrieving topical sentiments from Online document collection. In Proceedings of Document Recognition and Retrieval, 11: 27-34. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.84.9337.

23. David K.E, L.W. Ku, Y. Seki, H.H. Chen and N. Kando, 2007. Opinion analysis across languages: An overview of and observations from the NTCIR6 opinion analysis pilot task. International Workshop on Fuzzy Logic and Applications, 4578: 456-463. https://link.springer.com/chapter/10.1007/978-3-540-73400-0_57.

24. Soo-Min K. and E. Hovy, 2006. Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text. In: Proceedings of Workshop on Sentiment and Subjectivity in Text. Association of Computer Linguistics, Sydney, Australia, pp.1-8. https://dl.acm.org/citation.cfm?id=1654642.

25. Soo-Min K., and E. Hovy. Crystal: Analyzing predictive opinions on the web, 2007. In Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), Association for Computer Linguistics, Prague, June 2007, pp. 1056–1064. http://www.aclweb.org/anthology/D07-1113 (Scopus).

26. Das D. and S. Bandyopadhyay, 2010. Sentence Level Emotion Tagging on Blog and News Corpora. Journal of Intelligent System (JIS), 19(2) : 125-134. https://www.aclweb.org/anthology/Y10-1013.

27. Laura A.J and V.D Solovyev, 2009. What constructional profiles reveal about synonymy: A case

study of Russian words for SADNESS and HAPPINESS. Cognitive Linguistics, 20 (2): 367-393. DOI: https://doi.org/10.1515/COGL.2009.018 Q1, SJR 1.03.

28. Lukashevich N. and I. Chetverkin, 2013. Build a model for sentiment lexicon extraction in different subject areas. Modeling and analysis of Information Systems, 20(2): 70-79. https://istina.msu.ru/publications/article/4707093/.

29. Пазельская А.Г., and А.Н. Соловьев, 2011. A method of Sentiment Analysis in Russian Texts. Computational linguistics and intellectual technologies. (Dialogue-2011) 4:510-552. http://www.dialog-21.ru/digests/dialog2011/materials/ru/pdf/50.pdf.

30. He H., A. Grissom, J. Morgan, J.B Graber and H. Daumé, 2015. Syntax-based Rewriting for Simultaneous Machine Translation. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 17–21 September 2015 pp. 55-64. http://www.aclweb.org/anthology/D15-1006.

31. Shilpa A., E. Mayfield, C.P. Rose and E. Nyberg, 2010. Sentiment classification using automatically extracted subgraph features. In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Los Angeles, California, June 05 – 2010, pp. 131-139. https://dl.acm.org/citation.cfm?id=1860647.

32. Yi J., T. Nasukawa, W. Niblack and R. Bunescu, 2003. Sentiment analyzer: extracting sentiments about a given topic using natural language processing techniques. In Proceedings of the 3rd IEEE international conference on data mining (ICDM 2003), Florida, USA, November 19–22, pp. 427–434. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.102.7125.

УДК 004.93 004.89

И.А. Ботыгин, В.С. Шерстнёв, А.И. Шерстнёва,
Н.А. Солтаганов, В.А. Тартаковский

ИСПОЛЬЗОВАНИЕ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПОСТРОЕНИЯ ПРОГНОЗНЫХ МОДЕЛЕЙ МЕТЕОПАРАМЕТРОВ

Аннотация

В работе представлены программные эксперименты по построению прогнозных моделей с использованием одного из современных подходов в компьютерном моделировании, а именно, сложных технологических архитектур искусственных нейронных сетей. Для машинного обучения использовался фреймворк TensorFlow с надстройкой Keras в программной среде Jupyter Notebook. Для обработки и визуализации массивов входных данных и результатов эксперимента были использованы библиотеки языка программирования Python – Pandas, Numpy, Matplotlib. Надстройка Keras позволила спроектировать рекуррентную нейронную сеть, в логике которой нейроны взаимодействуют сами с собой. Это позволило смоделировать обучение искусственной нейронной сети на последовательных пространственных цепочках, т.е. на повторяющихся годовых выборках метеоданных. Эксперименты были проведены на открытых метеорологических данных, полученных из Всероссийского научно-исследовательского института гидрометеорологической информации – Мировой центр данных. Для обучения модели использовались девяносто процентов имеющихся данных. На оставшихся десяти процентах был проведен сравнительный анализ точности предсказаний значений исследуемых метеопараметров. Модель достаточно хорошо отображала общие колебания параметров, но пиковые значения иногда были неточными. В некоторых случаях форма линейного графика была предсказана достаточно точно, хотя были небольшие отставания от линейного графика фактических данных. Для повышения точности прогнозных моделей были использованы функции фильтрации входных данных с целью исключения аномальных выбросов и отсутствующих данных.

Ключевые слова: рекуррентная нейронная сеть, машинное обучение, метеорологические данные, фреймворк TensorFlow, надстройка Keras.

Введение

Научный прогноз погоды всегда актуален для мирового сообщества и представляет ценность не только для многих отраслей народного хозяйства, но и непосредственно для людей. Идет постоянное улучшение существующих моделей и поиск новых подходов, особенно, в численных моделях прогнозирования погоды. В частности, одним из таких перспективных подходов является машинное обучение, когда прямое решение задачи заменяется на процесс обучения по множеству сходных задач и последующей выдачей по возможным входным данным адекватного классифицирующего ответа, причем даже для данных, выходящих за пределы имеющейся обучающей выборки [1–6].

Наиболее эффективным способом реализации вычислительного интеллекта в машинном обучении является использование искусственных нейронных сетей (ИНС), позволяющих достаточно точно отражать нелинейный характер прогнозируемых явлений и событий [7, 8].

В настоящее время искусственные нейронные сети нашли применение и продемонстрировали свою эффективность во многих задачах. Исследовано применение ИНС для обработки текста на естественном языке и его классификации [9], для анализа тональности текста [10], для распознавания речи [11]. Используются нейронные сети для подавления шума [12]. В [13] описывается применение ИНС для автоматизации вождения (беспилотные автомобили). Безусловно, для создания полностью автономных систем управления транспортом еще далеко, но успехи беспилотных авто впечатляют. Корейские ученые использовали нейросетевой подход для преобразования объектов на фотографиях без искажения фона [14]. В [15] приведены прогнозные оценки, выполненные на базе нейросети, запоминания человеком изображений естественной сцены, в которых нет запоминающихся образов высокого уровня. Итальянские исследователи предложили метод анимации путем переноса движения из видео на целевое изображение [16]. Исследователи из Университета Западной Австралии создали сеть ClaRAN, которая определяет расположение галактик в глубоком космосе, анализируя излучение черных дыр, полученное с помощью радиотелескопов [17]. В [18] с использованием метода глубокого обучения представлено восстановление изображение лица в высоком разрешении из изображения с низким разрешением. В [19] использование нейросетей проиллюстрировано в исследованиях, связанных с коммуникационными интерфейсами мозга. На базе методов глубокого обучения разработана автоматизированная система подтверждения личности [20]. Сверточные нейронные сети используются командой разработчиков IBM при разработке программного обеспечения датчика, который контролирует состояние здоровья пациента, отслеживая малейшие движения и деформации ногтя [21].

Ввиду высокой точности, самыми популярными видами ИНС являются сверточные и рекуррентные нейронные сети. Именно использованию рекуррент-

ных нейронных сетей для построения прогнозных моделей метеорологических параметров и посвящена настоящая работа.

Экспериментальная часть

В экспериментах использовались метеорологические данные с 2015 г. по 2017 г. с метеостанции с синоптическим индексом 29430 (г. Томск). Набор данных включал восьми срочные наблюдения следующих метеорологических параметров (компонентов): сумма осадков, температура поверхности почвы, температура воздуха по сухому термометру, парциальное давление водяного пара, относительная влажность воздуха, температура точки росы, атмосферное давление на уровне станции.

Ниже (рис. 1, 2, 3) представлены исходные метеорологические параметры, для которых строились прогнозные модели.

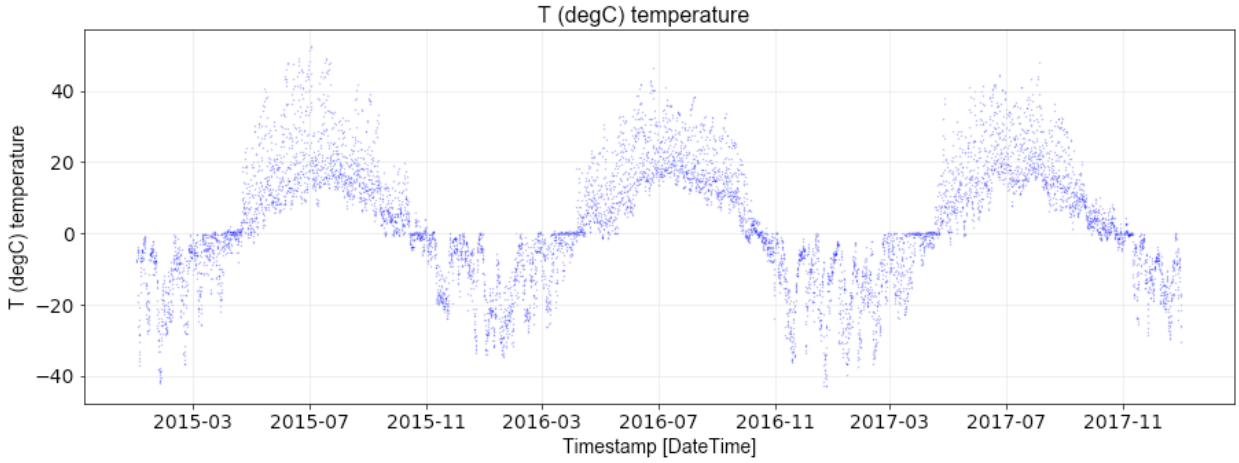


Рис. 1. Линейный график температуры воздуха по сухому термометру

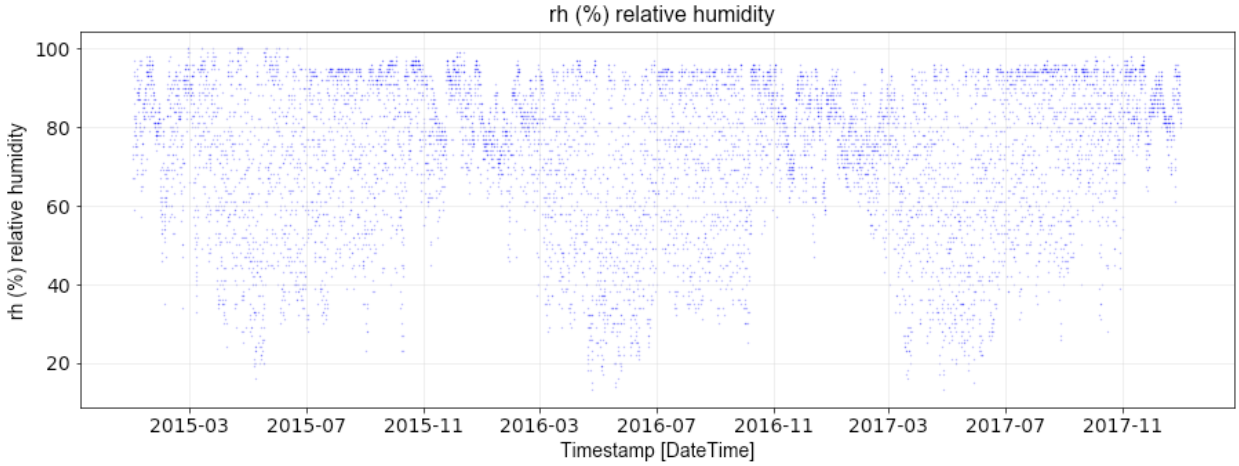


Рис. 2. Линейный график относительной влажности воздуха

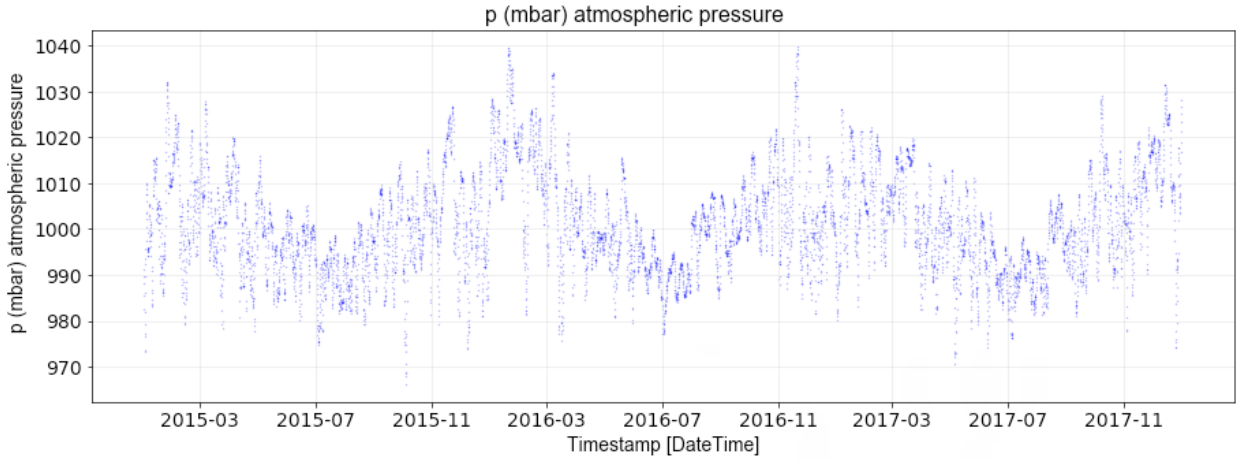


Рис. 3. Линейный график атмосферного давления на уровне станции

Для разработки моделей прогноза был использован язык программирования Python с использованием наиболее популярного фреймворка машинного обучения TensorFlow [22] с надстройкой Keras [23]. Для обработки и визуализации массивов входных данных и результатов эксперимента были использованы открытые библиотеки: Pandas, Numpy и Matplotlib [24].

Для реализации процесса обучения нейросети исходные данные были преобразованы – диапазоны значений параметров были нормированы от 0 до 1, аномальные выбросы параметров заменены на значения, похожие на фактические. Кроме того, на участках с отсутствующими данными они были искусственно добавлены. В совокупности такие данные не превышали 1 % всех данных эксперимента. В экспериментах каждый компонент набора данных состоял из 8 760 (8*365*3=8760) наблюдений. Обучающая выборка состояла из 7 884 (90 %) наблюдений и тестовой выборки из 876 (10 %) наблюдений по каждому компоненту.

Для процесса обучения нейросети были использованы семь указанных выше компонентов. Для построения прогнозных моделей были выбраны: температура воздуха по сухому термометру, относительная влажность воздуха, атмосферное давление на уровне станции. В процессе экспериментирования с гиперпараметрами (различными параметрами нейросети, необходимыми для улучшения качества обученной модели) были определены их оптимальные значения для достижения максимальной точности прогноза.

Ниже (рис. 4–9) представлены результаты обучения нейросети на обучающей выборке и результаты прогноза на тестовой выборке для 3-х, указанных выше исследуемых параметров. Зеленым цветом отображены фактические значения рядов наблюдений. Синим цветом отображены прогнозные значения рядов наблюдений. На всех графиках по оси X указаны трехчасовые моменты наблюдений. На обучающих выборках начальная точка отсчета соответствует 00 часов 01.01.2015 г. На тестовых выборках начальная точка отсчета соответствует 00 часов 12.09.2017.

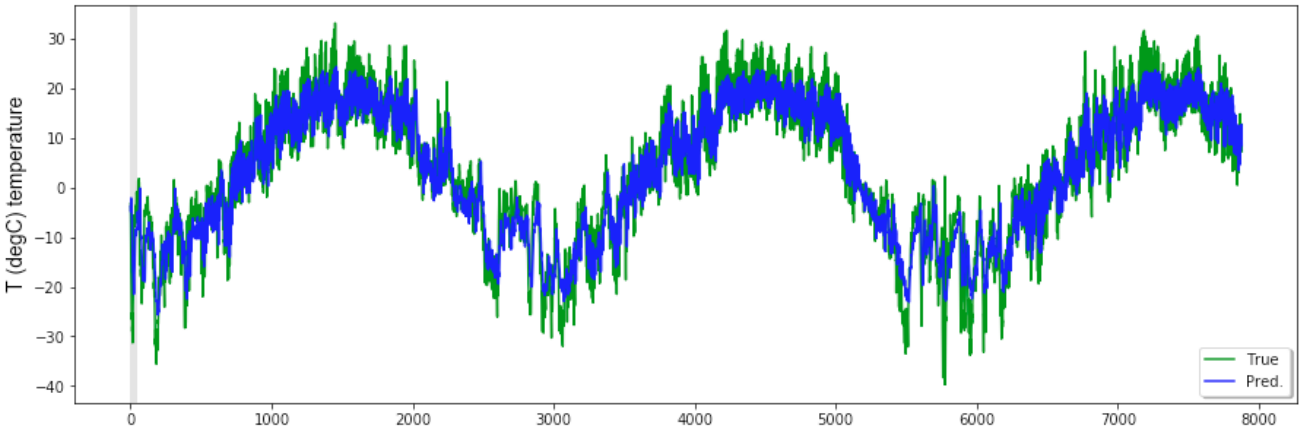


Рис. 4. Фактические и прогнозные значения температуры на обучающей выборке

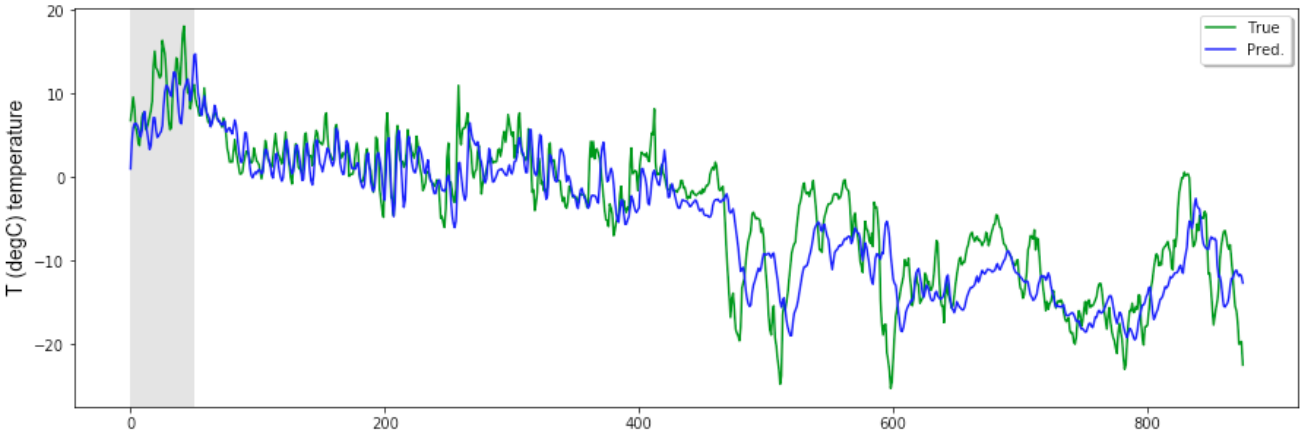


Рис. 5. Фактические и прогнозные значения температуры на тестовой выборке

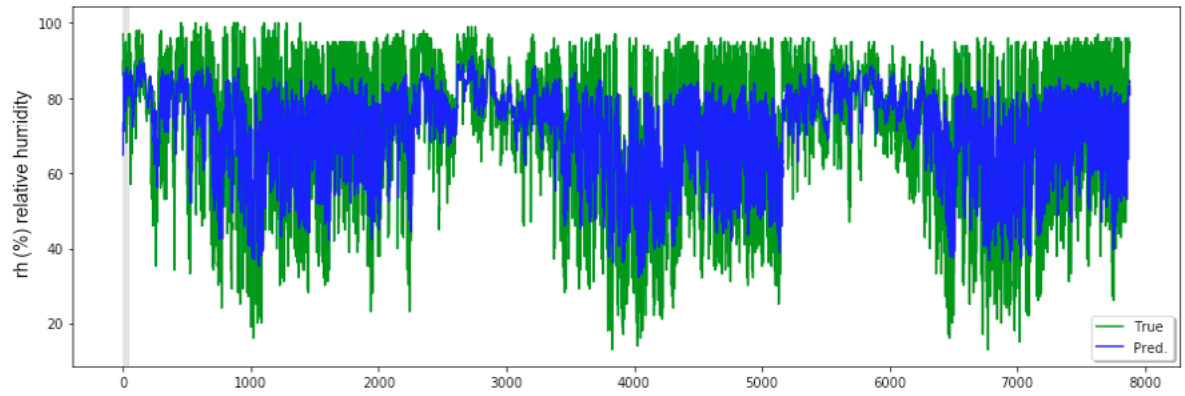


Рис. 6. Фактические и прогнозные значения относительной влажности воздуха на обучающей выборке

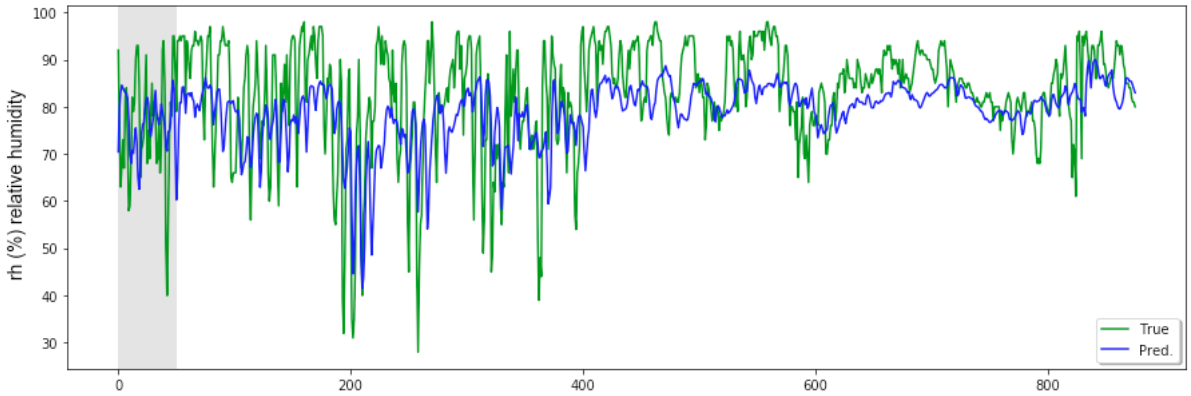


Рис. 7. Фактические и прогнозные значения относительной влажности воздуха на тестовой выборке

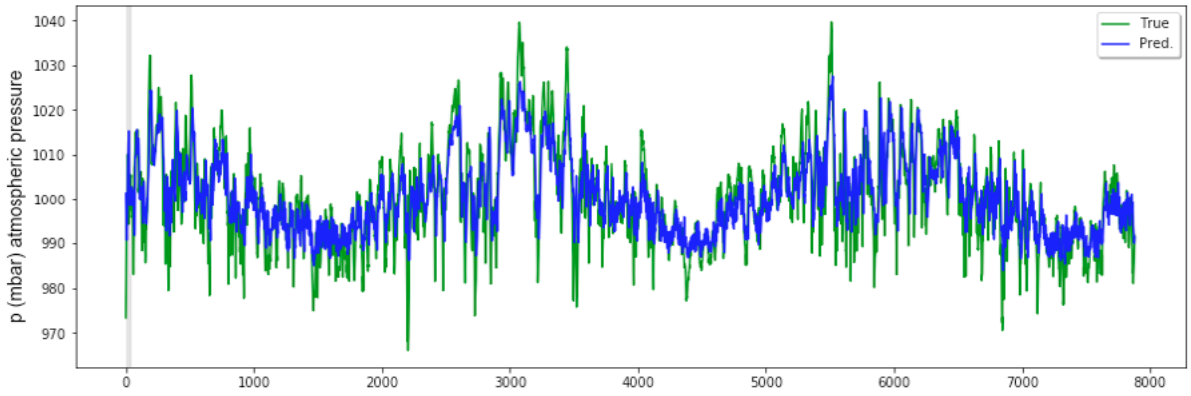


Рис. 8. Фактические и прогнозные значения атмосферного давления на обучающей выборке

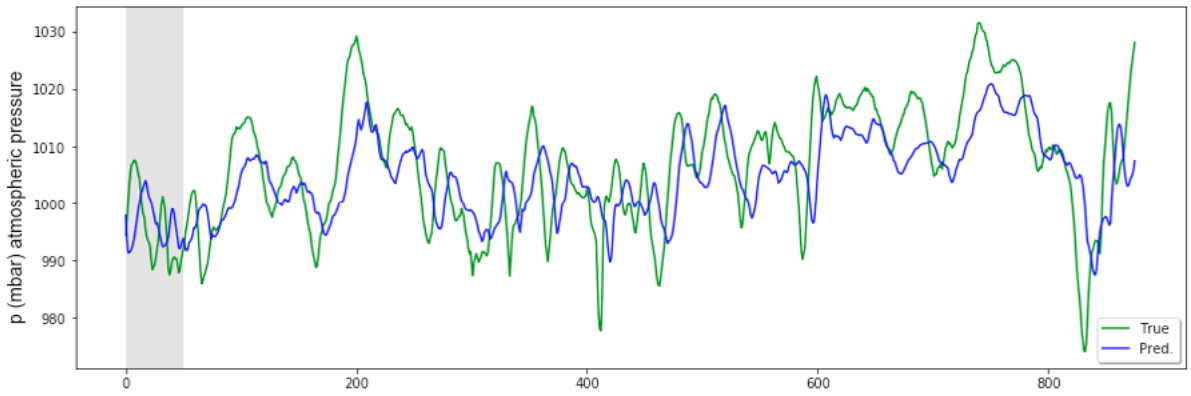


Рис. 9. Фактические и прогнозные значения атмосферного давления на обучающей выборке

Выводы

Модель достаточно хорошо отображала общие колебания параметров, но пиковые значения иногда были неточными. В некоторых случаях форма линейного графика была предсказана достаточно хорошо, хотя были небольшие отставания от линейного графика фактических данных. Эксперименты показали, что для повышения точности прогноза нужно больше входных компонентов (видов измеряемых значений) и интервал между наблюдениями был как можно меньше.

Благодарности

Исследование выполнено при частичной финансовой поддержке РФФИ и Администрации Томской области в рамках научного проекта № 18-47-700005р_а.

Список литературы

1. П. Флах, 2015. Машинное обучение. М.: ДМК Пресс, с: 400.

2. Х. Бринк, Д. Ричардс, М. Феверолф, 2017. Машинное обучение. Питер, с: 336.

3. Ричард С. Саттон, Эндрю Г. Барто, 2017. Обучение с подкреплением. Бином. Лаборатория знаний, с: 400.

4. D. Barber, 2010. Bayesian Reasoning and Machine Learning, pp: 610. Date Views 10.12.2017 http://web4.cs.ucl.ac.uk/staff/D.Barber/textbook/090310.pdf.

5. Martin T. Hagan, Howard B. Demuth, Mark H. Beale, Orlando De Jesús, 2014. Neural Network Design, pp: 1012. Date Views 10.12.2018 http://hagan.okstate.edu/NNDesign.pdf.

6. Н. Шакла, 2019. Машинное обучение и TensorFlow. Питер, с: 336.

7. H.S. Hippert, C.E. Pedreira and R.C. Souza, 2001. Neural networks for short-term load forecasting: a review and evaluation. IEEE Trans. Power Systems, 1(16): 44–55.

8. D.W. Bunn, 2000. Forecasting loads and prices in competitive power markets. Proc. IEEE, 88: 163–169

9. S. Lai, L. Xu, K. Liu and J. Zhao, 2015. Recurrent Convolutional Neural Networks forText Classification. AAAI, pp: 2267–2273.

10. C.N. dos Santos, M. Gatti, 2014. Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts. COLING, pp: 69–78.

11. V. Pratap, A. Hannun, Q. Xu, J. Cai, J. Kahn, G. Synnaeve, V. Liptchinsky and R. Collobert, 2018. AV2LETTER++: The fastest open-source speech recognition system. Date Views 10.12.2018 https://arxiv.org/pdf/1812.07625.pdf.

12. X. Chen, X.Liu, M.J.F. Gales and P.C. Woodland, 2015. Recurrent neural network language model training with noise contrastive estimation for speech recognition. In Acoustics, Speech and Signal Processing, pp: 5411-5415.

13. Who Leads the Self-Driving Cars Race? State-of-Affairs in Autonomous Driving. Date Views 30.01.2019 https://neurohive.io/en/state-of-the-art/self-driving-cars/

14. S. Mo, M. Cho and J. Shin, 2019. InstaGAN: Instance-aware image-to-image translation. Date Views 15.01.2019 https://arxiv.org/pdf/1812.10889.pdf.

15. J. Lu, M. Xu, R.Yang and Z. Wang, 2018. What Makes Natural Scene Memorable? Date Views 15.01.2019 https://arxiv.org/pdf/1808.08754.pdf.

16. A. Siarohin, S. Lathuili, S. Tulyakov, E. Ricci and N. Sebe, 2018. Animating Arbitrary Objects via Deep Motion Transfer. Date Views 15.01.2019 https://arxiv.org/pdf/1812.08861.pdf.

17. C. Wu, O. Wong, L. Rudnick, S. Shabala, M. Alger, J. Banfield, C.Ong, S. White, A. Garon, R. Norris, H. Andernach, J. Tate, V. Lukic, H. Tang, K. Schawinski and F. Diakogiannis, 2018. Radio Galaxy Zoo: CLARAN – a deep learning classifier for radio morphologies. Monthly Notices of the Royal Astronomical Society, 482: 211–230.

18. X. Yu, B. Fernando, B. Ghanem, F. Porikli and R. Hartley, 2018. Face Super-resolution Guided by Facial Component Heatmaps. In the Proceedings of the ECCV 2018: 15th European Conference, Munich, Germany, September 2018, pp. 219–235.

19. L. Jiang, A. Stocco, D. Losey, J. Abernethy, C. Prat and R. Rao, 2018. BrainNet: A Multi-Person Brain-to-Brain Interface for Direct Collaboration Between Brains. Date Views 15.01.2019 https://arxiv.org/pdf/1809.08632v1.pdf.

20. Y. Shi and A. Jain, 2018. DocFace+: ID Document to Selfie Matching. Date Views 15.01.2019 https://arxiv.org/pdf/1809.05620.pdf.

21. Fingernail Sensors and AI Can Help Clinicians to Monitor Health and Disease Progression. Date Views 15.01.2019 https://www.ibm.com/blogs/research/2018/12/fingernail-sensors/

22. TensorFlow: An open source machine learning framework for everyone. Date Views 15.01.2019 https://www.tensorflow.org/

23. Keras: The Python Deep Learning library. Date Views 15.01.2019 https://keras.io/

24. The Python Package Index (PyPI). Date Views 15.01.2019 https://pypi.org/

УДК 004.81

Ю.А. Егоров

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ КЛАССИЧЕСКИХ ПОДХОДОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ ПОЗ ЧЕЛОВЕКА С ИСПОЛЬЗОВАНИЕМ СКЕЛЕТНОЙ МОДЕЛИ

Аннотация

В работе представлено исследование эффективности применения классических подходов для решения задачи бинарной классификации позы людей на изображениях. Приведена постановка задачи классификации позы человека на изображении. Выдвинута гипотеза о том, что использование скелетной модели человека позволяет применять классические подходы к решению задачи классификации – подходы, не использующие глубокое обучение. Описан эксперимент и его результаты. Целью эксперимента является проверка выдвинутой гипотезы. Исходными данными являются необработанные изображения скелетных моделей людей, состоящих из опорных точек и соединяющих их линий. Данные были предварительно сгенерированы с использованием 3D моделей. В качестве алгоритмов и моделей классификации выбраны: SVM (метод опорных векторов), kNN (метод k ближайших соседей), многослойный персептрон. В результате проведенных экспериментов было установлено, что использование скелетной модели, представленной с помощью необработанных изображений, не позволяет решить поставленную задачу с помощью выбранных подходов с высокой точностью. Предложено альтернативное представление скелетной модели, позволяющее описать данную модель явно, и выдвинута гипотеза о применимости модели в новом представлении для решения задачи указанными методами.

Ключевые слова: классификации поз человека, метод опорных векторов, метод k ближайших соседей, многослойный персептрон.

Введение

Задача классификации поз человека (human pose classification problem) на изображении находит применение в таких сферах как обеспечение безопасности, здравоохранение, спорт, сфера развлечений.

Существующие решения поставленной задачи, как правило, в качестве исходных данных используют 3D изображения, полученные с помощью глубинных камер [1, 2]. Предложенные решения используют как подходы машинного обучения [3], так и классификаторы, использующие правила вывода для распознавания поз [4]. В [5] авторы предлагают использовать признаки Хаара как для 3D, так и для 2D изображений.

Также известны решения для 2D изображений: в [6] исследуется применение леса рандомизированных решающих деревьев; в [7] представлен метод классификации, основанный на каскаде рандомизированных заграждающих фильтров.

Большинство недавних работ посвящено решению другой задачи – human pose estimation – задачи по выделению на изображении областей, описывающих позу человека. Решения данной задачи, описанные в работах, разделяются на три группы:

1. использующие глубокое обучение [8, 9, 10, 11];
2. использующие иерархические модели и связи между частями тела людей [12, 13];
3. решения, основанные на выделении частей тела без выстраивания между ними иерархий [14].

Из описания позы человека строится скелетная модель, которую можно применить для решения задачи классификации поз человека.

В данной работе проверяется гипотеза о том, что использование скелетной модели человека позволит с высокой точностью решить задачу классификации поз человека на изображении с помощью классических подходов. Классическими подходами в работе называются алгоритмы и методы, не использующие глубокое обучение.

Скелетная модель

Скелетная модель человека представляет собой кортеж, $x_i = \langle x_i^1, \dots, x_i^j, \dots, x_i^m \rangle$, где x_i^j – соответствует опорной точке объекта. Опорные точки представляют собой суставы и три точки на лице, необходимые для отображения наклона и поворота головы. При графическом отображении скелетной модели, опорные точки соединяются линиями, как показано на рисунке 1.

Проверка гипотезы

Описание эксперимента

Был проведен эксперимент для проверки выдвинутой гипотезы. В эксперименте использовано графическое представление скелетных моделей. Каждое изображение получено с помощью ортогональной проекции 3D моделей (вид спереди). Сгенерированная выборка состоит из 600 изображений, разбитых на два класса: человек поднял руку (класс 1), человек принял другое положение (класс 2). Изображения были сгенерированы таким образом, чтобы граница между классами была нечеткой. 52 % изображений принадлежат классу 1, 48 % – классу 2. На рисунке 1 представлены примеры изображений.

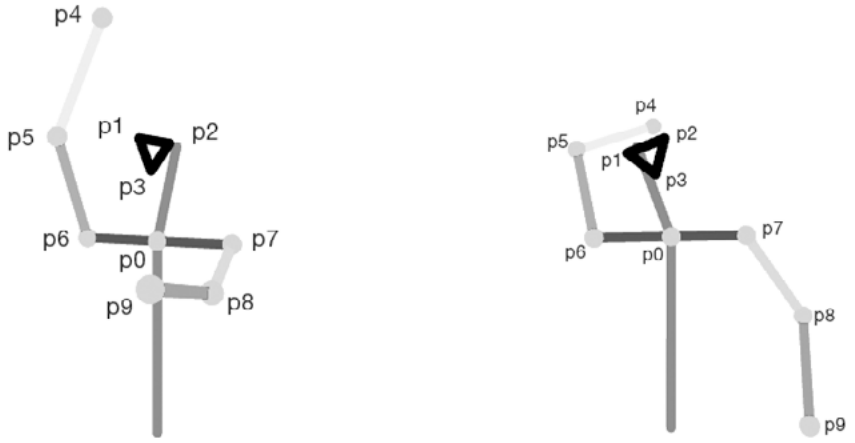


Рис. 1. Примеры исходных данных (слева класс 1, справа класс 2)

В эксперименте сравниваются три классификатора: kNN (k Nearest Neighbours, метод k ближайших соседей), SVM (Support Vector Machine, метод опорных векторов) и многослойный прецептрон. Производится подбор гиперпараметров для каждого классификатора и вычисление точности классификации. Оценка точности классификации производится с помощью F-меры и accuracy.

Для подбора оптимальных значений гиперпараметров использовался метод кросс-проверки с разделением выборки на 5 частей. Оптимальными гиперпараметрами классификатора считаются параметры, при которых в результате кросс-проверки было получено максимальное значение F-меры. В [15] представлен исходный код проекта по реализации эксперимента.

Метод k ближайших соседей

В эксперименте использовался модифицированный kNN [16, 17], использующий -мерное дерево для поиска ближайших соседей.

Производилась оптимизация следующий гиперпараметров:

1) метрика (расстояние между объектами): L1 (манхэттенское расстояние), L2 (евклидово расстояние) и D_{chess} (расстояние Чебышева);

2) количество соседей – k=1,15;

3) способ определения класса: выбор моды среди k ближайших соседей, взвешенный способ.

В ходе кросс-проверки были получены следующие оптимальные параметры для F-меры равной 0,540: расстояние Чебышева, k=13, способ определе-

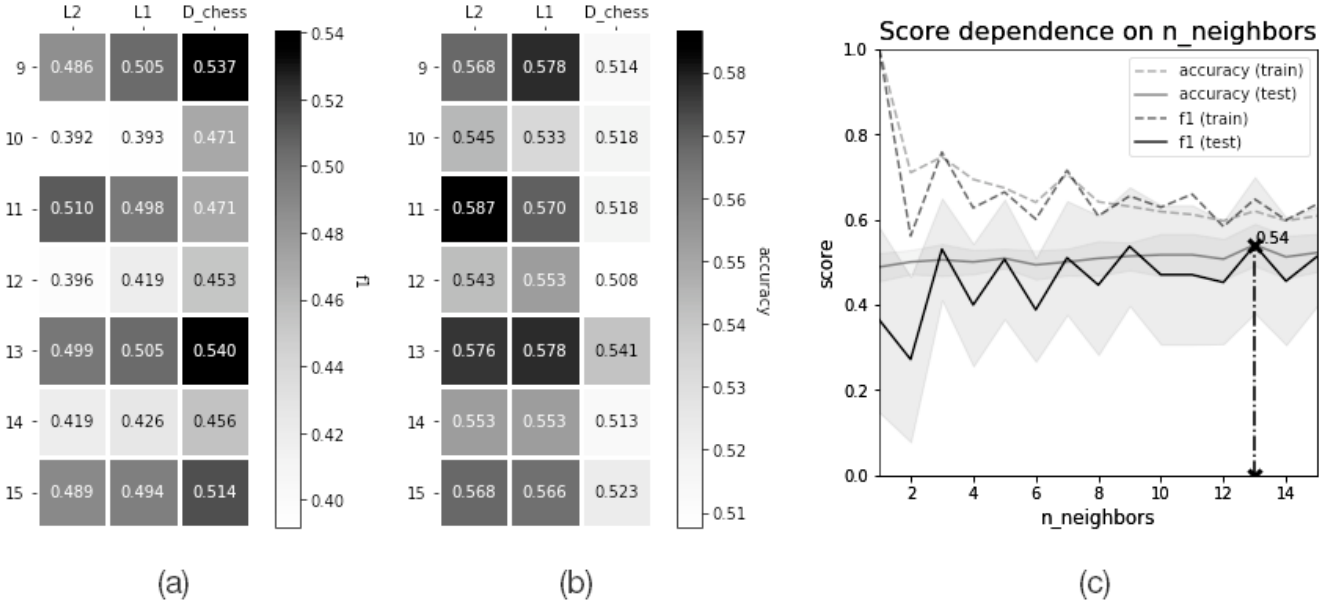


Рис. 2. a) Тепловая карта зависимости F-меры от выбора метрики (по горизонтали) и k (по вертикали);
b) Тепловая карта зависимости accuracy от выбора метрики (по горизонтали) и k (по вертикали);
c) График зависимости F-меры и accuracy от количества соседей для расстояния Чебышева

ния класса – выбор моды среди *k* ближайших соседей. Для accuracy равной 0.587 были получены параметры: евклидово расстояние, *k=11*, способ определения класса – выбора моды среди *k* соседей.

На рисунке 2 представлены графики, иллюстрирующие результаты проведенного эксперимента.

Метод опорных векторов

Для SVM классификатора [18] производилась оптимизация следующих гиперпараметров:

1) ядро классификатора: линейное, RBF (radial basis function, радиально-базисная функция), сигмоидное, полиномиальное;

2) штрафной параметр $C \in (0,1] \in R$;

3) коэффициент ядра для RBF, сигмоидного и полиномиального ядер [gamma] $\in (0,1] \in R$;

4) степень полинома для полиномиального ядра $p \in [2,5] \in N$.

В ходе кросс-проверки были получены следующие результаты. Оптимальное ядро для классификатора – полином степени $p = 2$. Коэффициент ядра [gamma]= $4{,}64 \cdot 10^{-5}$. Штрафной параметр *c* не оказал влияния на точность работы классификатора. При этом были получены значения F-меры равное 0,66 и accuracy равное 0,68.

На рисунке 3 представлены тепловые карты, иллюстрирующие результаты проведенного эксперимента.

Многослойный персептрон

Многослойный персептрон [19, 20] обучался со следующими параметрами:

1) функция активации – логистическая сигмоида;

2) алгоритм оптимизации – Adam [21];

3) штрафной параметр [alpha] равный 0,001;

4) максимальное количество эпох 1 000.

Персептрон имеет 62 208 входов (для изображения 144 на 144 px с тремя цветовыми каналами). Скрытые слои полносвязные.

В ходе кросс проверки была найдена оптимальная структура скрытых слоев персептрона, состоящая из двух слоев по 144 элемента в каждом слое. Для персептрона данной структуры значение F-меры составило 0,77, значение меры accuracy составило 0,8.

Выводы

Проведенные эксперименты показали, что использование классических подходов к решению задачи классификации с помощью графических изображений не дает точных результатов (accuracy равная 0,8 и F-мера равная 0,77 для многослойного персептрона). Но была выдвинута новая гипотеза о том, что повысить точность классификации можно за счет изменения представления скелетной модели.

Предлагается вместо графического представления скелетной модели использовать координаты опорных точек в декартовых координатах с центром в точке (рис. 1). Новая сформулированная гипотеза также требует проверки.

Список литературы

1. Zhu Y. and K. Fujimura, 2007. Constrained Optimization for Human Pose Estimation from Depth Sequences. Asian Conference on Computer Vision ACCV 2007, Tokyo, Japan, November 2007, pp. 408–418.

2. Jansen B., F. Temmermans and R. Deklerck, 2007. 3D human pose recognition for home monitoring of

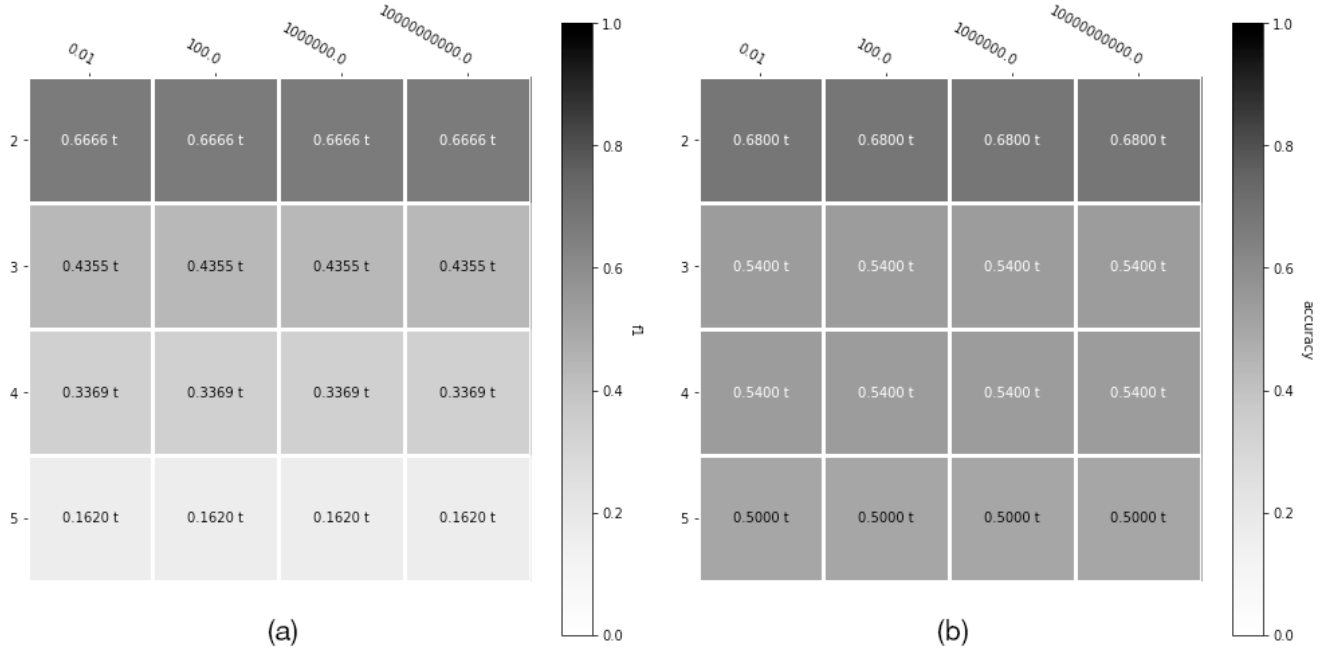


Рис. 3. a) Тепловая карта, иллюстрирующая зависимость F-меры от значения штрафного коэффициента c (по горизонтали) и степени полинома p (по вертикали); b) Тепловая карта, иллюстрирующая зависимость меры accuracy от значения штрафного коэффициента c (по горизонтали) и степени полинома p (по вертикали)

elderly. 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, February 2007, pp. 4049–4051.

3. Shotton J., A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman and A. Blake, 2011. Real-time human pose recognition in parts from single depth images. CVPR 2011, Colorado Springs, USA, June 2011, pp. 1297–1304.

4. Zainordin F.D., H.Y. Lee, N.A. Sani, Y.M. Wong and Ch. S. Chan, 2012. Human pose recognition using Kinect and rule-based system. World Automation Congress 2012, Puerto Vallarta, Mexico, June 2012, pp. 1–6.

5. Van den Bergh, M, E. Koller-Meier and L. Van Gool, 2009. Real-Time Body Pose Recognition Using 2D or 3D Haarlets. International Journal of Computer Vision, 83(1): 72–84.

6. Rogez G., J. Rihan, S. Ramalingam, C. Orrite and Ph.H.S. Torr, 2008. Randomized trees for human pose detection. 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, USA, June 2008, pp. 1–8.

7. Rogez G., J. Rihan, C. Orrite-Uruñuela and Ph.H.S. Torr, 2012. Fast Human Pose Detection Using Randomized Hierarchical Cascades of Rejectors. International Journal of Computer Vision, 99(1): 25–52.

8. Toshev A. and Ch. Szegedy, 2014. DeepPose: Human Pose Estimation via Deep Neural Networks. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2014), Columbus, USA, August 2014, pp. pp. 1653-1660.

9. Wei Sh.-En, V. Ramakrishna, T. Kanade and Y. Sheikh, 2016. Convolutional Pose Machines. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, USA, June 2016, pp, 4724–4732.

10. Cao Zh., T. Simon, Sh.-En Wei and Y. Sheikh, 2017. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, USA, July 2017, pp. 7291–7299.

11. Newell A., K. Yang and J. Deng, 2016. Stacked Hourglass Networks for Human Pose Estimation. European Conference on Computer Vision (ECCV 2016), Amsterdam, The Netherlands, October 2016, pp. 483–499.

12. Tian Y., C. Lawrence, Z. Srinivasa and G. Narasimhan, 2012. Exploring the Spatial Hierarchy of Mixture Models for Human Pose Estimation. European Conference on Computer Vision (ECCV 2012), Florence, Italy, August 2012, pp. 256–269.

13. Sun M. and S. Savarese, 2011. Articulated part-based model for joint object detection and pose estimation. 2011 IEEE International Conference on Computer Vision (ICCV 2011) (ICCV), Barcelona Spain, November 2011, pp. 723–730.

14. Felzenszwalb P.F. and D.P. Huttenlocher, 2005. Pictorial Structures for Object Recognition. International Journal of Computer Vision, 61(1): 55–79.

15. The project source code providing the results of the reserash. Date Views 01.02.2019 https://github.com/yordwynn/elementary-pose.

16. Sproull R.F., 1991. Refinements to nearest-neighbor searching in k-dimensional trees. Algoritmica, 1(6): 579–589.

17. Ramasubramanian V. and K.K. Paliwal, 1992. Fast K-dimensional tree algorithms for nearest neighbor search with application to vector quantization encoding. IEEE Transactions on Signal Processing, 40(3): 518–531.

18. Boser B.E., I.M. Guyon, V.N. Vapnik, 1992. A training algorithm for optimal margin classifiers. COLT ‘92 Proceedings of the fifth annual workshop on Computational learning theory, Pittsburgh, USA, July 1992, pp. 144-152.

19. Ruck D.W., S.K. Rogers, M. Kabrisky, M.E. Oxley and B.W. Suter, 1990. The multilayer perceptron as an approximation to a Bayes optimal discriminant function. IEEE Transactions on Neural Networks, 1(4): 296–298.

20. Tang J., C. Deng and G-B. Huang, 2016. Extreme Learning Machine for Multilayer Perceptron. IEEE Transactions on Neural Networks and Learning Systems, 27 (4): 809–821.

21. Kingma D.P., J. Ba, 2015. Adam: A Method for Stochastic Optimization. 3rd International Conference for Learning Representations, San Diego, USA, May 2015, pp. 1–13.

УДК 519.766.4

А.С. Еримбетова, Ф.А. Мурзин

РАЗРАБОТКА И АНАЛИЗ ТЕХНОЛОГИЙ ОПРЕДЕЛЕНИЯ РЕЛЕВАНТНОСТИ ТЕКСТА ПОИСКОВОМУ ЗАПРОСУ

Аннотация

В этой статье мы представляем разработку новой лингвистической и алгоритмической поддержки информационно-поисковых технологий. Методы, рассмотренные в этой статье, используются для ранжирования результатов выдачи из других систем, таких как хорошо известные поисковые системы Google, Yandex и т. д. В процессе решения задач были рассмотрены способы представления семантико-синтаксических отношений между словами в предложениях на основе диаграмм, полученных в результате работы программной системы Link Grammar Parser. Также рассматриваются методы вычисления степени совпадения конструкций на естественном языке. Изучаются вопросы фильтрации данных. Эта проблема может быть решена с помощью лексической, морфологической, синтаксической и семантической информации. Основная идея заключается в разработке технологий, позволяющих информационно-поисковым системам сравнивать полученные данные с вариантами пользовательских запросов, основываясь на анализе их синтаксических структур. Предложено обобщение алгоритма идентификации текстовой темы Н. Кумара, К. Сринатана и В. Вармы с помощью коннекторов системы Link Grammar Parser. В статье описаны реализация программы и результаты тестирования. Подобные программные продукты могут применяться в различных организациях или отдельными пользователями, осуществляющими регулярный поиск в сети различного рода информации.

Ключевые слова: обработка естественного языка, Link Grammar Parser, релевантность, сниппет, информационно-поисковая система.

Введение

Количество информации в сети интернет стремительно растёт с каждым днем. В связи с этим возникает проблема поиска нужной информации. Основная цель данной статьи – разработка методов, позволяющих сравнивать конструкции естественного языка и выявлять перефразированные варианты предложений на основе анализа его синтаксической структуры.

Существует достаточное количество поисковых систем и методов ранжирования страниц, но большинство из них являются закрытыми и представляют собой коммерческую тайну. Также коммерческие системы поиска заинтересованы в сохранении пользователей на своих страницах для показа рекламы, в связи с чем им это выгодно, когда пользователь получает нужную информацию не сразу, а после нескольких запросов и посещённых веб-страниц.

Представленные в этой работе методы используются для ранжирования результатов выдачи других поисковых систем, таких как Google, Yandex, системы полнотекстового поиска ElasticSearch, Solr, Sphinx и т. д. Описанные методы не подходят для массового полнотекстового поиска из-за вычислительных затрат, однако он хорошо работает как инструмент ранжирования и исключения не релевантных результатов выдачи.

В настоящее время созданы методы лингвистического анализа текстов и разработаны программные системы, позволяющие автоматически выполнять морфологический, синтаксический и семантический анализ [1–4] предложений текста: AOT, Solarix, NLTK, FreeLing, и т.п. Это дает возможность улучшения качества (полноты и точности) и вычислительной эффективности решения поисково-аналитических задач, посредством разработки и применения интеллектуальных методов оценки результатов работы известных поисковых систем [19, 20]. То есть речь идет о фильтрации полученных от них данных. Таким образом, можно принять во внимание [6–9] лексическую, морфологическую, синтаксическую и семантическую информацию. В данной работе нас интересует вопрос фильтрации выходных данных поисковой системы.

Алгоритм Нираджа Кумара основан на нескольких идеях [14]: текст может состоять из нескольких предложений и представлен в виде некоторого графа; учитывается порядок слов; алгоритм косвенно учитывает статистику; с помощью понятия центральности учитывается совместимость слов.

Мы стараемся и дальше продвигать этот алгоритм с учетом синтаксической структуры. Использовалась нестандартная теория синтаксиса, на которой основана программная система Link Grammar Parser, и эффективность которой в приложениях подтверждена экспериментами.

Основная идея алгоритма резюмирования

Основой для этой статьи послужило исследование [14], в котором авторы рассматривают процесс автоматического резюмирования. Основная идея состоит в том, что из текста могут быть выделены фрагменты, соответствующие данной теме. Фрагменты не обязательно следуют один за другим. В текст между ними могут быть вставлены фрагменты на другие темы. Далее выделенные фрагменты можно объединить в резюме по данной теме. Как правило, из текста можно выбрать несколько таких резюме, соответствующих различным темам. Одна из возникающих проблем состоит в том, что перестановка слов в предложении

может существенно изменить его смысл, что приводит к некорректной работе лингвистических алгоритмов, работающих с отдельными ключевыми словами, их частотами и т. д.

Алгоритм состоит из нескольких этапов: предварительная обработка текста статьи; расчет весов слов; разделить статью на темы (рис. 1); расчет весов тем; вычисление релевантности заранее заданных образцов и тем; расчет итоговой оценки.

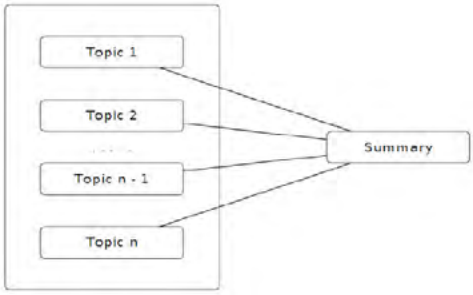


Рис. 1. Общая схема алгоритма

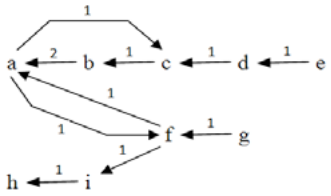
Краткое описание основного алгоритма идентификации

Для решения поставленных задач предлагается использовать метод представления семантико-синтаксических отношений между семантическими единицами предложения на основе диаграмм программной системы Link Grammar Parser [15, 16]. Диаграммы, полученные с помощью анализатора Link Grammar Parser, представлены в виде графов. Затем следует их предварительная подготовка к сравнению и само сравнение. Графы будут сравниваться не только на лексическом уровне, но и на уровне связей между словами. Особенно интересуют алгоритмы, основанные на тестировании ряда логических свойств.

Предположим, что мы работаем с деревьями, полученными в результате синтаксического анализа, взятого из системы Link Grammar Parser. Затем следует «обобщение» таких деревьев. На этом этапе производилась нормализация словоформ. Возможно, что произойдет преобразование нескольких других предложений. Например, рекурсивный порядок слов меняется на прямой. Сложные формы глагола «обрезаны» до простых форм. В результате мы получаем каркас дерева, в котором удаляются различные речеобразующие конструкции. Такие деревья сравнивается друг с другом, то есть проводится определение релевантности текста по запросам. Алгоритм сравнения выражений напоминает работу конечных автоматов, работающих на вершинах дерева [5].

Расчет веса слов

В общем случае мы предполагаем, что имеется некоторый фрагмент анализируемого текста или заранее подготовленный «тестовый» фрагмент, то есть последовательность предложений на естественном языке $S=<S_1,S_2,...,S_n>$. Кроме того, некоторый граф связан с последовательностями S. Метод можно объяснить с помощью примера, рассмотренного в [14]. Ниже показана последовательность предложений и соответствующий граф. Предложение: S_1, S_2, S_3. Слова: *a b c d e; c a f g; h i f a b*



Здесь *a, b, c, d, e* – слова соответствующих предложений. Множество всех слов, встречающихся в данной совокупности предложений, образует множество вершин графа. Ребро графа отображает факт непосредственного следования одного слова за другим. При этом направление ориентации ребра задается от последующего слова к предыдущему. Вес ребра – это количество вхождений данной пары слов, следующих друг за другом в данном порядке, в весь текстовый фрагмент $S=<S_1,S_2,...,S_n>$.

В работе [14] авторы предполагают, что проводится некоторая преобработка текста. А именно, удаляются разного рода служебные слова: артикли, предлоги, послелоги, частицы, междометия. Более того, текст пропускается через стеммер, и в итоге мы имеем только корни слов.

Нами предлагается некоторое обобщение алгоритма с учетом синтаксической структуры предложений, полученной на выходе системы Link Grammar Parser, в том числе на основе размытой логики Заде [18]. При нашем подходе целесообразно не проводить такого рода предобработку, в том числе не использовать стеммер. Важно исправить, сколько раз встречается связи (коннектор). Надо попытаться учесть количество входящих в вершину связей и количество выходящих из вершины связей. Для этого вводится понятие ранга вершины (веса каждого слова).

Веса слов могут быть рассчитаны различными способами, например:

$$S(V_i) = \frac{1-\lambda}{N} + \lambda \sum_{V_j \in IN(V_i)} \frac{S(V_j)}{|OUT(V_j)|}, \qquad (1)$$

где S (V_i) – ранг V_i; S (V_j) – ранг V_j, с которого ссылка направляется в узел V_i; $|OUT(V_j)|$ число узлов, соединенных ребрами, исходящими из V_j; N – количество узлов в графе; λ – коэффициент затухания (damping factor), он равен 0,85.

Приведенная выше формула означает, что вершина, на которую идет ссылка из других вершин с высоким рангом, сама получает высокий ранг. Таким образом, мы можем выделить в целом тексте наиболее важные, действительно ключевые, слова.

Модель оценки релевантности

Давайте обобщим этот алгоритм и попытаемся учесть синтаксические структуры предложений. Мы будем использовать систему Link Grammar Parser перед вычислением весов тем. Для данного предложения программная система назначает свою синтаксическую структуру в виде набора отмеченных ссылок (рис. 2) между парами слов.

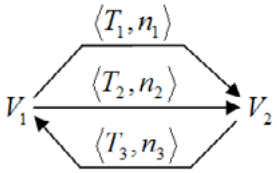
```
+----→WV----→+-----Os-----+
+---Wd--+---Ss---+       +Ds**c+
|       |        |       |      |
LEFT-WALL Tom.m opened.v-d the door.n
```

Рис. 2. Пример разбора предложения

Таким образом, мы присваиваем граф $G_i(V_i,E_i)$, созданный синтаксическим анализатором грамматики ссылок, на основе предложения S_i. В этом графе V_i – это набор слов, а E_i – набор триплетов $< v_1,v_2,t >$, где $v_1,v_2 \epsilon V_i$ – вершины, а t – тип связи. Таким образом, мы получаем $G_1,\ldots,G_k$ – графы предложений. На следующем шаге мы построим граф $G(V,E)$, объединив графы предложений $G_1,\ldots,G_k$.

Здесь $V=V_{1\le i\le k}V_i$ – множество всех слов из предложений; E – это набор из четырех частей $< v_1,v_2,t,n >$, где, как и ранее, $v_1,v_2 \epsilon V_i$ – это вершины, t – тип связи и дополнительный параметр $n=|\{i:<v_1,v_2,t > \epsilon E_i\}|$ номер вхождения триплета $< v_1,v_2,t >$.

Пример двух верхних соединений показан ниже.



Предполагам, что каждой связи T_i сопоставлено положительное число α_i, называемое ее весом или важностью. Далее полагаем: 1) $(T_i,1) = \alpha_i$, 2) $\mu(T_i,n)=n\cdot\mu(T_i,1)=n\cdot\alpha_i$, 3) $\mu(T_{i1},n_1,K,T_{ik},n_k)=\sum_j(T_{ij},n_j) =\sum_j n_j\cdot\alpha_j$.

Таким образом, если из вершины V_i в вершину V_j, идет множество параллельных ребер с пометками $<T_{i1},n_1,K,T_{ik},n_k>$, то все они могут быть заменены одним ребром, имеющим вес равный $w_{ij}=\mu(T_{i1},n_1,K,T_{ik},n_k)$.

Соответствующим образом может быть изменена формула вычисления ранга слова. При этом можно использовать два варианта формулы.

$$S(V_i) = \frac{1-\lambda}{N} + \lambda \sum_{V_j \in IN(V_i)} \frac{w_{ji} \cdot S(V_j)}{|OUT(V_j)|} \text{ и} \quad (2)$$

$$S(V_i) = \frac{1-\lambda}{N} + \lambda \sum_{V_j \in IN(V_i)} \frac{w_{ji} \cdot S(V_j)}{\sum_{V_k \epsilon OUT(V_j)} w_{jk}} \quad (3)$$

Определим LinkLength_{ij}, который будет использоваться в качестве веса ребер в конечном графе:

$$PathLengh_{ij} = \frac{1}{LinkLength_{ij}}. \quad (4)$$

Для каждой вершины мы вычисляем центральность близости. Смысл этого понятия – обратная величина среднего геодезического расстояния до других вершин:

$$C_c(v_i) = \frac{N-1}{\sum_{v_j \epsilon V, v_j \neq v_i} d_G(v_i,v_j)} \quad (5)$$

Мы вычисляем относительное изменение центральности близости для каждого узла:

$$Diff(v_i) = \frac{|C_c^1(v_i) - C_c^2(v_i)|}{C_c^1(v_i)}. \quad (6)$$

Далее рассчитаем параметр:

$Diff(V_i)=(|C_c(V_i)_2 - C_c(V_i)_1| / (C_c(V_i)_1)\times 100.)$ (7)

Он характеризует, насколько одинаково или по-разному используется данное слово в контексте первого и второго фрагментов текста. После этого мы устанавливаем порог, который служит критерием «сходства» употребления слова. Например, в качестве порога можно взять медиану этих размеров $Diff_{avg}f(V_i)$. В дальнейшем этот порог можно использовать при анализе других фрагментов.

Наконец, определяем предикат, указывающий, содержатся ли вершины аналогично в графиках для анализируемой темы и в стандарте для выбранной темы:

$\text{Sim}(v_i)=(Diff_{avg} < 0.5 \wedge Diff(v_i) < 0.5) \vee$
$(Diff_{avg} > 0.5 \wedge Diff(v_i) < Diff_{avg})$ (8)

В качестве оценки соответствия темы стандарту, мы берем частоту встречающихся слов, для которых предикат верен, ко всем словам:

$$Score(T) = \frac{|\{v \in V: Sim(v)\}|}{|Set(T)|}. \quad (9)$$

Окончательная оценка взвешенной суммы оценок по темам имеет вид: $Score = \sum_T W(T)Score(T)$, где могут использоваться дополнительные веса .

Таким образом, становится возможным определить, соответствует ли сообщение или последовательность сообщений данной статье. Может быть указан дополнительный коэффициент, который помогает оценить отношение пользователя к фильму, игре или новостям, описанным в статье. В этом случае эксперты предлагают значения весов. Если вычисленная оценка релевантности удовлетворяет заданному критерию, то она дает возможность получить более полную информацию об отношении пользователя к материалу для чтения.

Реализация программы

Программный продукт (рис. 3) должен обеспечивать: получение от пользователя поискового запроса и его предварительную обработку; запрос и получение от поисковой системы результатов поиска по поисковому запросу пользователя; анализ результатов поисковой системы; вывод результатов анализа в таблицу и сортировка их по убыванию коэффициента релевантности.

Алгоритм написан на C++ с использованием Link Grammar Parser и сторонних стеммеров. Интерфейс реализован на языке со статической типизацией TypeScript с использованием платформы для разработ-

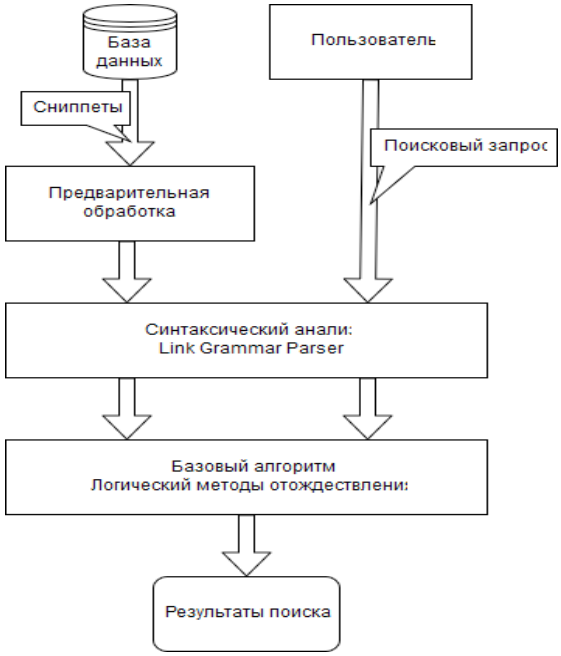


Рис. 3. Схема работы программной системы

ки веб-приложений Angular 5, библиотеки реактивного программирования Rx.js, библиотеки компонентов Material Angular, фреймворка для построения нативных графических приложений для настольных операционных систем с помощью веб-технологий Electron. Интерфейс после получения запроса от пользователя, запрашивает сниппеты из Google. После последовательно передаёт сниппеты алгоритму. После того, как получены данные о релевантности, интерфейс отображает их в виде таблицы в отдельной вкладке (рис. 4).

Результаты тестирования системы

В процессе тестирования система запрашивала у поисковой системы Google список сниппетов вместе с URL. Эксперт проанализировал фрагменты и результаты работы системы. Результаты тестирования разработанной системы представлены ниже в таблице: n_0 – общее количество фрагментов, полученных от Google; n_1 – количество релевантных ссылок, одобренных системой; n_2 – количество релевантных ссылок, пропущенных системой; n_3 – количество не релевантных ссылок, одобренных системой.

В таблице 1 приведены запросы на русском языке. По каждому запросу загружались списки адресов с их описанием, которые поисковики обычно выдают пользователю. По этим коротким описаниям (сниппет) производилась оценка ресурса. Для сравнения с поисковой системой была составлена статистика. Система оставляла релевантные ссылки, отбрасывая нерелевантные по ее мнению. В итоге выяснилось, что на проведенных тестах в среднем из 100 ссылок, полученных из поискового сервиса Google, система выделяла 5–15 качественных релевантных ссылок, до 5 ссылок система ошибочно принимала за релевантные и остальные отбрасывала как нерелевантные, что соответствовало действительности. Это показывает, что данная система смогла произвести фильтрацию на хорошем уровне. Мы реализовали этот алгоритм сами. Тестирование показало, что наш алгоритм, учитывающий Link Grammar Parser, работает намного лучше.

Выводы

В процессе решения задач были рассмотрены способы представления семантико-синтаксических отношений между словами в предложениях на основе диаграмм, полученных в результате работы программной системы Link Grammar Parser. Также рассматриваются методы вычисления степени совпадения конструкций на естественном языке. В результате предложен алгоритм фильтрации результатов выдачи поисковых систем Google, Yandex и т. д. Было проведено много тестирований. Система проанализировала фрагменты, полученные вместе с URL из поисковой системы Google. Эксперт подтвердил высокое качество обработки информации.

Благодарности

Работа выполнена в рамках грантового финансирования по научным и (или) научно-техническим проектам на 2018–2020 годы МОН РК № AP05133550 и проект РФФИ № 18-07-01457-а.

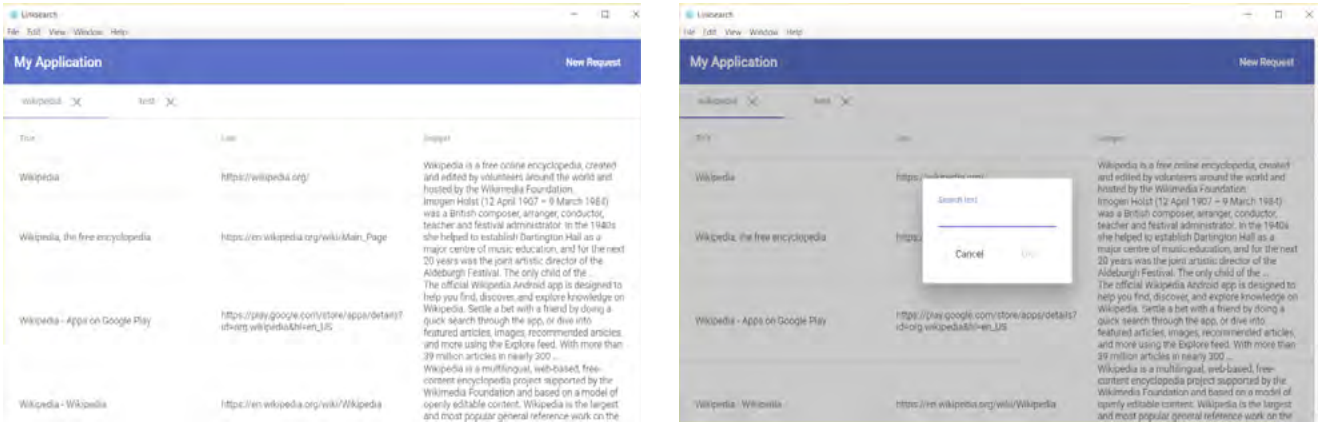


Рис. 4. Диалоговое окно поискового запроса и диалоговое окно нового запроса

Таблица 1

Результаты тестирования системы

п/п	Текст тестового запроса	n_0	n_1	n_2	n_3
1	Действие пряностей на организм человека	290	20	0	2
2	Ароматические вещества гвоздики и мускатного ореха	230	20	0	0
3	Природные фенолы, ванилин, воздействие на человека	176	19	1	4
4	Глюкоза, фруктоза, сахароза, лактоза и их влияние на организм человека	170	20	0	0
5	Аллергические реакции на пластмассы и полимерные ткани	210	19	1	0
6	Действие аскорбиновой кислоты на организм человека	290	19	1	0
7	Аллергические реакции на формальдегидные смолы	140	19	1	1
8	Аллергические реакции на красители	310	20	0	0
9	Пенициллины и другие антибиотики, привыкание к ним	100	17	1	0
10	Морфин, никотин, кофеин и их воздействие на человека	220	20	0	1

Список литературы

1. Автоматическая обработка текста (АОТ). Дата обращения: 25.01.2019 http://www.aot.ru

2. Компьютерная грамматика русского языка: лексика, морфология, синтаксис. Дата обращения 20.01.2019 http://www.solarix.ru.

3. Natural Language Toolkit. Date Views 25.12.2018 http://nltk.org.

4. Bird S., 2009. Natural Language Processing with Python. O'Reilly Media Inc. pp. 504.

5. Freeling: An Open Source Suite Of Language Analyzers. Date Views 28.05.2018 http://nlp.lsi.upc.edu/freeling.

6. Zobel J. and A. Moffat, 2006. Inverted files for text search engines. ACM Computing Surveys (CSUR), vol. 38, #2, Article 6. doi:10.1145/1132956.1132959.

7. Соченков И.В., 2013. Метод сравнения текстов для решения поисково-аналитических задач. Искусственный интеллект и принятие решений. М.: Федеральный исследовательский центр «Информатика и управление» Российской академии наук, #2, С. 32–43.

8. Dan S., and M. Lapata, 2007. Using Semantic Roles to Improve Question Answering. Proc. of the 2007 Joint Conf. on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL). Date Views 05.05.2018 http://www.aclweb.org/anthology/D/D07/D07-1002.pdf

9. Osipov G., I. Smirnov, I. Tikhomirov and O. Zavjalova, 2008. Application of linguistic knowledge to search precision improvement. Proc. of 4th Intern. IEEE Conf. on Intelligent Systems, vol. 2, pp. 17-2–17-5.

10. Verberne S., L. Boves, N. Oostdijk and P. Coppen, 2008. Using syntactic information for improving why-question answering. Proc. of the 22nd Intern. Conf. on Computational Linguistics (COLING '08), vol. 1, Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 953–960.

11. Чатуев М.Б. и А.М. Чеповский, 2011. Частотные методы в компьютерной лингвистике: учеб. пособие – М. МГУП.

12. Turney P.D. and P. Pantel, 2010. From frequency to meaning: Vector space models of semantics. Journal of artificial intelligence research, #37. pp. 141-188. Date Views 10.05.2018 http://www.aaai.org/Papers/JAIR/Vol37/JAIR-3705.pdf.

13. Sebastian F., 2012. Machine Learning in Automated Text Categorization. ACM Computing Surveys. V. 34. #1. pp. 1–47.

14. Kumar N., K. Srinathan and V. Varma, 2012. Using graph based mapping of co-occurring words and closeness centrality score for summarization evaluation. CICLing Proc. of the 13th Intern. Conf. on Computational Linguistics and Intelligent Text Processing, pp. 353-365.

15. Temperley D., D. Sleator and J. Lafferty, 1998. Link Grammar Documentation. Date Views 10.12.2018 http://www.link.cs.cmu.edu/link/dict/index.html.

16. Sleator D. and D. Temperley, 1991. Parsing English with a Link Grammar. Pittsburgh: School of Computer Science Carnegie Mellon University. p. 93.

17. Батура Т.В. и Ф.А. Мурзин, 2008. Машинно-ориентированные логические методы отображения семантики текста на естественном языке: моногр. Институт систем информатики им. А.П. Ершова СО РАН. Новосибирск: Изд. НГТУ, 248 с.

18. Zadeh L.A., 1975. The concept of a linguistic variable and its application to approximate resoning. Information Sciences, vol. 8, pp. 199–249.

19. Peters C., M. Braschler and P. Clough, 2012. Multilingual Information Retrieval: From Research To Practice. Berlin-Heidelberg: Springer-Verlag, p. 2017

20. Orasan C., V. Pekar and L. Hasler, 2004. A comparison of summarization methods based on term specificity estimation. Proc. of the Fourth International Conference on Language Resources and Evaluation (LREC-04). Lisbon, Portugal. pp. 1037–1041.

УДК 004.822

Ю.А. Загорулько, О.И. Боровикова

ПРОБЛЕМЫ ПОСТРОЕНИЯ ОНТОЛОГИЙ НАУЧНЫХ ПРЕДМЕТНЫХ ОБЛАСТЕЙ НА ОСНОВЕ ПАТТЕРНОВ ОНТОЛОГИЧЕСКОГО ПРОЕКТИРОВАНИЯ

Аннотация

В настоящее время онтологии являются основным средством формализации и систематизации знаний в различных областях знаний, включая научные предметные области (НПО). Для разработки онтологий предложено много различных подходов и методов. В последнее время набирает популярность подход, базирующийся на паттернах онтологического проектирования (паттернах ОП). Применение таких паттернов значительно облегчает процесс построения онтологий, но существуют проблемы их использования. Одна из проблем состоит в том, что при построении онтологий НПО требуются паттерны, ориентированные на представление научных знаний, а известные каталоги паттернов ОП таких паттернов, как правило, не содержат. Другие проблемы вызваны трудностями повторного использования паттернов ОП, разработанных другими разработчиками, из-за их большой сложности, а также отсутствием удобных инструментов, поддерживающих использование паттернов при разработке онтологий. В статье описывается подход к решению проблем использования паттернов ОП при построении онтологий научных предметных областей. Данный подход предлагает систему паттернов ОП и методику их использования. Его применение позволяет экономить ресурсы, затрачиваемые на разработку онтологий, избегать ошибок онтологического моделирования, а также обеспечивать согласованное представление всех сущностей онтологии.

Ключевые слова: онтология, научная предметная область, паттерн онтологического проектирования, методология построения онтологий.

Введение

В связи с проникновением технологий Semantic Web практически во все сферы человеческой деятельности появилась насущная потребность в формализации знаний различных предметных областей в виде онтологий. Особенно остро стоит проблема построения онтологий научных предметных областей (НПО), к которым обычно относят предметные области, охватывающие некоторую научную дисциплину или область научных знаний во всех ее аспектах.

Для разработки онтологий применяются различные методологии и подходы [1–4]. В последнее время интенсивно развивается подход, использующий паттерны онтологического проектирования (Ontology Design Patterns) или паттерны ОП (ODPs) [5–8]. Согласно нему паттерны ОП представляют собой документально зафиксированные описания проверенных на практике решений типовых проблем онтологического моделирования. Несмотря на то, что использование паттернов ОП позволяет сэкономить человеческие ресурсы и повысить качество разрабатываемых онтологий, они пока не нашли широкого практического применения из-за ряда проблем, возникающих при их использовании.

В статье дается краткий обзор паттернов онтологического проектирования, анализируются проблемы их использования, описывается предложенная авторами методология построения онтологий научных предметных областей, базирующаяся на паттернах ОП. Паттерны ОП, используемые в данной методологии, появились в результате решения проблем онтологического моделирования, с которыми авторы статьи столкнулись в процессе разработки онтологий для различных научных предметных областей [9].

Краткий обзор паттернов онтологического проектирования

Паттерны онтологического проектирования имеют в качестве своих прародителей шаблоны проектирования (design pattern), широко используемые в разработке программного обеспечения. Аналогично шаблонам проектирования, паттерны ОП предназначены для описания решений типичных проблем, возникающих при разработке онтологий [7].

В зависимости от проблем, для решения которых предназначены паттерны онтологического проектирования, различают структурные паттерны (Structural ODPs), паттерны соответствия (Correspondence ODPs), паттерны содержания (Content ODPs), паттерны логического вывода (Reasoning ODPs), паттерны представления (Presentation ODPs) и лексико-синтаксические паттерны (Lexico-Syntactic ODPs). (Заметим, что эта типология паттернов была предложена в рамках проекта NeOn [10]).

При разработке онтологий чаще всего используются структурные паттерны, паттерны содержания и представления.

Структурные паттерны либо фиксируют способы решения проблем, вызванных ограничениями выразительных возможностей языков описания онтологий, либо задают общую (модульную) структуру и вид онтологии. Паттерны первого типа называются логическими паттернами (Logical ODPs), паттерны второ-

го типа – архитектурными паттернами (Architectural ODPs).

Паттерны содержания задают способы представления типовых фрагментов онтологий, на основе которых могут строиться онтологии различных предметных областей.

Паттерны представления определяют рекомендации по именованию и аннотированию элементов онтологии, применение которых должно повысить читаемость онтологии, а также удобство и простоту ее использования.

В настоящее время создано и развивается несколько каталогов паттернов ОП [11-13]. Наиболее представительный из них размещен на портале Ассоциации ODPA (Association for Ontology Design & Patterns) [11], созданном в рамках проекта NeOn [10].

Паттерны ОП чаще всего описываются в формате, предложенном на портале ассоциации ODPA [11]. Согласно нему описание паттерна включает сведения о его авторе и области применения, его графическое представление, текстовое описание, набор сценариев, примеров и ссылки на другие паттерны. Паттерны содержания также могут снабжаться набором вопросов компетенции (Competency questions) [6, 7], которые могут использоваться как при разработке паттернов, так и для поиска нужных паттернов при разработке конкретной онтологии.

Проблемы использования паттернов онтологического проектирования

Первая проблема повторного использования паттернов обусловлена их сложностью – зачастую разработчику новой онтологии трудно понять семантику, которую заложили в паттерн его авторы. В последнее время наблюдается тенденция к упрощению паттернов [14]. Появились даже, так называемые, метапаттерны, которые описывают очень простые сущности [15]. Однако такие простые паттерны не могут существенно облегчить построение онтологий НПО.

Другая проблема вызвана отсутствием удобных инструментов разработки онтологий, поддерживающих использование паттернов ОП. Здесь можно отметить плагины для инструмента разработки онтологий проекта NeOn [10] и редактора онтологий WebProtégé [16]. Однако первый плагин доступен только для участников проекта NeOn, а второй может использоваться только в редакторе WebProtégé, который не очень популярен среди разработчиков онтологий из-за его ограниченной функциональности (по сравнению с настольной версией).

Третья проблема связана с отсутствием систематизированных наборов паттернов, ориентированных на специалистов в предметной области. Существующие каталоги паттернов не отвечают этому требованию.

На наш взгляд, наиболее близка к решению этой проблемы библиотека OTTR (Reasonable Ontology Templates) [17], которая обеспечивает язык для представления паттернов ОП и поддерживающие его программные средства. Эта библиотека предоставляет разработчикам онтологий паттерны в виде высокоуровневых OWL-макросов, что делает возможным их использование специалистами в предметной области.

Что касается наличия в широком доступе паттернов, которые можно было бы использовать при разработке онтологий НПО, то упомянутые выше каталоги даже частично не покрывают потребностей построения онтологий научных областей.

Методология разработки онтологий научных предметных областей

Главными особенностями методологии разработки онтологий НПО являются: структурирование онтологии НПО в виде ее разделения на несколько относительно независимых онтологий; использование базовых онтологий, включающих только самые общие сущности, не зависящие от конкретной НПО, и большей частью представленные паттернами; построение онтологии НПО на основе базовых онтологий путем их достройки и развития; пополнение онтологии НПО путем конкретизации (означивания) содержательных паттернов.

Структура и содержание онтологии НПО и базовых онтологий

Онтология любой НПО содержит не только описания присущих ей системы понятий и методов обработки и анализа информации, но и описания релевантных ей информационных ресурсов. В связи с этим онтологию НПО можно представить в виде системы взаимосвязанных онтологий, отвечающих за представление указанных выше трех компонентов знаний, а именно: онтологии области знаний, онтологии задач и методов и онтологии научных интернет-ресурсов.

Онтология области знаний задает систему понятий и отношений, предназначенных для детального описания моделируемой НПО и выполняемой в ее рамках научной и исследовательской деятельности. Онтология задач и методов описывает задачи, решаемые в данной НПО, и методы их решения. Онтология научных интернет-ресурсов служит для описания, представленных в сети Интернет информационных ресурсов, релевантных данной НПО.

Так как разработка онтологии НПО «с нуля» является непростой задачей, предложен метод ее построения на основе небольшого, но представительного набора базовых онтологий, включающих только самые общие сущности, не зависящие от конкретных НПО. В этот набор входят: (1) онтология научного знания, (2) онтология научной деятельности, (3) базовая онтология задач и методов, (4) базовая онтология информационных ресурсов.

Все базовые онтологии имеют спецификации на языке OWL [18].

Онтология научного знания содержит классы, задающие структуры для описания понятий, входящих в любую НПО. Такими понятиями являются Раздел науки, Объект исследования, Предмет исследования, Метод исследования, Научный результат и др.

Онтология научной деятельности включает классы понятий, относящиеся к организации научно-исследовательской деятельности, такие как Персона, Организация, Событие, Научная деятельность, Проект, Публикация и др.

Базовая онтология информационных ресурсов включает класс Информационный ресурс в качестве основного класса. Набор атрибутов и связей этого класса основан на стандарте Dublin core [19].

С помощью понятий и отношений базовой онтологии задач и методов могут быть описаны задачи, решаемые в данной НПО, методы их решения и реализующие их программные компоненты и алгоритмы.

Набор паттернов онтологического проектирования

Для поддержки рассматриваемой методологии был разработан набор паттернов ОП [20], реализованных на языке OWL. Этот набор включает паттерны трех типов: структурные логические паттерны, паттерны представления и содержания.

Необходимость в использовании структурных логических паттернов возникла из-за отсутствия в языке OWL выразительных средств для представления сложных сущностей и конструкций, актуальных при построении онтологий НПО, в частности, областей допустимых значений, многоместных и атрибутированных отношений (бинарных отношений с атрибутами).

Паттерн представления области допустимых значений предназначен для задания таких конструкций, которые в реляционной модели данных называются доменами и характеризуются названием и множеством элементарных значений. Домены удобно использовать при описании возможных значений свойств класса, когда весь набор таких значений известен заранее. В этом паттерне домен задается перечислимым классом, который является наследником специально введенного служебного класса *Домен* и состоит из конечного набора различных индивидов (объектов), определяющих возможные значения некоторого свойства (рис. 1).

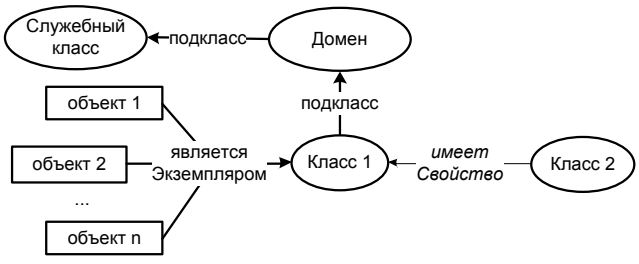


Рис. 1. Структурный паттерн представления области допустимых значений

Примерами таких доменов являются «Географический тип», «Должность», «Тип организации», «Тип публикации», которые включают соответственно виды населенных пунктов, виды должностей, типы организаций и публикаций.

Для представления атрибутированных отношений предложен структурный паттерн, показанный на рис. 2.

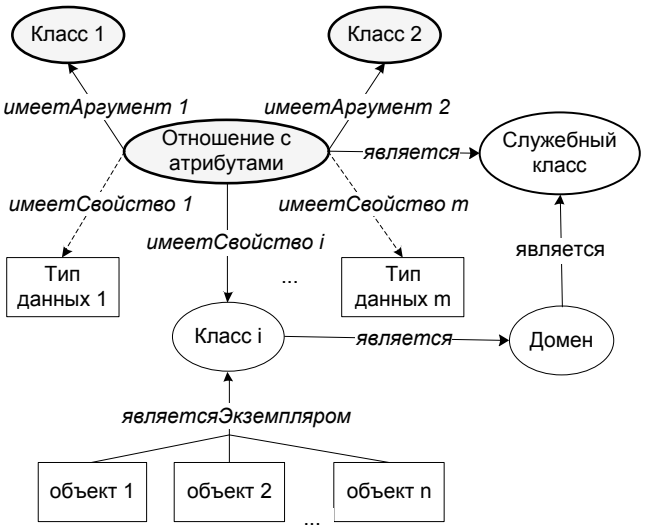


Рис. 2. Структурный паттерн бинарного атрибутированного отношения

Центральное место в данном паттерне занимает служебный класс *Отношение с атрибутами*, с которым связываются классы, моделирующие аргументы бинарного отношения. При этом атрибуты бинарного отношения моделируются свойствами данного класса (в нотации языка OWL либо *DataProperty*, либо *ObjectProperty*). Для представления конкретного типа отношения вводится новый класс – наследник класса *Отношение с атрибутами*.

Подобным образом построен паттерн для многоместного отношения.

Для единообразного и непротиворечивого представления используемых в НПО понятий и их свойств были разработаны паттерны содержания, описывающие основные понятия базовых онтологий. Благодаря этому, разработка онтологии конкретной НПО в основном состоит в специализации паттернов содержания и построении на их основе фрагментов целевой онтологии.

В качестве примера приведем паттерн, предназначенный для описания прикладных задач, решаемых в рамках научной предметной области (рис. 3).

С этим паттерном связывается следующий набор квалификационных вопросов:

Какими методами решается задача?

Какие данные используются для решения задачи?

Что является результатом решением задачи?

К какому разделу науки относится задача?

Кто формулирует задачу? и др.

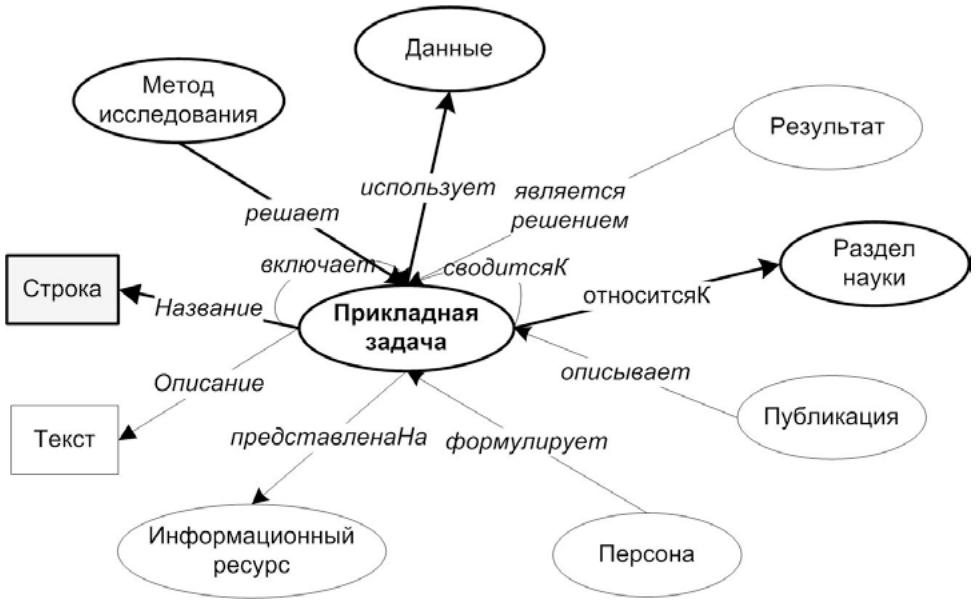


Рис. 3. Паттерн для описания прикладной задачи

Методика построения онтологии

Построение онтологии НПО включает два основных этапа:

1. Построение компонентов онтологии НПО на основе базовых онтологий путем их достройки и развития;

2. Пополнение онтологии НПО путем конкретизации (означивания) паттернов содержания, представленных в базовых онтологиях.

При этом составная часть онтологии НПО – онтология области знаний – строится на основе онтологий научного знания и научной деятельности, онтология задач и методов – на основе базовой онтологии задач и методов, онтология научных интернет-ресурсов – на основе базовой онтологии интернет-ресурсов.

Использование паттернов содержания поддерживается специальным редактором, который позволяет специалистам в предметной области пополнять онтологию фактическими данными – объектами классов и их свойствами. При этом пользователь из представленной ему иерархии классов выбирает нужный класс, редактор по имени класса находит соответствующий паттерн и на его основе строит форму, содержащую поля для заполнения свойств объекта этого класса пользователем.

Выводы

В статье рассмотрены проблемы применения паттернов онтологического проектирования для разработки онтологий научных предметных областей. Представлена методология разработки онтологий НПО, решающая большинство из этих проблем. Данную методологию поддерживают набор паттернов онтологического проектирования, описывающих основные конструкции и сущности, необходимые для описания научных предметных областей, и редактор данных, позволяющий пополнять онтологию фактическими данными путем означивания паттернов содержания. Благодаря простоте паттернов и редактора данных ими могут пользоваться не только инженеры знаний, но и специалисты в моделируемой области знаний.

Данная методология показала свою практическую полезность при разработке онтологий различных научных предметных областей («Поддержка принятия решений» [21], «Активная сейсмология» [22] и др.).

Благодарности

Работа выполнена при частичной финансовой поддержке РФФИ (проект № 19-07-00762).

Список литературы

1. Fernández-López M., Gómez-Pérez, A., Pazos, A., Pazos, J., 1999. Building a Chemical Ontology Using Methontology and the Ontology Design Environment. IEEE Intelligent Systems & their applications, 4(1): 37–46.

2. Sure Y., Staab, S., Studer, R., 2009. Ontology Engineering Methodology. In Handbook on Ontologies, Eds., Staab S., Studer R. Berlin: Springer Verlag, pp: 135–152.

3. Pinto H., Staab, S., Tempich, C., 2004. DILIGENT: Towards a fine-grained methodology for DIstributed, Loosely-controlled and evolvInG Engineering of oNTologies. In the proceedings of the 16th European Conference on Artificial Intelligence. Frontiers in Artificial Intelligence and Applications, 110. IOS Press, pp. 393–397.

4. De Nicola A., Missikoff, M., Navigli, R.A., 2005. Proposal for a Unified Process for Ontology Building: UPON. In: Database and Expert Systems Applications, Eds., Andersen, K.V., J. Debenham, R. Wagner DEXA 2005. Lecture Notes in Computer Science, 3588: 655–664.

5. Gangemi A., Presutti V., 2009. Ontology Design Patterns. In Handbook on Ontologies, Eds., Staab, S. and R. Studer. Berlin: Springer Verlag, pp: 221–243.

6. Blomqvist E., Hammar K., Presutti V., 2016. Engineering Ontologies with Patterns: The eXtreme Design Methodology, In Ontology Engineering with Ontology Design Patterns. Studies on the Semantic Web, Eds., Hitzler, P., and A. Gangemi, K. Janowicz, A. Krisnadhi, V. Presutti, IOS Press, pp: 23–50.

7. Karima N., Hammar K., Hitzler P., 2017. How to Document Ontology Design Patterns. In Advances in Ontology Design and Patterns, 32. IOS Press, Kobe, Japan, pp: 15–27.

8. Ломов П.А., 2015. Применение паттернов онтологического проектирования для создания и использования онтологий в рамках интегрированного пространства знаний. Онтология проектирования, Т.5, 2(16): 233–245.

9. Borovikova O., Globa L., Novogrudska R., Ternovoy M., Zagorulko G., Zagorulko Yu., 2012. Methodology for knowledge portals development: background, foundations, experience of application, problems and prospects. Joint NCC&IIS Bulletin, Series Computer Science, 34: 73–92.

10. NeOn Project. Date Views 07.02.2019 http://neon-project.org/nw/Welcome_to_the_NeOn_Project.html.

11. Портал ассоциации ODPA (Association for Ontology Design & Patterns). Date Views 07.02.2019 http://ontologydesignpatterns.org.

12. Ontology Design Patterns (ODPs) Public Catalog. Date Views 07.02.2019 http://odps.sourceforge.net.

13. Dodds, L., Davis, I. Linked Data Patterns. 2012. Date Views 07.02.2019 http://patterns.dataincubator.org/book.

14. Krisnadhi A., Hitzler P., 2017. A Core Pattern for Events. In Advances in Ontology Design and Patterns, 32. IOS Press, Kobe, Japan, pp: 29–37.

15. Krisnadhi A., Hitzler P., 2017. The Stub Metapattern. A Core Pattern for Events. In Advances in Ontology Design and Patterns, 32. IOS Press, Kobe, Japan, pp: 29–45.

16. Hammar K., 2015. Ontology Design Patterns in WebProtégé. In the proceedings of 14th International Semantic Web Conference (ISWC-2015). Posters & Demonstrations Track. CEUR Workshop Proceedings, 1486: Date Views 07.02.2019 http://ceur-ws.org/Vol-1486/paper_50.pdf.

17. Skjæveland M.G., Forssell H., Klüwer J.W., Lupp D., Thorstensen E., Waaler A., 2017. Pattern-Based Ontology Design and Instantiation with Reasonable Ontology Templates. In the proceedings of the 8th Workshop on Ontology Design and Patterns (WOP 2017). Vienna, Austria, October 21. CEUR Workshop Proceedings, 2043. Date Views 07.02.2019 http://ceur-ws.org/Vol-2043/paper-04.pdf.

18. Antoniou G., Harmelen F., 2009. Web Ontology Language: OWL. In Handbook on Ontologies, Eds., Staab, S. and R. Studer. Berlin: Springer Verlag, pp: 91–110.

19. DCMI Metadata Terms. Date Views 07.02.2019 http://dublincore.org/documents/dcmi-terms.

20. Zagorulko Y., Borovikova O., Zagorulko G., 2018. Development of Ontologies of Scientific Subject Domains Using Ontology Design Patterns. In Communications in Computer and Information Science, 822: 141–156.

21. Загорулько Г.Б., 2016. Разработка онтологии для интернет-ресурса поддержки принятия решений в слабоформализованных областях. Онтология проектирования, Т. 6, 4(22): 485–500.

22. Braginskaya L., Kovalevsky V., Grigoryuk A., Zagorulko G., 2017. Ontological approach to information support of investigations in active seismology. In the proceedings of the 2nd Russian-Pacific Conference on Computer Technology and Applications (RPC),Vladivostok, Russky Island, Russia, 25–29 September, pp. 27–29.

УДК 519.25 004.852

М.А. Карпов

РАСПОЗНАВАНИЕ ШУМА В ОТВЕТАХ ОРАКУЛА В ПРОЦЕССЕ АКТИВНОГО ОБУЧЕНИЯ

Аннотация

В современном мире мы обладаем огромным количеством оцифрованной информации связанной с различными сферами жизни, но не можем применять ее в следствие отсутствия меток, ассоциированных с этими данными. Одним из путей решения этой проблемы является применение методов активного обучения, также известных как обучение с частичным привлечением учителя. Методы активного обучения используют оракула для получения меток. Оракулом в алгоритмах машинного обучения зачастую выступает человек, а человеку свойственно ошибаться. В данной работе предлагается метод выявления случайного шума в ответах оракула с целью последующего уточнения. Метод не требует информации извне, поэтому может применяться достаточно широко. Исследуется влияние этого метода на качество получаемой модели, а так же экспериментально оценивается качество распознавания ошибок оракула. Рассматриваются характеристики и метрики связанные с эффективностью активного обучения.

Метод показал высокие показатели точности распознавания ошибок при достаточно низких значениях полноты. Было отмечено положительное влияние алгоритма на качество получаемой модели, однако, несмотря да достаточно высокую точность распознавания ошибок оракула, оказываемое влияние оказалось не значительным.

Ключевые слова: активное обучение, несовершенный оракул, распознавание шума, машинное обучение, обучение с частичным привлечением учителя

Введение

С ростом количества не размеченных данных все больше возрастает потребность в методах активного обучения. Методы активного обучения, так же известные как обучение с частичным привлечением учителя, позволяют достичь значительного прироста качества работы модели при затратах значительно меньшего количества человеческих ресурсов для разметки данных [1]. Ключевая гипотеза заключается в том, что если позволить алгоритму машинного обучения выбирать данные для обучения (быть «любопытным»), его производительность возрастет при меньших затратах на обучение [2]. Большинство эмперических наблюдений подтверждает эту гипотезу [3]. Активное обучение, помимо прочего, эффективно справляется с задачами классификации несбалансированных наборов данных [4–6].

Оракулом в алгоритмах машинного обучения зачастую выступает человек. Минус состоит в том, что человеку свойственно ошибаться. Это могут быть как систематические ошибки, основанные на искажениях восприятия, так и случайные, основанные на невнимательности. Существуют исследования активного обучения с несовершенным оракулом. В некоторых случаях ошибки вносились случайно и равномерно [7]. В других работах предпринимались попытки имитации человеческих ошибок [8–10]. Также проводились исследования с применением нескольких несовершенных оракулов с различной «стоимостью» получения ответа [11]. В работах предпринимаются попытки создания моделей, которые могли бы быть устойчивыми к вносимому шуму [12]. Так же предлагаются методы обнаружения шума с использованием доменной адаптации [13].

Упомянутые выше работы либо не предпринимают попыток исправлять ошибки оракула, либо исправляют их за счет внешней информации. Как пример, при использовании доменной адаптации необходима обученная на другом домене модель.

Целью данной работы является исследование возможности распознавания случайного шума в ответах оракула в процессе активного обучения без использования дополнительной (внешней) информации. Ставятся следующие задачи:

- Разработать метод распознавания случайных ошибок оракула без использования дополнительных источников информации;
- Экспериментально оценить качество распознавания ошибок;
- Оценить степень влияния исправления ошибок на полученную в результате модель.

Методика

Формализуем проблему активного обучения с несовершенным оракулом. Обозначим пространство входных векторов как X, а множество меток Y. В данной работе рассматривается задача бинарной классификации, поэтому $Y=\{-1,1\}$. Буквой H обозначим множество моделей машинного обучения, т.е. отображений $X \rightarrow Y$, при том каждой метке y_i', полученной с помощью модели $h(x_i)$ вычисляется степень уверенности c_i в диапазоне $\overline{[0,1]}$. Символом $O: \forall x_i \in X, O(x_i) = y_i, y_i \in Y$ обозначим оракул.

Дана конечная не размеченная выборка и набор размеченных примеров $A \subset (X \times Y)$. На начальном этапе набор A должен быть не пуст. Так же известен набор истинных меток $Y_n \subset Y : \forall x_i \in X_n \exists y_i \in Y_n$, который

является истинной меткой. Набор Y_n не участвует в алгоритме активного обучения и используется только для внешней оценки качества предсказаний модели.

Алгоритм активного обучения состоит в следующем: на основе данных A обучается новая модель $h \in H$. Для каждого $x_i \in X_n$ вычисляется $y_i' = h(x_i)$, а также степень уверенности c_i. Затем выбирается такое x_i', что c_i минимально. Значение x_i' извлекается из X_n. Значение $(x_i', O(x_i'))$ помещается в A. Далее алгоритм повторяется вновь. Одно повторение описанных действий будем называть шагом активного обучения.

Истинным ответом оракула для примера x_i будем называть такое $y_i = O(x_i)$, что $y_i \in Y_i$. Ложным ответом оракула для примера будем называть такое $y_i = O(x_i)$, что $y_i \neq y_i$.

На каждом шаге активного обучения подсчитывается качество модели. В качестве метрики используется F мера. Для подсчета этой оценки используется набор Y_n в качестве истинных меток. Для оценки качества полученной модели в результате эксперимента в целом для последних k значений F меры подсчитывается медиана и ее среднеквадратичное отклонение. Чем больше значение медианы, тем точнее в среднем работает полученная модель. Чем меньше значение среднеквадратичного отклонения медианы, тем выше вероятность получить хорошую модель остановившись на случайном шаге из последних k. Обозначим символом F_k медиану F меры на последних k шагах активного обучения.

Основная часть

В данной работе рассмотрен метод идентификации ложных ответов оракула *0* в процессе активного обучения. В каждом эксперименте оракулу подавалось на вход фиксированное количество примеров из X_n. Кроме того задается фиксированное и определенное количество примеров, на которые оракул дает ложный ответ. Номера ложных ответов выбираются случайным образом. Каждый эксперимент повторялся дважды: в первом случае алгоритм доверяет ответам оракула, а во втором пытается обнаружить и исправить ложные ответы. При этом гарантируется, что при повторном обращении оракул всегда дает истинный ответ.

Предлагается следующий способ обнаружения ложных ответов оракула: для каждого $x_i \in X_n$ задается $e_i = 0$. На каждом шаге алгоритма активного обучения, после получения новой модели $h \in H$, для каждого $(x_i, y_i) \in A$ вычисляется $y_i' = h(x_i)$. Если какой-то $y_i \neq y_i'$, то значение e_i увеличивается на единицу. Как только значение e_i превышает заданный порог t – пример x_i снова подается на вход оракулу для повторной разметки. Затем пример (x_i, y_i) извлекается из A, а на его место записывается $(x_i, 0(x_i))$, где $0(x_i)$ уже гарантированно является истинным ответом. Значению e_i вновь присваивается значение 0 и алгоритм активного обучения продолжается.

Гипотеза состоит в том, что ложные ответы оракула будут противоречить его предыдущим истинным ответам, и модель, полученная в результате обучения на наборе A будет обнаруживать это противоречие.

Эксперимент

В качестве входных данных был использован классический набор данных с классификацией операций по кредитным картам на мошеннические и подлинные [14]. В качестве модели был выбран случайный лес [15]. Обоснованно это тем, что для этой модели удобно применять методы подсчета уверенности для заданных предсказаний [16].

В качестве исходной выборки в A помещалось 2 примера из X_n с истинными значениями метки (положительный и отрицательный). В каждом эксперименте производилось 100 шагов активного обучения. Порогу t эмпирически было присвоено значение 3. В каждом эксперименте оракул давал от 0 до 34 ложных ответов. На рисунке 1 изображен график одного

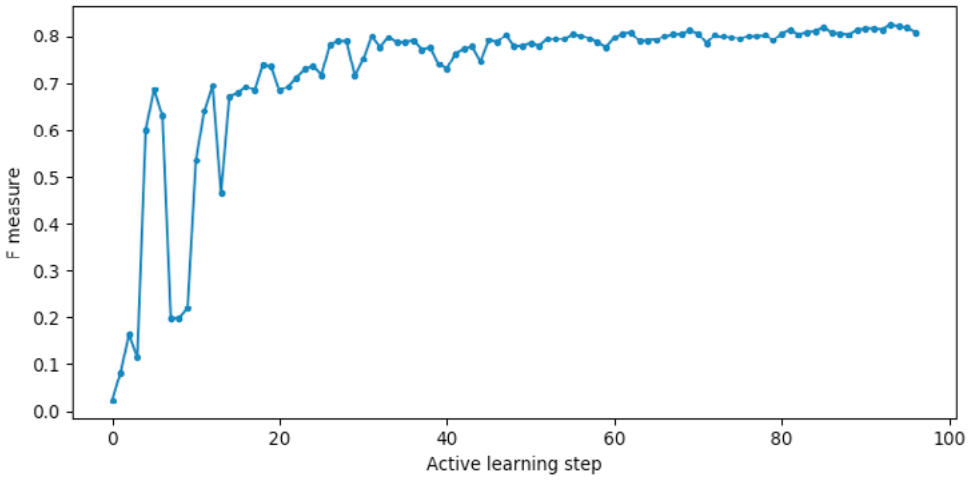


Рис. 1. Эксперимент без внесения ложных ответов. Каждая точка отражает значение F меры на данном шаге активного обучения

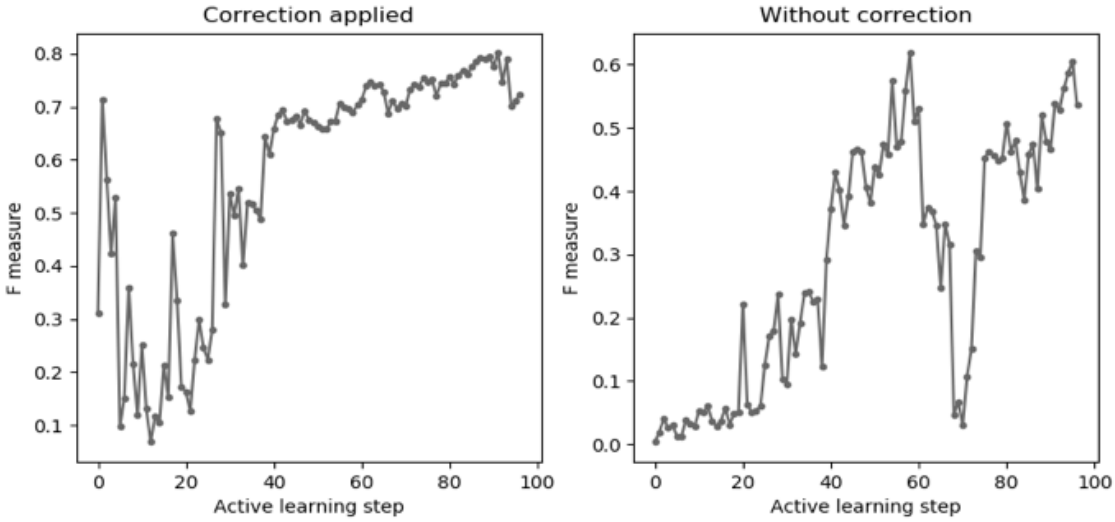


Рис. 2. Пример эксперимента в котором исправление ложных ответов положительно повлияло на качество полученной модели

эксперимента без внесения ложных ответов. Для каждого набора вносимых ложных ответов эксперимент проводился дважды: с применением распознавания и исправления ошибок и без него.

Эксперименты оказались не однозначны. Рисунок 2 показывают ход эксперимента с внесением 11 ложных ответов, с распознаванием ложных ответов и без, соответственно. В этой паре экспериментов исправление ошибок позволило стабилизировать качество модели на достаточно высоком уровне. Однако есть и обратные примеры (рис. 3), когда исправление ложных ответов не только не помогает, но и ухудшает результаты.Для исследования влияния исправления ложных ответов было проведено более 800 пар экспериментов (с распознаванием ложных ответов и без).

Исследовано качество распознавания ложных ответов. На рисунке 4 изображена доля распознанных ложных ответов по отношению к общему количеству внесенных ошибок. Можно заметить, что в среднем данный алгоритм способен распознать чуть больше четверти внесенных ложных ответов. Однако следует отметить, что количество ложно-положительных примеров в среднем не превышает 1. Таким образом среднее значение точности (precision) по всем экспериментам составляет 0.85, а полнота (recall) 0.27. Проведен анализ изменения качества полученной модели в случае применения алгоритма исправления ложных ответов. Эксперименты группировались по количеству внесенных ложных ответов. В каждой группе вычислялись средние значения медианы и стандартного отклонения метрики (медиана меры на последних 50 шагах активного обучения). Для медианы вычислялся доверительный интервал ([alpha]=0.99). На рисунке 5 можно наблюдать как изменялось качество полученной модели в

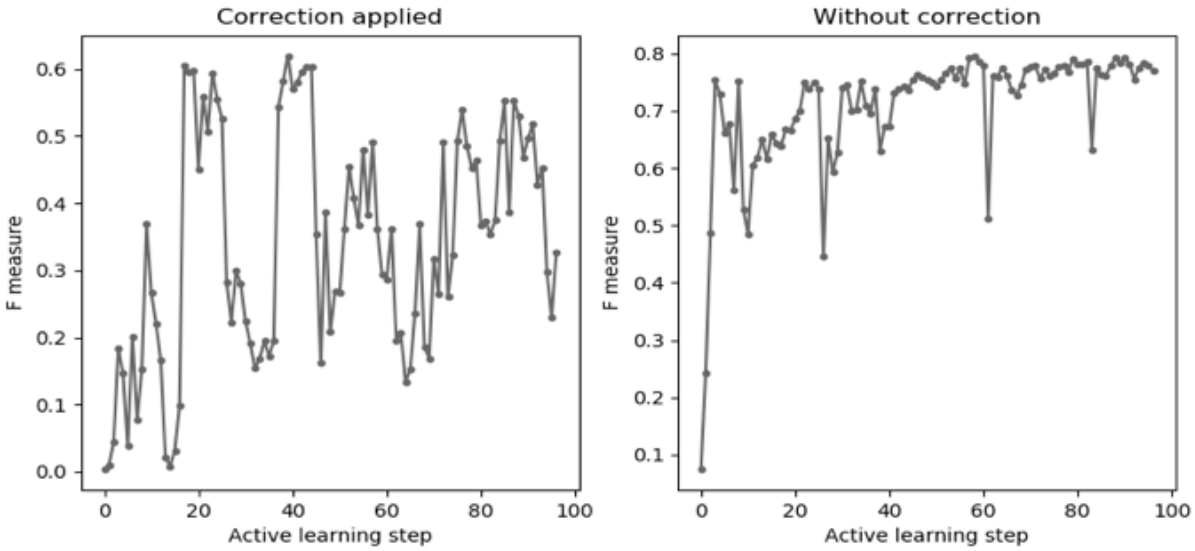


Рис. 3. Пример эксперимента в котором исправление ложных ответов отрицательно повлияло на качество полученной модели

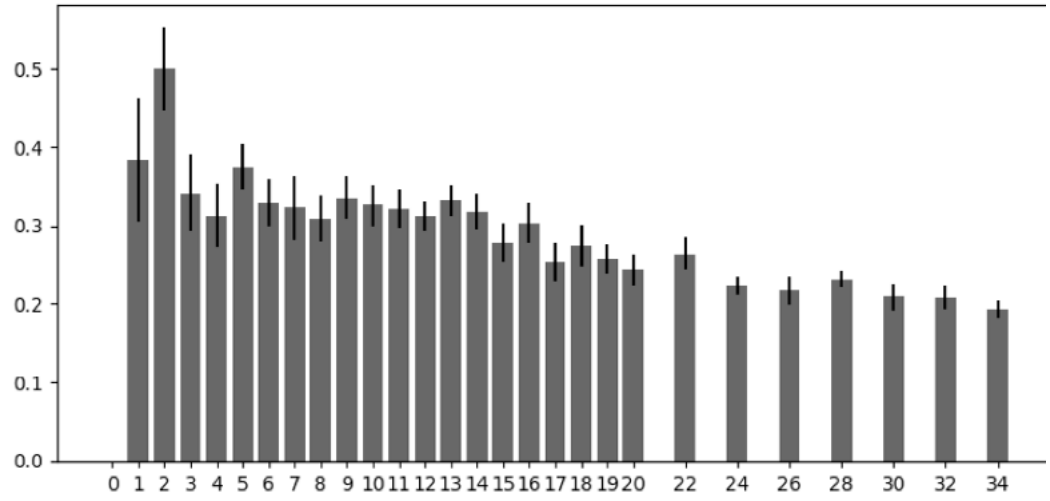


Рис. 4. Доля исправленных ложных ответов по отношению к общему количеству. Тонкие линии показывают среднеквадратичное отклонение показателя. Ось x отражает количество внесенных ошибок

зависимости от количества внесенных ложных ответов. Можно заметить, что несмотря на то, что исправление ложных ответов в среднем улучшает качество модели, оно не значительно меняет ситуацию в целом.

Также была выдвинута гипотеза, что успех распознавания ложных ответов зависит от распределения ложных ответов по шагам активного обучения. Результат эксперимента считался успешным в случае, если стандартное отклонение медианы не превышает 0.05 и медиана по данному эксперименту превышает среднюю медиану по данному классу количества внесенных ложных ответов. Для проверки гипотезы все эксперименты были разбиты на 6 групп:

1. Успешный результат без исправления ложных ответов.

2. Не успешный результат без исправления ложных ответов.

3. Успешный результат с исправлением ложных ответов.

4. Не успешный результат с исправлением ложных ответов.

5. Успешный результат с исправлением ложных ответов при условии, что аналогичный эксперимент без исправления ошибок был не успешен.

6. Не успешный результат с исправлением ложных ответов при условии, что аналогичный эксперимент без исправления ошибок был успешен.

В каждой группе интервал из 100 ответов разбивался на 5 равных частей и подсчитывалось, сколько ложных ответов попало в каждую из них. На рисунке 6 изображены усредненные распределения ложных ответов для каждой из групп. Явной зависимости не обнаружено.

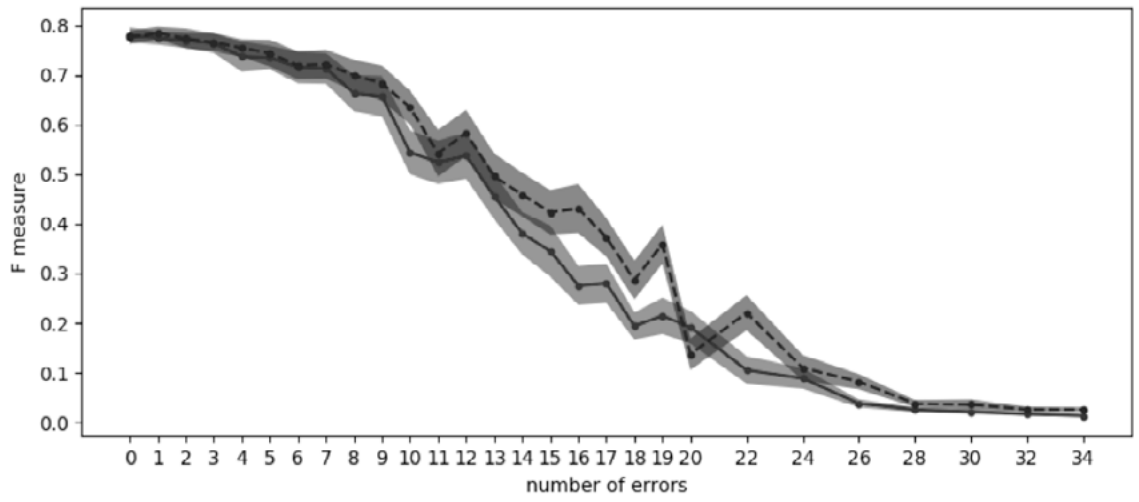


Рис. 5. Средняя F мера полученной модели в зависимости от количества вносимых ложных ответов. Прерывистая линия отражает эксперименты с применением исправления ложных ответов. Сплошная линия – эксперименты без использования исправлений. Фон линии показывает доверительный интервал

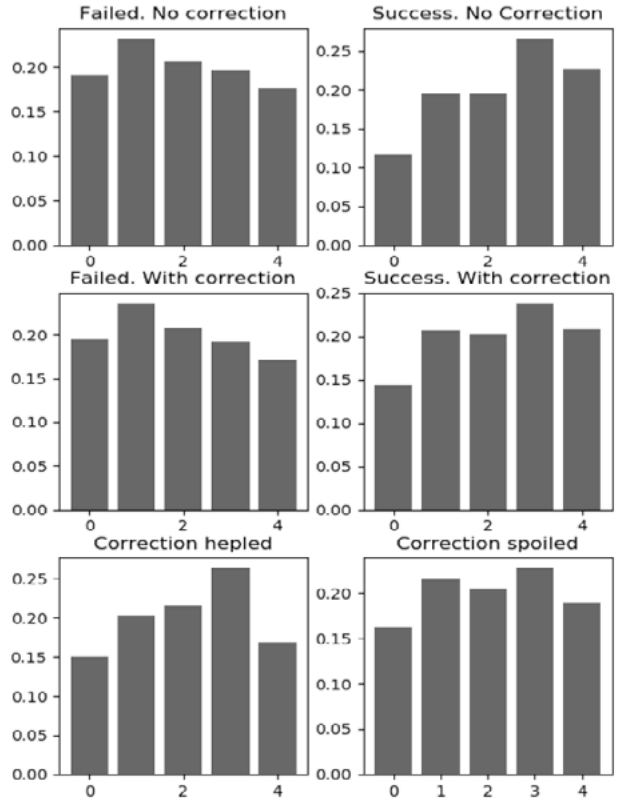


Рис. 6. Распределение ошибок ложных ответов по шагам активного обучения

Выводы

Был предложен метод распознавания и исправления случайного шума в ответах оракула без использования дополнительной (внешней) информации в процессе активного обучения. Метод позволяет распознавать небольшую часть ложных ответов с низким уровнем ложно-положительных предсказаний. Метод не способен значительно повлиять на качество полученной модели, однако в среднем не ухудшает результаты. Данный подход может быть использован совместно с другими подходами для комплексной борьбы с шумом оракула, описанными в работах [7, 12, 13].

Следующим шагом в данном исследовании может быть проверка данного алгоритма на наборах данных из других доменов для проверки доменной независимости. Кроме того имеет смысл попробовать заменить модель случайного леса на другие модели, например случайное дерево принятия решений, SVM, глубокие нейронные сети и т. д., а так же ансамбли из различных комбинаций моделей. Помимо этого имеет смысл исследовать данный алгоритм при различных значениях параметра t. Таким образом метод имеет перспективы, которые будут рассмотрены в дальнейших исследованиях.

Список литературы

1. Ireson N., Ciravegna F., Califf M.E., Freitag D., Kushmerick N., & Lavelli A., 2005. Evaluating machine learning for information extraction. In Proceedings of the 22nd international conference on Machine learning (pp. 345-352). ACM.

2. Settles B., 2014. Active learning literature survey. 2010. Computer Sciences Technical Report, 1648.

3. Tomanek K., & Olsson F., 2009. A web survey on the use of active learning to support annotation of text data. In Proceedings of the NAACL HLT 2009 workshop on active learning for natural language processing (pp. 45–48). Association for Computational Linguistics.

4. Ertekin S., Huang J., Bottou L., & Giles L., 2007. Learning on the border: active learning in imbalanced data classification. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management (pp. 127–136). ACM.

5. Trittenbach H., Englhardt A., & Böhm K., 2018. An Overview and a Benchmark of Active Learning for One-Class Classification. arXiv preprint arXiv:1808.04759.

6. Sarawagi S., & Bhamidipaty A., 2002. Interactive deduplication using active learning. In Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 269–278). ACM.

7. Natarajan N., Dhillon I.S., Ravikumar P.K., & Tewari A., 2013. Learning with noisy labels. In Advances in neural information processing systems (pp. 1196–1204).

8. Beygelzimer, A., Hsu, D. J., Langford, J., & Zhang, T., 2010. Agnostic active learning without constraints. In Advances in Neural Information Processing Systems (pp. 199–207).

9. Balcan M.F., Beygelzimer A., & Langford J. (2009). Agnostic active learning. Journal of Computer and System Sciences, 75(1), 78–89.

10. Du J., & Ling C.X., 2010. Active learning with human-like noisy oracle. In 2010 IEEE International Conference on Data Mining (pp. 797–802). IEEE.

11. Donmez P., & Carbonell J.G., 2008. Proactive learning: cost-sensitive active learning with multiple imperfect oracles. In Proceedings of the 17th ACM conference on Information and knowledge management (pp. 619–628). ACM.

12. Zhu X., Loy C.C., & Gong S., 2013. Constrained clustering: Effective constraint propagation with imperfect oracles. In Data Mining (ICDM), 2013 IEEE 13th International Conference on (pp. 1307–1312). IEEE.

13. Liu G., Yan Y., Subramanian R., Song J., Lu G., & Sebe N., 2016. Active domain adaptation with noisy labels for multimedia analysis. World Wide Web, 19(2), 199–215.

14. Credit Card Fraud Detection. Anonymized credit card transactions labeled as fraudulent or genuine. https://www.kaggle.com/mlg-ulb/creditcardfraud

15. Breiman L., 2001. Random forests. Machine learning, 45(1), pp. 5–32.

16. Zhou Z.H., & Li M., 2010. Semi-supervised learning by disagreement. Knowledge and Information Systems, 24(3), 415–439.

УДК 519.766.4 004.852

Ю.Д. Кленин

ОБЗОР ПРИМЕНЕНИЯ ПОДХОДОВ ГЛУБОКОГО ОБУЧЕНИЯ В ЗАДАЧЕ СТРУКТУРИРОВАНИЯ ТЕКСТОВОЙ ИНФОРМАЦИИ

Аннотация

Объемы информации, доступные современным системам, постоянно и с большой скоростью возрастают. Большая часть этой информации представлена в виде простых текстов (“plain text”), являющихся, так называемыми, “сырыми” (“raw”) данными в задачах обработки натурального языка. Такие данные по большей части слабоструктурированны, или совершенно не имеют жесткой структуры, что делает их сложными для анализа и использования как информационными системами, так и человеком. Кроме того, такие тексты принадлежащие различным предметным областям знания, также склонны иметь большое количество различий и большую вариацию в используемой лексике и стилях речи. Эти факторы затрудняют перенос и повторное использование ранее обученных систем на новые задачи. В результате, обучение моделей и ML-алгоритмов (machine learning, машинное обучение) требует значительных объемов ручного труда экспертов на подготовку размеченных наборов учебных данных для каждой отдельно взятой задачи. В данной работе предлагается создание автоматизированного средства преобразования слабоструктурированного простого текста к более жесткой онтологической модели графа знания, устойчивого к смене предметных областей. Рассматриваются возможности использования подходов переноса знания и активного обучения для сокращения человеческих трудозатрат в процессе обучения нейросетевой модели, а также различные варианты применения архитектуры seq2seq (sequence-to-sequence, последовательность-в-последовательность) в задаче структурирования данных.

Ключевые слова: нейронная сеть, sequence-to-sequence, семантическая сеть, извлечение сущностей, граф знаний, активное обучение, структурирование текста.

Введение

Объемы данных, производимых человечеством, возрастают все быстрее и быстрее. Большая часть этих данных представлена или представима (средствами распознавания речи и изображения) в виде текстовой информации. Причем главным образом это так называемая «сырая» (“raw”) информация, оформленная в виде «простого» (“plain”), неструктурированного текста. Основной недостаток таких данных – низкая скорость и высокая сложность их обработки как человеком, так и информационными системами. Например, одна из крупных групп методов машинного обучения – так называемые алгоритмы «с учителем» (“supervised”) – требуют наличия заранее подготовленных, размеченных данных, которые используются для обучения моделей. Такие данные, в отличие от неструктурированных текстов, достаточно редки и, как правило, требуют больших затрат ручного труда экспертов, повышая сложность и стоимость процесса создания и обучения таких моделей.

Даже алгоритмы способные обучаться без размеченной учебной выборки («без учителя», “unsupervised”) встречают определенные трудности с неструктурированными данными. Во-первых, такие алгоритмы требуют существования некоторой процедуры оценки качества работы системы, что также сводится к использованию ручного труда экспертов. Во-вторых, обучение таких алгоритмов требует подбора таких данных, которые были бы достаточно различны чтобы быть представительными, но достаточно похожи, чтобы система могла научиться находить в них закономерности.

Кроме того, за исключением крупных компаний работающие с большими данными, большинство приложений информационных систем требуют решения задач в рамках достаточно узких предметных областей, для которых, зачастую, невозможно подобрать достаточно большую для “unsupervised” модели выборку учебных документов, а подготовка данных для “supervised” модели требует больших затрат труда.

Таким образом, чтобы найти баланс между качеством работы систем, решающих конкретные задачи и затратами на их применение, необходимо автоматизировать процесс структурирования документов. «Структурирование» (“structuring”) является распространенным в литературе термином применительно к структурированию текстов – приданию текстам некоторой структуры. Здесь и далее, термин «структурированное знание» определяет организованную, систематизированную совокупность знаний, в то время как понятие «структура» используется для определения самой такой организации или системы – принципов и правил определяющих порядок структурирования знания. С данной позиции, подавляющее большинство текстов на натуральных языках являются «неструктурированными» – не имеющими явно выраженного структурированного знания.

В данной работе рассматриваются возможности применения двух популярных в литературе подходов к задаче автоматизации структурирования документов:

• transfer learning («перенос обучения», также известен как «передача обучения»);

• active learning («активное обучение»).

Также, рассматриваются возможности использования набирающей популярность методики seq2seq, (sequence-to-sequence, последовательность в последовательность) для решения задачи структурирования документов.

Transfer learning

Основной принцип transfer learning состоит в обучении модели для задачи с достаточным количеством учебных данных, с последующим переносом всей модели или ее части, на модель, которая применяется уже непосредственно для дообучения на целевой задаче. Эффективность переноса обучения зависит от того, насколько близки два задания. Так, в работе [1], авторы рассматривают переход от одного языка к другому как экстремальный пример адаптации предметной области, используя уже подготовленную базу знания на языке-доноре, для выявления отношений в отдельных, неструктурированных фразах. Следует отметить, что качество практических результатов достаточно низкое (20 баллов по метрике F1).

В [2] авторы показывают, на примере NER (Named Entity Recognition), что применив transfer learning из другой предметной области и дообучившись на размеченных вручную 16 % от целевой коллекции, можно добиться того же качества, что и обучившись на 34 % целевой коллекции, что означает более чем двукратное сокращение требуемого объема труда экспертов.

В то время как предыдущая работа использует transfer learning чтобы преодолеть промежуток между предметными областями, в [3] рассматриваются возможности переноса знания между задачей языкового моделирования (предсказание слова по предшествующим) и NER, показывая что предварительное обучение позволяет повысить общее качество работы системы.

Более широкий анализ возможностей transfer learning производится в работе [4], где авторы оценивают его эффективность для обобщенной задачи разметки последовательностей. Конкретно, ими выделяется три подзадачи: перенос знания между предметными областями на примере POS-разметки двух разных коллекций, перенос знания между приложениями на примере POS-разметки и NER-разметки, и перенос знания между языками на примере NER-разметки коллекций на английском и испанском языках. По результатам сравнения на различных коллекциях документов, transfer learning наиболее эффективен в первой задаче и наименее эффективен в последней.

Если рассмотренные выше задачи являются относительным простыми примерами структурирования документа, то более близка к проблематике данного исследования, задача описанная в [5], где вопросно-ответная система, имея некоторый исходный вопрос, получает представление о предметной области в процессе диалога с пользователем, постепенно наполняя свой граф знания концептами и отношениями, необходимыми для ответа на этот вопрос. Для каждого высказывания пользователя строится граф отношений между сущностями этого высказывания, который добавляется к графу всего диалога. Система затем использует получившийся граф диалога, чтобы выбрать ответ на исходный вопрос. После этого, полученный граф диалога используется для расширения глобального графа знания. Одним из важных недостатков предлагаемой авторами системы является принцип определения концептов. Концепт определяется как набор концептных ключевых слов, имеющих общий корень, а для получения таких концептов, авторы пользуются стеммингом. Подобный подход, основывающийся, в лучшем случае, на морфологической близости, а в худшем – на форме символьной последовательности, порождает две существенные проблемы. Во-первых, не все однокоренные (а в худшем случае – просто похожие по написанию) слова описывают один и тот же концепт. Во-вторых, слова, описывающие один и тот же концепт могут иметь радикально отличные корни.

Active learning

Если перенос знания пытается снизить затраты на подготовку данных повторно используя менее подходящие, но уже подготовленные ранее данные, то активное обучение пытается максимально эффективно использовать те затраты труда, без которых не обойтись, повышая качество тех данных, которые приходится подготавливать вручную. Это достигается путем предоставления модели возможности интерактивно взаимодействовать с экспертом в процессе обучения, определяя разметка каких именно данных будет наиболее полезна системе. Соответственно, одним из важнейших вопросов в активном обучении является выбор принципа отбора данных для обработки экспертом.

Классическими алгоритмами отбора являются случайный отбор и отбор тех примеров, в которых модель наименее уверена. Уверенность для каждого примера определяется как вероятность назначаемая системой наиболее вероятному ответу. Так, в [6], посвященной задаче NER в медицине, авторы классические алгоритмы с предлагаемым ими алгоритмом, кластеризующим все примеры по тематике, а затем применяющим для каждого кластера отбор по неуверенности. Эмпирически, чем сложнее алгоритма, тем более сложные примеры он выбирает, что проявляется в больших временных затратах эксперта на разметку каждой сущности, а также общем количестве сущностей.

В [7], авторы математически показывают, что традиционная неуверенность имеет предвзятость к

выбору более длинных предложений, подтверждая предположения [6]. Для борьбы с этим, вводится MNLP (Maximum Normalized Log-Probability), нормализующий неуверенность по длине последовательности слов, и BALD (Bayesian Active Learning by Disagreement), оценивающий неуверенность модели как несогласие – число моделей в наборе со случайно примененным к их узлам дропаутом (dropout, исключение), не выбравших наиболее популярную метку для каждого слова. Авторы показали, что активное обучение достигает state-of-the-art используя лишь 25 % учебных данных, а также, что уже обучившийся алгоритм способен адаптироваться к изменению предметной области, выбирая при дообучении примеры ниболее отличающиеся от ранее встреченных.

Еще один подход на основе идеи разнообразия, схожий с [6] предлагается в [8], также применительно к задаче NER. В этой работе, алгоритм группирует примеры по близости их структур (учитывая части речи и синтаксические роли элементов текста). Аналогичное исследование проводится в [9], за исключением использования большего количества атрибутов: помимо POS-меток используются различные лингвистические, орфографические, морфологические, лексические, контекстные, и семантические атрибуты. В целом, основным недостатком этих подходов, являются именно эти атрибуты, требующие использования дополнительных инструментов для их извлечения, что может быть проблематично для предметных областей со специфическим жаргоном, стилем, или малоресурсным языком. Альтернативой может послужить использование векторных представлений, получаемых на основе конкретной коллекции.

Другой подход к активному обучению предложен в [10]. Здесь используется повторное запрашивание разметки некоторого примера. Несмотря на то, что такой подход сокращает общее количество размеченных данных, предполагается, что переразметка может повысить качество, исправляя потенциальные ошибки экспертов. Следует отметить, что по результатам статьи, после определенного объема размеченных данных, прирост качества становится незначительным, однако, на малых объемах данных, лидируют традиционные алгоритмы выбора примеров, что заставляет усомниться в эффективности подхода.

Если предыдущие алгоритмы использовали некоторую статистическую оценку каждого примера, то в [11] выбор рассматривается как марковский процесс принятия решений, в котором состояние определяется уже размеченными данными и параметрами самой модели. Для каждого примера алгоритм принимает решение либо отправить его эксперту, либо перейти к следующему примеру, используя reinforcement learning (обучение с закреплением), для вознаграждения алгоритма на основе качества аннотации на тестовых данных. Данный алгоритм поддерживает transfer learning, продемонстрированный на кросс-языковом применении модели. Важно отметить, что определенные сложности могут возникнуть в обучении данных моделей в задачах, где невозможно составление золотого стандарта тестовой выборки, и результаты работы системы должны оцениваться экспертом.

Авторы [12] используют информацию, содержащуюся в DBpedia для повышения качества извлечения сущностей связанных с медицинскими препаратами и их побочными реакциями из текстов в социальных медиа. Применив векторные представления концептов и отношений графа знаний позволяет получить предварительный, «серебряный» стандарт. Поверх этого стандарта становится возможным эффективно применить активное обучение для достижения «золотой» разметки. Таким образом, в этой работе происходит перенос общего знания DBpedia на конкретную предметную область и дообучение с помощью активного обучения позволяет достичь высокого уровня качества с минимальным использованием человеческих ресурсов для разметки целевой коллекции.

Подытоживая: и активное обучение, и перенос знания могут эффективно применяться в задаче структурирования документов. Более того, их совместное использование также позволяет добиться более высоких результатов. **Sequence-to-sequence**

Один из набирающих популярность способов обработки простых текстов, рассматривает структурирование информации как преобразование ее из одной формы в другую. Обобщенное название данного подхода: sequence-to-sequence («последовательность в последовательность» или просто “seq2seq”), и в его основе лежит кодирование документа в некоторое векторное представление, с последующим декодированием в последовательность меток.

Так, [13] посвящена применению seq2seq в режиме автоэнкодера (autoencoder – модель, обучаемая воспроизводить входную последовательность) для повышения качества векторных представлений сущностей и отношений в условиях нехватки учебных данных. Авторы затем осуществляют transfer learning, переносся параметры энкодера на модель извлечения отношений. Предобученные таким образом модели извлечения отношений сравнимы по качеству с моделями без переноса, однако, нуждаются лишь в половине, или даже четверти объема тренировочных данных.

Радикально отличный подход предложен в [14]. Здесь, seq2seq модель используется напрямую для генерации триплетов отношений и сущностей, получив на вход неструктурированный текст. Эта модель превосходит state-of-the-art. Одна из предложенных архитектур, MultiDecoder, использует по одному декодеру на каждое отношение. Эта модель способна адаптироваться к изменениям, поскольку при введении новых отношений или удалении старых, достаточно лишь заменить набор декодеров.

Проблема адаптации рассматривается также и в [15]. Эта работа посвящена задаче распознавания

именованных сущностей и необходимым преобразованиям, связанным с адаптацией модели под расширяющийся набор типов извлекаемых сущностей. Для этого, производится модификация архитектуры seq2seq: на выходе энкодера добавляется полносвязный слой, выдающий численный вектор с длиной, определяемой числом типов сущностей. При появлении новых сущностей, производится расширение данного слоя до необходимого размера, после чего, полученная модель обучается выявлять новые типы сущностей, одновременно тонко настраивая этот процесс для уже известных типов сущностей.

В [16] модель на основе seq2seq получает на вход пару сущностей и набор предложений, содержащих эти сущности, обучаясь генерировать набор отношений между этими сущностями, найденных в предложенных примерах. Для этого модель использует сверточный энкодер, объединяющий все переданные модели примеры в общее представление. Закодированная информация затем несколько раз проводится через декодировщик, который на каждом проходе генерирует одно отношение. Экспериментальная оценка качества показывает, что эти модели извлечения отношений показывают более высокое качество, чем другие нейросетевые подходы.

В целом, можно отметить плотную связь между всеми тремя рассмотренными в данном обзоре подходами и задачей адаптивного структурирования информации.

Выводы

Рассмотрев основные подходы, которые могут применяться в данном исследовании, возможно определить следующие шаги развития предложенной концепции. Первичный эксперимент должен быть нацелен на построение базовой модели знания, и должен фокусироваться на самых простейших отношениях. Для этого было принято решение произвести извлечение упоминаний и примитивного отношения “is-a” («это»). Наиболее подходящими для этого эксперимента коллекциями являются наборы словарных и энциклопедических определений, сфокусированных на отношении некоторой сущности с ее «типом» или «классом», определяемым еще одной сущностью. Такие коллекции являются частично структурированными, что также может понизить сложность задачи. Конкретно, для проведения этого эксперимента предлагается использование толкового словаря русского языка, содержащего более 40 тысяч определений.

Построив граф из сущностей, связанных отношениями «это», можно получить достаточно крупный, если и примитивный граф. При полноценном применении данной системы, предполагается, что такой первичный граф должен наполнится упоминаниями основных сущностей моделируемой предметной области, что позволит в дальнейшем фокусировать систему на их поиск в неструктурированных текстах. Такой эксперимент предполагает использование коллекции документов на некоторую тематику, например, корпуса новостных статей.

Основной проблемой для данных экспериментов является оценка качества модели. Поскольку уже размеченные коллекции такого типа крайне редки, и трудоемки в производстве, то весьма проблематично предоставить такую коллекцию для оценки качества работы модели. Кроме того, поскольку рассматриваемые выше нейросетевые подходы являются, как правило, генеративными моделями, то такие размеченные коллекции будет некорректно сравнивать с результатом работы модели, поскольку конкретные извлекаемые концепты и отношения могут не совпадать дословно, но быть близки, если не идентичны по смысловой составляющей.

Ввиду отмеченных трудностей, предлагается произведение экспертной оценки непосредственно результатов работы модели, нежели предварительной экспертной разметки данных. Также следует отметить, что подход активного обучения уже задействует экспертов непосредственно, препятствуя использованию заранее подготовленной коллекции документов.

Список литературы

1. Verga P., D. Belanger, E. Strubell, B. Roth, and A. McCallum. 2015. Multilingual relation extraction using compositional universal schema. arXiv preprint arXiv:1511.06396.

2. Lee J.Y., F. Dernoncourt, and P. Szolovits. 2017. Transfer Learning for Named-Entity Recognition with Neural Networks. arXiv preprint arXiv:1705.06273.

3. Sachan D.S., P. Xie, M. Sachan, and E.P. Xing. 2018. Effective use of bidirectional language modeling for transfer learning in biomedical named entity recognition. In Machine Learning for Healthcare Conference, pp. 383–402.

4. Yang Z., R. Salakhutdinov, and W.W. Cohen. 2017. Transfer learning for sequence tagging with hierarchical recurrent networks. arXiv preprint arXiv:1703.06345.

5. Hixon B., P. Clark, and H. Hajishirzi. 2015. Learning knowledge graphs for question answering through conversational dialog. In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 851–861.

6. Chen Y., T.A. Lask, Q. Mei, Q. Chen, S. Moon, J. Wang, K. Nguyen et al. 2017. An active learning-enabled annotation system for clinical named entity recognition. BMC medical informatics and decision making 17, no. 2 (2017): 82.

7. Shen Y., H. Yun, Z.C. Lipton, Y. Kronrod, and A. Anandkumar. 2017. Deep Active Learning for Named Entity Recognition. arXiv preprint arXiv:1707.05928.

8. Kim S., Y. Song, K. Kim, J.-W. Cha, and G.G. Lee. 2006. Mmr-based active machine learning for bio named entity recognition. In Proceedings of the Human Language

Technology Conference of the NAACL, Companion Volume: Short Papers, pp. 69–72.

9. Kholghi M., L. Sitbon, G. Zuccon, and A. Nguyen. 2015. Active learning: a step towards automating medical concept extraction. Journal of the American Medical Informatics Association 23, no. 2 (2015): 289–296.

10. Lin C.H., M. Mausam, and D.S. Weld. 2016. Re-Active Learning: Active Learning with Relabeling." In AAAI, pp. 1845–852.

11. Fang M., Y. Li, and T. Cohn. 2017. Learning how to active learn: A deep reinforcement learning approach. arXiv preprint arXiv:1708.02383.

12. Stanovsky G., D. Gruhl, and P. Mendes. 2017. Recognizing mentions of adverse drug reaction in social media using knowledge-infused recurrent models. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers, vol. 1, pp. 142–151.

13. Li Z., L. Qu, Q. Xu, and M. Johnson. 2016. Unsupervised Pre-training With Seq2Seq Reconstruction Loss for Deep Relation Extraction Models. In Proceedings of the Australasian Language Technology Association Workshop 2016, pp. 54–64.

14. Zeng X., D. Zeng, S. He, K. Liu, and J. Zhao. 2018. Extracting relational facts by an end-to-end neural model with copy mechanism. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), vol. 1, pp. 506–514.

15. Chen L., and A. Moschitti. 2018. Learning to Progressively Recognize New Named Entities with Sequence to Sequence Models. In Proceedings of the 27th International Conference on Computational Linguistics, pp. 2181–2191.

16. Su S., N. Jia, X. Cheng, S. Zhu, and R. Li. 2018. Exploring Encoder-Decoder Model for Distant Supervised Relation Extraction. In IJCAI, pp. 4389–4395.

УДК 004.93'1 004.852

С.А. Пачганов

СОЗДАНИЕ МОДЕЛИ ДЛЯ РАСПОЗНАВАНИЯ ЗАПРЕЩЕННОЙ СИМВОЛИКИ НА ИЗОБРАЖЕНИЯХ ПРИ ОГРАНИЧЕННОЙ ОБУЧАЮЩЕЙ ВЫБОРКЕ

Аннотация

В работе представлен алгоритм, способный распознавать на изображениях запрещенную в Российской Федерации символику, а именно флаг Исламского Государства и свастику. Алгоритм представляет собой сверточную нейронную сеть с архитектурой MobileNetV2, реализованную с использованием пакетов Keras и Tensorflowдля языка программирования Python. Обучающая выборка содержала 400 изображений, содержащих запрещенную символику. Для улучшения результатов использовалась техника transferlearning: модель была предварительно обучена на данных ImageNet. Данные для обучения также включали изображения без распознаваемых объектов, для расширения выборки проводились случайные аугментации: повороты, отражения, коррекцияцвета, обрезания и растягивания.При этом как минимум один распознаваемый объект всегда оставался на изображении. Проведен анализ паттернов, распознаваемых сетью на изображениях, представленных в обучающей выборке и тестовом наборе, а также примеры, найденные в сети интернет с использование полученной модели.Точность распознавания (precision) составила 0.957 и 0.976, полнота (recall) 0.615 и 0.904 для свастики и флагаИсламского Государства соответственнона независимой тестовой выборке. В обсуждениях обозначены недостатки полученной модели и способы их устранения и улучшения точности классификации.

Ключевые слова: Компьютерное зрение, распознавание образов, аугментация, MobileNetV2, запрещенная символика, transferlearning, сверточные сети.

Введение

Миллионы изображений ежедневно размещаются в сети Интернет в различных социальных сетях по всему миру [1]. Часть этих изображений носит дискриминирующий, оскорбительный, призывающий к ненависти характер. Известно, что социальные сети используются террористами для пропаганды и вербовки людей [2, 3]. Человек, публикующий такие изображения, потенциально может быть вовлечен в террористическую деятельность. Правоохранительные структуры заинтересованы в мониторинге данной информации для предотвращения террористических атак. Ручной осмотр изображений – рутинный и малоэффективный способ решения поставленной задачей. Для автоматизации процесса могут быть применены методы машинного обучения и компьютерного зрения.

Для работы с изображениями с 2012 года успешно применяются сверточные нейронные сети [4]. Однако, использование нейронных сетей для решения задач компьютерного зрения предполагает наличие больших обучающих выборок (тысячи изображений для каждого класса). Одни из наиболее популярных сегодня архитектур нейронных сетей реализованы в популярных пакетах, используемых для реализации моделей машинного обучения.

Современные методы обучения позволяют бороться с недостаточными размерами тренировочной выборки. Метод transfer learning предполагает использование предобученной нейронной сети для обучения на новой задаче [5]. Эффективность такого подхода объясняется тем, что обученные ранее слои нейронной сети уже умеют выделять низкоуровневые признаки из изображений, после чего новая выборка необходима лишь для того, чтобы обучить сеть для вычленения высокоуровневых признаков и работы с полноразмерными объектами. При этом, чем ближе классы объектов предобученной сети и новой задачи, тем выше вероятность успешного применения такого метода.

Другим популярным методом является применение аугментаций к исходной обучающей выборке [6]. Так как сверточные нейронные сети предполагают подбор коэффициентов фильтров, используемых для свертки, то можно заметить, что для идентификации одного и того же объекта на зеркальном изображении требует как минимум двух разных фильтров.

Методика

Данные для обучения модели были собраны в сети Интернет, использовались следующие запросы в поисковых системах Яндекс и Google: ИГ, ИГИЛ, Исламское Государство, Флаг ИГ, террористы, теракт, ISIS, IslamicState, terrorism, наци, нацизм, фашизм, гитлер, националисты, СС, свастика, Nazi, fascism, Hitler, nationalism, swastikaи другие. Для набора данных отбирались изображения, содержащие искомые объекты: флаг ИГИЛ или свастику. Также отбирались похожие изображения, не содержащие искомые объекты для использования в качестве пустого класса. К набору с пустым классом также были добавлены многочисленные изображения из социальных сетей, не содержащие искомые объекты и отдаленные по своему наполнению.

Изображения размечались с использованием VGG Image Annotator [7], на изображения наносились bounding boxes, добавлялась аннотация. В таблице 1 представлена статистика итогового набора данных.

Таблица 1

Сводная статистика собранных данных

Выборка\ Набор	Пустой класс	Флаг ИГИЛ	Свастика
Обучающая	400	200	200
Тестовая	213	65	227

Для расширения выборки и улучшения качества классификатора был имплементирован генератор, применяющий случайные аугментации к изображениям, подаваемым в модель для обучения. Аугментации были реализованы с использованием пакета Tensorflow [8] со следующим требованием: при аугментации изображений с наличием запрещенной символики хотя бы один объект должен остаться на итоговом изображении. Список применяемых техник к изображений был следующим: поворот (tf.image. rot90), отражение справа налево (tf.image.flip_left_ right), сверху вниз (tf.image.flip_up_down), изменения яркости (tf.image.random_brightness), контраста (tf. image.random_contrast), насыщения (tf.image.random_ saturation), оттенка (tf.image.random_hue), разрешения (tf.image.random_jpeg_quality), добавление шума (0.01 tf.random_normal), размытия (scipy.ndimage. filters.gaussian_filter).

В качестве исходной модели была выбрана архитектура MobileNetV2 [9], которая была реализована с использованием пакета Keras [10]. Для инициализации использовались веса обученной на наборе ImageNet [11] модели. Взамен выходных слоев были добавлены слои: GlobalAveragePooling2D, Dense (512 нейронов, функция активации sigmoid), Dense (3 нейрона, функция активации softmax).

Пакет изображений (batch) состоял из 20 объектов, для обучения применялся оптимизатор RMSProp [12] с learning rate = 1e-4, в качестве функции потери использовалась категориальная кроссэнтропия. Процесс обучения происходил следующим образом:

1. Обучались последние 4 слоя сети в течение 20 эпох. Значение функции потерь составило 0.4042, а точности (accuracy) 0.8450.

2. Открывались последние 20 слоев сети для обучения, выставлялось новое значение learning rate= 5e-5. После обучения в течение 20 эпох значение функции потерь составило 0.2517, а точности (accuracy) 0.8916.

3. На последнем этапе обучались последние 40 слоев со значением learning rate = 5e-5. За 5 эпох было достигнуто значение функции 0.1025, а точности (accuracy) 0.9667.

В качестве метрик качества использовались точность (precision) и полнота (recall), рассчитанные по формулам (1) и (2):

$$Precision = \frac{TP}{TP+FP}, \quad (1)$$

$$Recall = \frac{TP}{TP+FN}, \quad (2)$$

где TP – истинно-положительные ответы (true positive), FP – ложноположительные (false positive), FN – ложноотрицательные (false negative).

Основная часть

Для исследования было накоплено 427и 265 изображений свастики и флага ИГИЛ соответственно. В обучении использовались по 200 фотографий из каждого набора, также для обучения был накоплен набор данных изображений без искомой символики в размере 400 изображений.

Целью работы было исследование возможности обучения сверточной нейронной сети для обнаружения запрещенной на территории РФ символики при малых размерах обучающей выборки. Для достижения цели применялись методы transfer-learning, аугментации выборки. За счет закрытия слоев сети для обучения сохранялись возможности предобученной модели к извлечению низкоуровневых признаков.

Для исследования возможностей полученной модели в сети Интернет в автоматическом режиме был собран большой набор данных из десятков тысяч изображений. В работе представлены изображения, на которых была обнаружена запрещенная символика, и описаны основные паттерны, которые использовала сеть для принятия решений.

Результаты и их обсуждение

Модель была обучена с использованием GPU Nvidia Tesla. В таблице 2 показаны полученные результаты измерения точности (precision) и полноты (recall) обученной модели на тестовой выборке, которая содержала 227 изображений со свастикой, 65 с флагом ИГИЛ и 213 изображений без запрещенной символики.

Таблица 2

Результаты обучения модели

Метрика\ Данные	Пустой класс	Флаг ИГИЛ	Свастика
Полнота (Recall)	0.954	0.904	0.615
Точность (Precision)	0.977	0.957	0.976

Из таблицы 2 видно, что обученная модель получилась сбалансированной и способна к правильному распознанию классов примерно в 9 из 10 случаев на тестовой выборке, изображения из которой не были использованы в обучении модели. Однако, можно

заметить низкое значение полноты в случае свастики, что может быть объяснено большим количеством похожих изображений со свастикой, что было связано с затруднительностью поиска разнородных изображений.

Для анализа возможностей модели был собран большой набор изображений из сети Интернет в автоматическом режиме, изображения которого были пропущены через полученную нейронную сеть. На рисунке 1 показаны некоторые примеры изображений, на которых были обнаружены запрещенные символы. При этом данные изображения отсутствовали в обучающей выборке.

На рисунке 1 помимо изображений показаны области, на которые реагируют выходные слои полученной модели, которые получены описанным в литературе методом [13].

Как видно из рисунка модель срабатывает на символику, однако помимо этого находит и другие паттерны, не связанные с символикой. В случае со свастикой часто встречаются срабатывания на большое скопление людей, в случае с флагом ИГИЛ на черный фон, людей с оружием и бородами, а также в камуфляжной одежде.

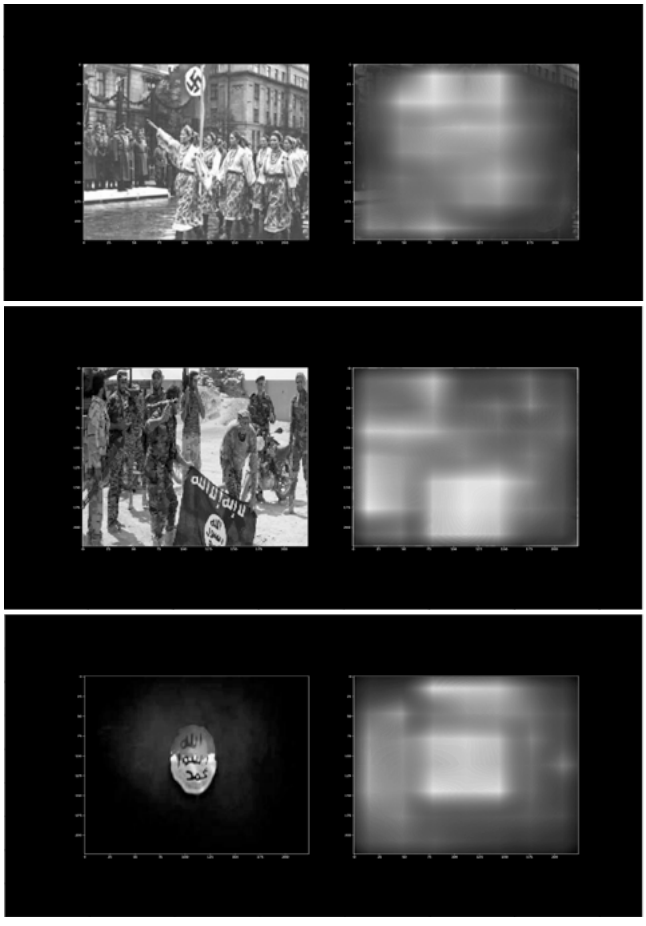


Рис. 1. Найденные в сети Интернет изображения и области этих изображений, которые были использованы моделью для принятия решения

Полученные результаты позволяют сделать вывод, что основная цель работы была достигнута, так как модель, обученная с использованием 800 изображений, способна идентифицировать запрещенную символику. Для улучшения точности модели могут быть использованы данные обнаруженные моделью. Описанный протокол обучения модели может быть использован для применения в других задачах с недостаточными по размеру наборами данных.

Выводы

В работе описан процесс обучения нейросетевой модели, способной распознавать на изображениях запрещенную в Российской Федерации символику, а именно флаг Исламского Государства и свастику. Модель представляет собой сверточную нейронную сеть с архитектурой MobileNetV2, реализованную с использованием пакетов Keras и Tensorflow. Обучающая выборка содержала 400 изображений, содержащих запрещенную символику. Для улучшения результатов использовалась техника transferlearning: модель была предварительно обучена на данных ImageNet. Данные для обучения также включали изображения без распознаваемых объектов, для расширения выборки проводились случайные аугментации: повороты, отражения, коррекция цвета, обрезания и растягивания. Продемонстрированы результаты анализа паттернов, распознаваемых сетью на изображениях, представленных в обучающей выборке и тестовом наборе, а также примеры, найденные в сети интернет с использование полученной модели. Точность распознавания (precision) составила 0.957 и 0.976, полнота (recall) 0.615 и 0.904 для свастики и флага Исламского Государства соответственнона независимой тестовой выборке. Для улучшения точности модели могут быть использованы данные обнаруженные моделью. Описанный протокол обучения модели может быть использован для применения в других задачах с недостаточными по размеру наборами данных.

Список литературы

1. Edwards J. Planet Selfie: We’re Now Posting a Staggering 1.8 Billion Photos Every Day //Business Insider. – 2014.

2. Klausen J. The Role of Social Networks in the Evolution of Al Qaeda-inspired Violent Extremism in the United States, 1990–2015. – Brandeis University, 2016.

3. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks //Advances in neural information processing systems. – 2012. – С. 1097–1105.

4. Lidberg L. Object Detection using deep learning and synthetic data. – 2018.

5. Pan S. J., Yang Q. A survey on transfer learning // IEEE Transactions on knowledge and data engineering. – 2010. – Т. 22. – №. 10. – С. 1345–1359.

6. Ding J. et al. Convolutional neural network with

data augmentation for SAR target recognition //IEEE Geoscience and remote sensing letters. – 2016. – T. 13. – № 3. – C. 364–368.

7. Dutta A., Gupta A., Zissermann A. VGG image annotator (VIA) //URL: http://www. robots. ox. ac. uk/~ vgg/software/via. – 2016.

8. Chen H.Y. et al. TensorFlow–A system for large-scale machine learning //OSDI. – 2016. – T. 16. – C. 265–283.

9. Sandler M. et al. Mobilenetv2: Inverted residuals and linear bottlenecks //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2018. – C. 4510–4520.

10. Keras: The Python Deep Learning Library. Date Views 01.02.2019 https://keras.io/

11. Deng J. et al. Imagenet: A large-scale hierarchical image database. – 2009.

12. Tieleman T., Hinton G. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude //COURSERA: Neural networks for machine learning. – 2012. – T. 4. – № 2. – C. 26–31.

13. Zhou B. et al. Learning deep features for discriminative localization //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – C. 2921–2929.

УДК 004.8

С.В. Смирнов

МЕТОДИКА РАЗРАБОТКИ НЕЧЕТКИХ КОГНИТИВНЫХ КАРТНА ОСНОВЕ ОНТОЛОГИЧЕСКОГО АНАЛИЗА ЭКСПЕРТНЫХ ДАННЫХ

Аннотация

Областью интереса данной работы является проблема идентификации популярных моделей слабоструктурированных ситуаций – когнитивных карт. Конкретно рассматривается построение нечетких когнитивных карт Коско на основе суждений экспертов, непосредственно участвующих в мониторинге и регулировании моделируемых ситуаций. В центре внимания предлагаемой методики – преодоление неполноты и противоречивости мнений у привлекаемых к разработке экспертов относительно состава значимых факторов предметной области и их взаимовлияния. В основу методики положены модели и методы, используемые на отдельных этапах онтологического анализа данных. Данные экспертизы формализуются как соответствие «упорядоченные пары факторов – влияния», где учитываются факторы ситуации, указанные хотя бы одним экспертом, и для каждого эксперта фиксируется его мнение о влиянии первого фактора на второй. Влияние предложено описывать лингвистической переменной, значения которой характеризуют как действительно указываемые экспертами оценки влияния (каждое из этих «действительных» оценок влияния – нечеткая переменная в универсуме [1, 1]), так и реалии коллективной экспертизы: отсутствие мнения эксперта для конкретной пары факторов. Взвешивание и интеграция мнений экспертов с помощью методов многозначной векторной логики позволяет сформировать нестрогое соответствие «упорядоченные пары факторов – влияния». Рациональное альфа-сечение нестрогих соответствий, разработанное в рамках онтологического анализа данных, обеспечивает получение непротиворечивого нечеткого соответствия «упорядоченные пары факторов – влияния» в рамках установленных ограничений существования различных оценок влияния для каждой пары факторов. Окончательно, агрегированный вес влияния дефаззифицируется известным методом центра тяжести.

Ключевые слова: нечеткая когнитивная карта, коллективная экспертиза, онтологический анализ данных, ограничение сосуществования свойств.

Введение

Нечеткие когнитивные карты (НКК; *fuzzy cognitive maps* – FCM) – результат эволюции формализма, предложенного Аксельродом [1]. Он одним из первых усмотрел возможность использования графов для описания восприятия людьми сложных слабоструктурированных ситуаций в политических и социальных системах. В НКК, введенных Коско [2], достигнуто сочетание этой идеи с нечеткой логикой и нейронными сетями. Такой подход сразу вызвал значительный научный интерес, и уже более 30 лет соответствующая методология моделирования сложных динамических систем активно развивается и широко применяется в различных прикладных областях, таких как компьютерные науки, инженерия, науки об окружающей среде, поведенческие науки, медицина, бизнес, информационные системы и информационные технологии (см., например, обзоры [3–6]). Есть основания рассматривать современные версии НКК в качестве платформы для объединения интеллектуального и когнитивного подходов в управлении сложными системами [7].

Большинство работ, связанных с НКК, сосредоточены на обсуждении моделей и методов анализа влияний по путям между вершинами-факторами, а также посвящены практическим приложением НКК. Вместе с тем известно, что использованию формальных вычислительных методов и технологий предшествует весьма трудоемкая экспертная работа по структурированию проблемных ситуаций (т.е. формирование системы факторов и связей между ними), параметризации системы факторов и связей, выбору функций влияния. Специалисты единодушно отмечают, что наряду с несомненными достижениями в развитии математического аппарата анализа этап описания представлений о слабоструктурированной динамической ситуации – этап построения НКК экспертами – слабо поддержан формальными моделями и методами [3–6, 8–11]. Особенно это касается вопросов увязки понимания разными экспертами анализируемой проблемной ситуации, когда создаваемая НКК является коллективным продуктом, и субъективное мнение одного эксперта другой эксперт может расценить как неверное или несущественное. Согласование знаний и мнений множества экспертов достигается при этом различными эвристическими и технологическими приемами, в большей или меньшей степени преодолевающими принятие решений простым большинством или волюнтаристским вердиктом лидера. Например, предлагают многократно повторять и резюмировать работу экспертов с понижением на каждом шаге значимости мнения тех из них, чьи оценки значительно отличаются от суждений большинства [12]. Лидерство пытаются заменить координацией работы и поддержкой неформальных действий выделенного координатора современными программно-техническими

средствами удаленного взаимодействия работников [13]. Стремятся реализовать интерсубъективное управление разработкой с присущим такому управлению дискурсом и стремлением к консенсусу [14, 15].

В [16] для обработки данных, представляющих субъективные мнения экспертов о структурировании проблемной ситуации в форме когнитивных карт, предложено использовать модели и методы онтологического анализа данных (ОАД) [17–19] и показано их использование на примере формирования знакового графа Аксельрода. В данной работе исследуется тот же вопрос применительно к построению более сложной модели – НКК.

Гипотетико-дедуктивная природа онтологического анализа данных

Рисунок 1 объясняет сопряжение в ОАД априорной работы субъекта по выдвижению гипотез об изучаемой предметной области (ПрО) и последующего выполнения дедуктивных процедур обработки данных измерений, ведущей к построению формальной онтологии.

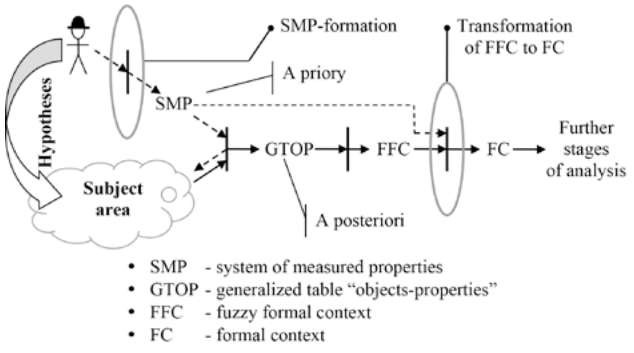


Рис. 1. Этапы онтологического анализа

Формируя систему измеряемых свойств (СИС) объектов ПрО, субъект, осознавая или нет, выдвигает гипотезы о понятийной структуре исследуемой ПрО. Сложность возникающей системы гипотетических понятий определяется качеством и количеством актов концептуального шкалирования [20] интересующих субъекта свойств объектов ПрО. При этом между подлежащими измерению свойствами образуются две и только две зависимости: несовместимость и обусловленность. Поэтому СИС – множество измеряемых свойств с двумя заданными на нем экзистенциональными отношениями. Модель СИС имеет многоуровневый рекурсивный характер и формируется путем применения ограниченного набора шаблонов [21].

В результате измерений в ПрО формируется обобщенная таблица «объекты-свойства» (ОТОС), содержащая, вообще говоря, неполные и противоречивые данные. Консолидация этих данных возможна лишь с помощью многозначных логик. Поэтому формальный контекст (ФК) ПрО – (G^*, M, I), где G^* – выборка объектов, M – множество измеряемых у объектов свойств, I – эмпирически выявляемое соответствие «объекты-свойства», – будет в итоге «мягким» (нестрогим, нечетким). Конструктивная интерпретация таких апостериорных данных возможна лишь после трансформации «мягкого» ФК в однозначный. Трансформация будет корректной лишь при учете априорно зафиксированных в СИС ограничений существования свойств [18, 19].

Формирование нечеткой когнитивной карты

НКК может быть описана кортежем (F, W), где $F = (f_i)_{i=1,\ldots,n}$ – множество вершин-факторов ситуации, $W = (w_{ij})_{i=1,\ldots,n;\, j=1,\ldots,n}$ – матрица каузальных связей графа, задающая одновременно смежность вершин и степень влияния вершин друг на друга.

Например, степень влияния может описываться экспертами лингвистически с использованием значений из множества T = {“очень сильно негативно (*nvs*)”, “сильно негативно (*ns*)”, “негативно (*nm*)”, “слабо негативно (*nw*)”, “не влияет (*z*)”, “слабо позитивно (*pw*)”, “позитивно (*pm*)”, “сильно позитивно (*ps*)”, “очень сильно позитивно (*pvs*)”}, где каждое значение – нечеткая переменная в универсуме [-1, 1] со свой функцией принадлежности [12].

Рассматривая задачу построения НКК в терминах ОАД, легко установить следующее:

- в качестве объектов интересующей нас ПрО выступают упорядоченные пары факторов $P = \{(f_i, f_j)\}$, где $f_i, f_j \in F$, $i \neq j$;

- актуальными семантическими суждениями в нашей ПрО будут независимые заключения экспертов о степени влияния i-го фактора на j-й.

- СИС задачи в соответствии с принципами построения канонической НКК имеет вид, показанный на рисунке 2.

Данные, поступившие от экспертов, составят ОТОС, структуру которой иллюстрирует таблица 1. Символ **X** указывает на наличие, а **None** – на отсутствие у объекта того или иного свойства (полагаем, что рациональный эксперт в каждой строке своего «ответа» отметит символом X только одну позицию, соответственно отмечая символом **None** все остальные, но в ОАД подобные ошибки, вообще говоря, не катастрофичны). **Символ NM** (*not measured*) указы-

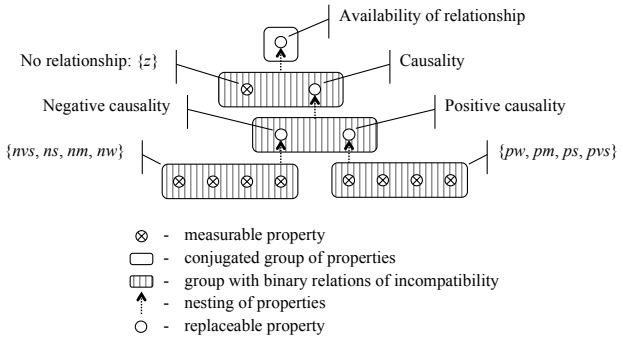


Рис. 2. Модель системы измеряемых свойств в задаче построения нечеткой когнитивной карты

Таблица 1

Данные экспертизы влияний факторов ситуации

	Expert 1				…	Expert *k*			
	nvs	*ns*	…	*pvs*		*nvs*	*ns*	…	*pvs*
(f_1, f_2)	**None**	**X**		**None**	…	**NM**	**NM**		**NM**
(f_1, f_3)	**F**	**F**		**F**		**NM**	**NM**		**NM**
…									
(f_n, f_{n-1})	**None**	**None**		**X**		**F**	**F**		**F**

вает, что экспертом оценка соответствующего влияния не проводилась (например, потому, что в его оценке ситуации отсутствует один или оба описываемых в строке ОТОС фактора влияния) и занимает, естественно, все позиции «ответа» эксперта. Символ **F** (*failure*) фиксирует наблюдаемый на практике случай сомнения и отказа эксперта от определенной оценки влияния *i*-го фактора на *j*-й, и для СИС данной задачи должна заполнять весь «ответ» эксперта.

Для консолидации неполных и противоречивых данных экспертизы, собранных в ОТОС, необходимым балансом качеств обладает многозначная векторная логика V^{TF}, где истинность отдельного семантического суждения о ПрО оцениваться вектором (Истина, Ложь) [22]:

$$b_{ij} = (b^+_{ij}; b^-_{ij}); b^+_{ij}, b^-_{ij} \in [0, 1]. \quad (1)$$

Нестрогое соответствие «упорядоченная пара факторов – влияние по мнению разных экспертов» получается путем преобразования ОТОС, когда смысловой интерпретацией значений **X**, **None**, **NM** и **F** служат истинностные константы V^{TF}-логики (1;0) – «Истина», (0;1) – «Ложь», (0;0) – «Неопределенность», (0,5;0,5) – «Недоопределенность» соответственно. Совмещение векторных оценок истинности интересующих нас семантических суждений о ПрО окончательно определит формальный контекст задачи построения НКК по результатам коллективной экспертизы: (P, T, I), где $I \subseteq P \times T$ – нестрогое соответствие «упорядоченная пара факторов – влияние».

Стандартное альфа-сечение $I^{(\alpha)}$ – однозначное (строгое) соответствие «упорядоченная пара факторов – влияние» при пороге доверия к данным экспертизы $\alpha = (\alpha^+, \alpha^-)$, $\alpha^+, \alpha^- \in [0, 1]$, – в общем случае не будет корректным из-за игнорирования экзистенциональных отношений в СИС. Поэтому в ОАД разработан метод рационального альфа-сечения «мягких» ФК, доставляющий альфа-сечение $\check{I}^{(\alpha)}$, в котором учтены ограничения существования свойств у объектов исследуемой ПрО. Именно при формировании $\check{I}^{(\alpha)}$ происходит дискриминация тех семантических суждений, которые, отражая мнение некоторой части экспертов, противоречат другим имеющимся суждениям с большими оценками истинности. Но подобная дискриминация может возникнуть лишь локально, для суждений относительно отдельно взятого объекта $(f_i, f_j) \in P$.

В задаче построения НКК этот метод ОАД применяется в особом режиме: не учитывается несовместимых «листовых» свойств модели СИС (рис. 2), т.к. на последующих этапах выявления степени влияния между факторами имеется возможность интеграции всех степеней позитивного или негативного влияния.

Кроме того, рассматриваемый метод модернизируется так, чтобы вместо однозначного отношения $\check{I}^{(\alpha)}$ сформировать нечеткое соответствие $\tilde{I}^{(\alpha)}$:

- отсекаемые векторные оценки трансформируемого соответствия I как обычно заменяются нулями;

- векторные оценки $(b^+_{ij}; b^-_{ij})$, которые по логике альфа-сечения должны были бы замещены единицей, сохраняются, а затем замещаются нечеткой оценкой

$$v_{ij} = \frac{b_{ij}^{\ +} - b_{ij}^{\ -} + 1}{\sqrt{2}} \quad (2)$$

(этот вполне обоснованный прием замещения ведет, тем не менее, к определенной потере информации, т.к. функция $(v_{ij} \rightarrow (b^+_{ij}; b^-_{ij}))^{-1}$ многозначна).

В получаемом контексте $(P, T, \tilde{I}^{(\alpha)})$ могут быть обнаружены пустые строки для всех объектов (f_i, f_j) и (f_j, f_i), i=1,…, n. Это означает, что при выбранном пороге доверия к данным экспертизы фактор f_j оказывается незначимым для структурирования проблемной ситуации, и указанные объекты должны быть исключены из контекста задачи.

Контекст $(P^-, T, \tilde{I}^{(\alpha)})$, где P^- – редуцированное указанным выше способом множество пар значимых факторов, содержит все сведения для формирования НКК (F, W):

- множество F легко извлекается из P^-;

- редукция соответствия $\tilde{I}^{(\alpha)}$ до двух столбцов, когда позитивные и, отдельно, негативные оценки влияния различной степени дефаззифицируется известным методом центра тяжести, определяет все необходимые данные для построения матрицы W.

Выводы

К достоинствам предложенного подхода к построению НКК следует отнести простоту алгоритмизации, формирование результата за один проход (по единовременно полученным данным экспертизы), отсутствие какой-либо искусственной маргинализации отдельных экспертов (дискриминация мнений пол-

ностью локализована в задаче оценки влияния пары конкретных факторов).

Следует также отметить, что ОТОС, представляющая данные экспертизы, может иметь более богатую структуру, чем было описано. Дополнительно в нем могут быть зафиксированы и обработаны априорные степени доверия к отдельным экспертам, независимые серии оценок одного и того же эксперта (возникающие, например, в результате нескольких его «подходов» к задаче), причем при различном количестве серий «измерений» у каждого эксперта и отличающейся достоверности данных, фиксируемых в каждой серии [18, 19].

Развитие предложенной методики для НКК, допускающих не только позитивные, но и негативные влияния одного фактора на другой [2], представляется возможным на основе формального описания такой НКК, предложенный Силовым [23]. Построение адекватной СИС должна опираться на имеющие в этой модели ограничения на заполнение клеток матрицы влияний $W^{\pm}$ [3, 23].

Список литературы

1. Axelrod R., 1976. Structure of Decision: The Cognitive Maps of Political Elites. Princeton University Press, 395 p.

2. Kosko B., 1986. Fuzzy cognitive maps. International Journal of Man-Machine Studies, 24(1): 65–75.

3. Papageorgiou E. and J. Salmeron, 2013. A Review of Fuzzy Cognitive Map Research at the Last Decade. IEEE Transactions on Fuzzy Systems, 21(1): 66–79.

4. Glykas M., 2010. Fuzzy Cognitive Maps: Advances in Theory, Methodologies, Tools and Applications (Introduction), vol. 247. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, pp. IX–XV.

5. Кузнецов О.П., А.А. Кулинич и А.В. Марковский, 2006. Анализ влияний при управлении слабоструктурированными ситуациями на основе когнитивных карт. В кн.: Н.А. Абрамова, К.С. Гинсберг, Д.А. Новиков (ред.) Человеческий фактор в управлении. Москва: КомКнига, с. 313–344.

6. Aguilar J., 2005. A survey about fuzzy cognitive maps papers (Invited paper), Int. J. Computational Cognition, 3(27): 27–33.

7. Groumpos P.P., 2018. Intelligence and Fuzzy Cognitive Maps: Scientific Issues, Challenges and Opportunities. Studies in Informatics and Control, 27(3): 247264.

8. Кулинич А.А., 2010. Компьютерные системы моделирования когнитивных карт: подходы и методы. Проблемы управления, #3: 2–16.

9. Stach V., L. Kurgan and W. Pedrycz, 2010. Expert-Based and Computational Methods for Developing Fuzzy Cognitive Maps. In Glykas M. (Eds) Advances in Theory, Methodologies, Tools and Applications, vol. 247. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, pp. 23–41.

10. Абрамова Н.А. и С.В. Коврига, 2008. Некоторые критерии достоверности моделей на основе когнитивных карт. Проблемы управления, #6: 23–33.

11. Schneider M., E. Shnaider, A. Kandel, and G. Chew, 1998. Automatic construction of FCMs. Fuzzy Sets Syst., 93(2): 161–172.

12. Stylios D.C. and P.P. Groumpos, 2004. Modeling Complex Systems Using Fuzzy Cognitive Maps. IEEE Transaction on Systems, Man and Cybernetics, Part A: Systems and Humans, 34: 155–162.

13. Коростелёв Д.А., Д.Г. Лагерев и А.Г. Подвесовский, 2008. Система поддержки принятия решений на основе нечетких когнитивных моделей “ИГЛА”. В трудах 11-й нац. конф. по искусственному интеллекту с международным участием КИИ-2008, Дубна, Россия, 28 сентября–03 октября. Т. 3. Москва, ЛЕНАНД, с. 329336.

14. Vittikh V.A., 2015. Introduction to the Theory of Intersubjective Management. Group Decision and Negotiation, 24(1): 67–95.

15. Виттих В.А., Т.В. Моисеева и П.О. Скобелев, 2013. Принятие решений на основе консенсуса с применением мультиагентных технологий. Онтология проектирования, #2: 20–25.

16. Смирнов С.В., 2017. Модели и методы формирования когнитивных карт при их коллективной разработке. В трудах Шестой междунар. науч. конф. «Информационные технологии и системы ИТиС-2017», Банное, Россия, 1–5 марта, Ред.: Ю.С. Попков, А.В. Мельников. – Челябинский гос. ун-т, с. 281–283.

17. Смирнов С.В., 2014. Многозначная и нечеткая логики в онтологическом анализе данных. В трудах Третьей междунар. науч. конф. «Информационные технологии и системы ИТиС-2014», Банное, Россия, 26 февраля – 2 марта, Ред.: Ю.С. Попков, А.В. Мельников. Челябинский гос. ун-т, с. 90-91.

18. Коварцев А.Н., В.С. Смирнов и С.В. Смирнов, 2015. Интеллектуализация формирования контекста для вывода понятийной структуры предметной области. В трудах Четвертой междунар. науч. конф. «Информационные технологии и системы ИТиС-2015», Банное, Россия, 25 февраля – 1 марта, Ред.: Ю.С. Попков, А.В. Мельников. Челябинский гос. ун-т, с. 82-83.

19. Самойлов Д.Е., В.А. Семенова и С.В. Смирнов, 2016. Анализ неполных данных в задачах построения формальных онтологий. Онтология проектирования, 6(3): 317–339.

20. Samoilov D.E., V.A. Semenova and S.V. Smirnov, 2018. Multilevel recursive model of properties existence constraints in machine learning. Journal of Physics: Conf. Series 1096 (2018) 012096.

21. Ganter, B. and R. Wille, 1989. Conceptual scaling. / In: F. Roberts (Ed.): Applications of Combinatorics and Graph Theory to the Biological and Social Sciences. – New York Springer-Verlag, pp. 139–167.

22. Аршинский Л.В., 2007. Векторные логики: основания, концепции, модели. Иркутский гос. ун-т, 228 с.

23. Силов, В.Б., 1995. Принятие стратегических решений в нечеткой обстановке. Москва, ИНПРОРЕС, 228 с.

УДК 004.451.2

А.В. Тузов

ИСПОЛЬЗОВАНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ МЕТОДОВ ПРОГНОЗИРОВАНИЯ ЗАГРУЖЕННОСТИ ПРОЦЕССОРА В ВЫСОКОНАГРУЖЕННОЙ WEB СИСТЕМЕ

Аннотация

В данной статье рассматривается актуальная проблема планирования ресурсов, в частности проблема планирования процессорного времени в высоконагруженных информационных системах. Цель данной работы проверить гипотезу, о возможности предсказания загруженности процессора в информационной системе с помощью внешних параметров среды таких как температура воздуха, наличие осадков, уровень пробок и прочее. Для проверки была создана система, в которой работали корпоративный сайт компании, система мониторинга ресурсов и общедоступное приложение для vk.com. На наш взгляд указанные внешние параметры в той или иной степени могут оказывать влияние на загруженность системы. Например, при более низких температурах воздуха нагрузка возрастает, а при более высоких уменьшается. Данные о загруженности процессора собирались в процентном соотношении в течение полугода с интервалом в пять минут. Внешние параметры были получены с таких сайтов как gismeteo.ru, яндекс погода и прочие. Для идентификации модели были выбраны несколько методов машинного обучения – линейная регрессия, метод опорных векторов и случайный лес. Все собранные параметры были отфильтрованы – проверены на наличие кросскореляции и нормализованы. Используя построенные модели, мы предсказали процентный расход процессора в высоконагруженной информационной системе. Для каждой модели было посчитано среднеквадратичное отклонение, по результатам которого эти модели были сравнены между собой. Анализ показал, что каждая из полученных моделей имеет ряд проблем. В качестве рекомендаций по улучшению результатов, указано увеличение качества и количества собираемых данных, а так же количество параметров на основании которых строится модель.

Ключевые слова: Машинное обучение, линейная регрессия, методом опорных векторов, случайный лес, регрессия.

Введение

В последние годы использование машинного обучения для предсказания и управления ресурсами систем является одной из наиболее быстро развивающихся отраслей. Предсказание расхода электроэнергии для домов, электростанций или облачных систем является наиболее популярными областями применения. Мы будем использовать машинное обучение на стыке с другой быстрорастущей отраслью – высоконагруженные информационные системы, в частности информационной web-системой. Для подобных систем исследования по предсказанию расхода ресурсами являются не частыми явлением, однако такие работы есть [1, 2]. Имеются исследования по использованию новых политик управления в системе [3], а так же использование энергонезависимых процессоров [4]. Использование внешних факторов, в которых функционирует системы, например погоды, характерно только для энергетической отрасли [5, 6, 7]. Поэтому в этой работе мы решили проверить гипотезу о том, что нагрузка на процессоры высоконагруженных информационных систем зависит от таких внешних параметров как температуры воздуха, уровня пробок и прочее. Если данная гипотеза окажется верна, то грамотное управление ресурсами положительно скажется на работе системы. В частности можно значительно сэкономить на системе охлаждения [1], а так же использовать для планирования стека задач. Для проверки гипотезы, мы будем использовать различные методы машинного обучения, такие как линейная регрессия, случайный лес и метод опорных векторов. Линейная регрессия, как самая простой и часто используемый метод. Случайный лес потому, что в некоторых работах, по предсказанию расхода ресурсов в облачных системах, он показывает самые хорошие результаты. И метод опорных векторов мы используем потому, что именно этот алгоритм показывает хорошие результаты во многих задачах регрессии.

Методика

В данной работе для построения модели мы будем использовать такие методы машинного обучения как линейная регрессия, случайный лес и метод опорных векторов.

Линейная модель [8–10] имеет вид:

$$y=Xb$$

Где $y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}$ – вектор значений зависимой переменной y, где y – значение процент загруженности процессора в системе.

$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix}$ – матрица внешних параметров.

Количество выбранных нами параметров (k) составляет 33. А общий объем данных (n), которые

мы собирали в течение четырех месяцев, получился 15 000.

Метод опорных векторов [11–13]. Подробное описание данного метода является слишком объемным и выходит за рамки данной статьи.

Случайный лес [14–16] – (random forest) это множество решающих деревьев. Так как в нашем случае решается задача регрессии их ответы усредняются. Алгоритм построения случайного леса, состоящего из N деревьев, выглядит следующим образом:

Для каждого $n = 1,\ldots, N$:

- Сгенерировать выборку X_n с помощью бутстрэпа;
- Построить решающее дерево b_n по выборке :
  - по заданному критерию мы выбираем лучший признак, делаем разбиение в дереве по нему и так до исчерпания выборки
  - дерево строится, пока в каждом листе не более объектов или пока не достигнем определенной высоты дерева
  - при каждом разбиении сначала выбирается m случайных признаков из исходных, и оптимальное разделение выборки ищется только среди них.

Итоговый классификатор $a(x) = \frac{1}{N}\sum_{i=1}^{N} b_i(x)$, т.е. для задачи классификации, мы выбираем решение голосованием по большинству, а в задаче регрессии – средним. Рекомендуется в задачах классификации брать m=$\sqrt{n}$, а в задачах регрессии – $m = \frac{n}{3}$, где n – число признаков. Также рекомендуется в задачах классификации строить каждое дерево до тех пор, пока в каждом листе не окажется по одному объекту, а в задачах регрессии – пока в каждом листе не окажется по пять объектов.

Все выбранные нами параметры для построения модели в той или иной степени оказывают влияние на высоконагруженную информационную систему. Например:

- Чем ниже температура воздуха, тем больше нагрузка. Если температура воздуха в прошлый день в это же время была выше, то на текущий момент нагрузка будет больше.
- При отклонении атмосферного давления от нормы, большее количество людей испытывают недомогание и проводят меньше времени за компьютерами и смартфонами.
- Если уровень пробок повышается, то больше людей проводят время в интернете, стараясь себя занять.
- Скорость ветра так же оказывает небольшое влияние на нагрузку. Ведь чем выше скорость ветра, тем больший дискомфорт доставляет прогулка по улице или занятия на воздухе. Соответственно, нагрузка на информационные системы возрастает.

Для каждого выбранного нами метода было посчитано среднеквадратичное отклонение по формуле:

$$A = \sqrt{\frac{\sum_{i=0}^{n}(y1_i - y2_i)^2}{n}} \quad (1)$$

где, $y1_i$ – предсказанное значение загруженности процессора, $y2_i$ – реальная загруженность процессора в процентах, n – размер собранной выборки.

Основная часть

Операционная система (ОС) – это набор инструментов, которые позволяют пользователю взаимодействовать с «железом». Одной из главных задач ОС является управление ресурсами, причем ресурсы может быть как физический объект – процессор, оперативная память, дисковые накопители, так и логический объект, который существует только в пределах самой ОС, например таблица выполняемых процессов или сетевых подключений [17]. При этом нужно понимать, что все физические ресурсы ограничены и в условиях высоконагруженных систем это является очень важной проблемой.

Для ее решения, есть несколько способов, один из которых – грамотных выбранный алгоритм работы процессора. Помимо простейших и базовых алгоритмов в системах, которые решают мультиклассовые задачи, например облачные сервисы и системы рендеринга, используются такие алгоритмы как дискриминационное совместное использование процессора (DPS), дискриминационное обслуживание в произвольном порядке (DROS), обобщенное совместное использование процессора (GPS), взвешенная справедливая организация очередей (WFQ), мультиклассовое совместное использование процессора (MPS), мультиклассовая служба случайного порядка (MROS) [3].

Однако, несмотря на наличие различных алгоритмов распределения процессорного времени между процессами, полученных аналитическим способом, существует ряд проектов, которые пытаются использовать машинного обучение для предсказания загруженности процессора и управления процессорным временем. Чаще всего такие исследования проводятся для крупных облачных решений. В качестве данных на основе которых мы будем строить свои модели, выступают внешний факторы в которых функционирует информационная система. Если результаты работы окажутся приемлемыми, то это покажет зависимость загруженности процессора в таких системах, от перечисленных в предыдущем разделе параметров.

Результаты экспериментальных исследований

Для проведения эксперимента был собран сервер linux ubuntu 14.04.5 LTS. Процессор Intel(R) Xeon(R) CPU X5650 2.67GHz. 16 гигабайта оперативной памяти, 105 гигабайт жесткий диск. В системе работают сервисы Apache/2.4.7 и mysql(Server version: 5.5.62-38.14-log Percona Server (GPL), Release 38.14) для поддержки корпоративного сайта компании и приложения для vk.com и zabbix, который следит за ресурсами системы.

На сервере был запущен скрипт, который в течение каждых десяти минут обращался к сайтам:

• yandex.ru, на котором собирались температура-воздуха по данным, температура воздуха по ощущениям, какая температура была вчера в это же время, скорость ветра, атмосферное давление и влажность.

• с сайтов transport.74.ru и www.ekburg.ru мы получали текущий уровень пробок на дорогах в Челябинск и Екатеринбурге соответственно.

Учитывая сложившиеся трудности и замечания в предыдущей работе, собираемые параметры были тщательно отобраны. В связи с тем, что нагрузка на информационную систему оказывалась только из определенных населенных пунктов, то и внешние данные брались только из них. В частности, это такие города как:

- Челябинск
- Екатеринбург
- Магнитогорск
- Москва
- Санкт-Петербург.

Скрипт по сбору данных работал в течение четырех месяцев, и общий объем собираемых данных составил 15 000 строк, которые записаны в csv файл. При некоторых попытках получить информацию с сайтов происходили ошибки, в таком случае значение параметра вычислялось как среднее между предыдущим и следующим значением. Итоговые строки в файле имели вид чисел разделенные спецсимволом, где последнее значение в строке – загруженность процессора в процентах.

Собранные данные были проверены на наличие кросскореляции относительно друг друга с использованием критерия Пирсона. Анализ показал, что корреляция все выбранных нами параметров не превышает средний уровень корреляции, что позволяет нам использовать всю полученную информацию для построения математической модели.

Получившийся файл мы разделили на два – обучающая, 10 000 записей, и проверочная выборка, 5 000 записей. После чего из-за того, что некоторые значения параметров меньше единицы, а другие значения превышают тысячу, была произведена нормализация данных, через логарифмическое преобразование. Для построения математической модели были выбраны линейная регрессия, метод опорных векторов и случайный лес. Данные собранные о загруженности процессора собирались в процентах, поэтому и среднеквадратичное отклонение (1) считалось в процентах.

При использовании линейной регрессии наилучший достигнутый результат среднеквадратичного отклонения (1) получился 3,32.

Построение модели с помощью метода опорных векторов показал себя гораздо лучше. Среднеквадратичное отклонение при данном методе составило 1,01

Наилучший результат среднеквадратичного отклонения при использовании случайного леса для построения математической модели составил 2,18.

Анализ полученных результатов показывает, что наихудший метод для построения модели на основе внешний данных в которых функционировала высоконагруженная информационная система, является линейная регрессия. Чуть лучше себя показал случайный лес и лишь метод опорных векторов показал хорошие результаты. Однако самой главной проблемой является то, что построенная модель не применима к другой системе работающей в других внешних условиях. Если нагрузка на наш сервер шла из указанных выше городов, то для системы охватывающей другие города, страны количество параметров надо подбирать соответствующие.

Выводы

За последние годы машинное обучение активно развивается, входит в различные сферы нашей жизни и используется для различных прикладных задач. В данной работе мы использовали методы машинного обучения на стыке с другой бурно развивающейся отраслью – ВС, так как в таких системах существует ряд проблем связанных с расходом ресурсов. Была проверена гипотеза о том, что загруженность процессора в информационной системе зависит от внешних параметров окружающей среды. Для проверки мы использовали такие методы машинного обучения как линейная регрессия, метод опорных векторов и случайный лес. По результатам работы методов были построены математические модели, посчитано среднеквадратичное отклонение. Наилучший результат показал метод опорных векторов со значением 1,01, тогда как у линейной регрессии он составил 3,32, а у случайного леса 2,18. Однако анализ результатов показывает, что построенная модель не универсальна и для каждой информационной системы придется выбирать свои параметры. Все замечания к предыдущей работе были учтены, что значительно сказалось на итоговом результате данной работы.

Список литературы

1. R. Evans and J. Gao. DeepMind AI Reduces Google Data Centre Cooling Bill by 40%. // DeepMind blog(2016), vol. 20.

2. Yang, Renyu, et al. Intelligent Resource Scheduling at Scale: a Machine Learning Perspective.// Service-Oriented System Engineering (SOSE), 2018 IEEE Symposium on. IEEE, 2018.

3. Avrachenkov K., Bodas T. On the equivalence between multiclass processor sharing and random order scheduling policies //ACM SIGMETRICS Performance Evaluation Review, 2018, vol. 45, no. 4, pp. 2–6.

4. Li, Jinyang, et al. PATH: Performance-Aware Task Scheduling for Energy-Harvesting Nonvolatile Processors. // IEEE Transactions on Very Large Scale Integration (VLSI) Systems (2018).

5. H. Zheng, J. Yuan, L. Chen. [Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a

Xgboost Algorithm for Feature Importance Evaluation]. Energies, 2017, vol 10, no. 8, 1168 p.

6. Divina F., Gilson, A., Goméz-Vela, F., García Torres, M., & Torres, J. F. [Stacking Ensemble Learning for Short-Term Electricity Consumption Forecasting]. Energies, 2018, vol 11, no. 4, 949p

7. Dahua Gan, Yi Wang, Ning Zhang, Wenjun Zhu. [Enhancing short-term prob-abilistic residential load forecasting with quantile long–short-term memory]. The Journal of Engineering, 2017, 6p.

8. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Издательский дом «Вильямс». 2007. 912 с.

9. Стрижов В.В. Методы индуктивного порождения регрессионных моделей. М.: ВЦ РАН. 2008. 55 с.

10. Дж. Себер. Линейный регрессионный анализ. – М.: Мир, 1980, 456 с.

11. Fine S., Scheinberg K. INCAS: An incremental active set method for SVM: Tech. rep.: 2002.

12. Osuna E., Freund R., Girosi F. An improved training algorithm for support vector machines // Neural Networks for Signal Processing VII. IEEE Workshop, 1997, pp. 276–285.

13. Bartlett P., Shawe-Taylor J. Generalization performance of support vector machines and other pattern classifiers // Advances in Kernel Methods. — MIT Press, Cambridge, USA, 1998.

14. A. Liaw, M. Wiener. Classification and Regression by randomForest //R News (2002) Vol. 2/3 p. 18.

15. A. Natekin, A. Knoll Gradient boosting machines, a tutorial // FronNeurorobot. 2013.

16. [16] И. Генрихов. О критериях ветвления, используемых при синтезе решающих деревьев // Машинное обучение и анализ данных, 2014, vol. 1, № 8, p. 988–1017.

УДК 553.98.04

А.В. Якимчук

РЕАЛИЗАЦИЯ МОДЕЛИ ОЦЕНКИ И ПРОГНОЗИРОВАНИЯ РИСКОВ СОСТОЯНИЯ ЛЕСНОГО ФОНДА НЕФТЕДОБЫВАЮЩЕГО РЕГИОНА

Аннотация

В статье рассматривается применение нейронных сетей для реализации модели оценки и прогнозирования состояния лесного фонда на примере Ханты-Мансийского автономного округа – Югры. В качестве сведений используется накопленный массив данных Природнадзором Югры о ранее случившихся инцидентах на территории округа в период с 2010–2018 гг. В основе реализованной модели лежат методы машинного обучения, которые позволяют получить вещественное число, являющееся степенью риска возникновения повторного инцидента на территории определенного участка. Для построения тепловой карты используются методы геоинформационного анализа.

Ключевые слова: машинное обучение, прогнозирование рисков, keras, нейронные сети, данные о нефтеразливах, лесные вырубки, риск-ориентированный подход, ГИС.

Введение

С каждым годом проблема контролирования состояния лесного фонда, связанная с нефтеразливами, незаконными лесными вырубками и загрязнениями, приобретает всё более серьёзные масштабы [1]. Так, в период за 2017–2018 гг. в России произошло более 100 тысяч крупных и мелких инцидентов [2]. Выявлением инцидентов занимаются органы федерального государственного экологического надзора и надзора в области промышленной безопасности, однако количество инспекторов ограничено, и невозможно проверить абсолютно все подконтрольные участки.

В связи с этим всё более активно в практике контрольно-надзорной деятельности начинает применяться риск-ориентированный подход. Данный подход основан на методологии анализа риска и позволяет оптимизировать методы и частоту проверок со стороны регуляторов с учетом степени риска контролируемых объектов.

Рассмотрим несколько работ, написанных ранее по данной проблеме. В работе «Методика расчета экологических рисков» [3] представлена методика, учитывающая все возможные последствия наступления риска. Также проанализирована информационная основа оценки экологических рисков и необходимость классификации факторов экологической опасности. Приведена структура полного экологического ущерба. Рассмотрены методы оценки ущерба, как для окружающей природной среды, так и для здоровья человека. Показана связь экологического ущерба с социальным. Приведены основные принципы природоохранной политики. Показана зависимость выбора метода количественной оценки ущерба для здоровья человека от продолжительности и уровня интенсивности воздействия. Рассмотрены экономические оценки экологических воздействий. Приведены основные элементы процесса управления рисками.

В работе [4] авторами предлагается новый подход к организации процесса анализа и мониторинга экологических рисков в условиях нечеткости, в основе которого лежит качественный анализ экологических рисков, выполняемый экспертными методами с использованием ГИС технологий.

В статье «Принятие решений в геоинформационных системах» Р.Г. Болбаков рассматривает методы и технологи принятия решений при помощи геоинформационных систем. Также в работе показана возможность многоаспектного применения системного анализа при управлении с помощью ГИС. Показан информационный канал управления с помощью ГИС. Управление с помощью ГИС осуществляется в цифровом и визуальном варианте.

В работе «Эколого-ориентированные математические модели управления рисками и обеспечением безопасности» [6] анализируются различные варианты применения методов теории игр, линейного и нелинейного программирования для принятия оптимальных эколого-ориентированных решений в сфере управления рисками и обеспечением комплексной безопасности. Представлены примеры математических моделей оценки качества продовольствия, минимизации его вредного воздействия на человека, а также оптимизации функционирования биоэнергетики, основанной на использовании энергетических лесов. Сформулированы рекомендации по внедрению и развитию разработанных эколого-ориентированных методик в интересах рационального природопользования, ресурсосбережения, обеспечения безопасности человека и природно-техногенных комплексов.

Приведенный выше анализ исследований по применению риск-ориентированного подхода в экологической сфере показал, что в настоящее время используются различные подходы для определения вероятности возникновения рисков, и каждый способен решить определенно конкретную задачу. В связи с этим предлагается применить новый подход с использованием современных методов машинного обучения для прогнозирования участков наиболее подверженных риску на основе исторических и геоинформационных данных.

Для разработки модели оценивания и прогнозирования экологических рисков будут использованы данные с территории Ханты-Мансийского автономного округа – Югры.

Исходные данные и их анализ

В информационных системах, применяемыми различными ведомствами Ханты-Мансийского автономного округа – Югры, накоплены большие объемы разнородной информации об имеющихся техногенных объектах нефтегазового сектора, об экологическом состоянии территорий, о выявленных экологических нарушениях и др. Одним из главных источников такой информации является государственная информационная система «Эконадзор», которая используется Службой по контролю и надзору в сфере охраны окружающей среды, объектов животного мира и лесных отношений Ханты-Мансийского автономного округа –Югры (Природнадзор Югры) [7]. В данную систему инспекторы Природнадзора Югры записывают информацию по выявленным в ходе обследования территории инцидентам, связанным с лесными пожарами, незаконными лесными рубками, самовольным занятием и использованием лесной территории, незаконной транспортировкой древесины, незаконными свалками, загрязнениями земель и водных ресурсов нефтепродуктами и их производными.

Картографическая информация о природных и техногенных объектах, в том числе полученная по результатам дешифрирования космических снимков, содержится в информационных ресурсах Центра космических услуг Ханты-Мансийского автономного округа – Югры (ЦКУ), созданного на базе автономного учреждения Ханты-Мансийского автономного округа – Югры «Югорский научно-исследовательский институт информационных технологий». ЦКУ является региональным центром компетенций в области использования данных дистанционного зондирования Земли и других результатов космической деятельности, и активно взаимодействует с правоохранительными и контрольно-надзорными органами автономного округа в сфере выявления экологических нарушений, что позволило ему аккумулировать значительные объемы пространственной и атрибутивной информации о состоянии окружающей среды в регионе.

Анализ имеющихся в различных информационных ресурсах данных выявил основные факторы, учет которых позволил сделать оценку рисков изменения состояния лесного фонда [8–12]. Для рисков загрязнения территории лесного фонда нефтью и нефтепродуктами основными факторами являются следующие:

- наличие техногенных объектов;
- наличие объектов размещения отходов;
- приближенность к населенным пунктам;
- наличие транспортной инфраструктуры;
- наличие заявок на использование лесных участков;
- наличие нарушений правил санитарной безопасности;
- наличие неисполненных предписаний;
- сообщения об авариях и инцидентах;
- результаты дешифрирования загрязненных участков по космическим снимкам;
- рассчитанные ущербы.

Для анализа рисков из перечисленных выше факторов выделены ключевые показатели, оказывающие наибольшее влияние на соответствующие риски (табл. 1).

Таблица 1

Ключевые показатели рисков нефтеразливов

	Показатель	Группа показателей
1	Удаленность от малого населенного пункта (население менее 5 000 чел.)	Населенные пункты
2	Удаленность от среднего населенного пункта (население от 5 000 до 30 000 чел.)	
3	Удаленность от крупного населенного пункта (население свыше 30 000 чел.)	
4	Удаленность от дороги с твердым покрытием	Транспортная инфраструктура
5	Удаленность от дороги без покрытия	
6	Удаленность от реки	Водные объекты
7	Удаленность от озера	
8	Удаленность от болота	
9	Тип леса	Лес
10	Класс леса	
11	Дела	Инциденты на лесном участке
12	Нарушения	
13	Ущербы	
14	Аварии	

Концептуальная схема построения модели оценки рисков состояния лесного фонда

Для построения модели оценки рисков изменения состояния лесного фонда, в частности нефтяных разливов, будут использоваться описанные выше данные, которые обладают следующими характерными свойствами:

1) Данные обладают множеством характеристик, т.е. являются многомерными;

2) Данные описываются как числовыми характеристиками (непрерывные значения), так и текстовыми (дискретные значения);

3) Данные обладают пространственными характеристиками, которые указываются либо в виде географических координат, либо в виде текстового описания местоположения (лесничество, квартал, выдел и т.д.);

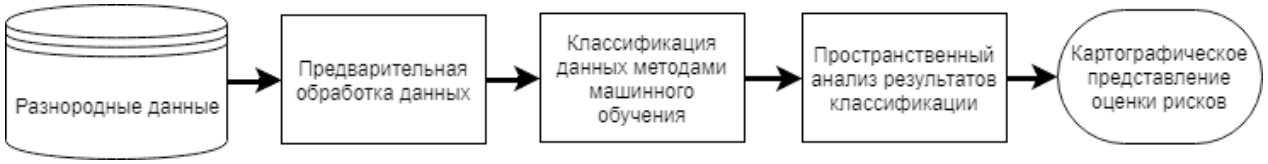


Рис. 1. Концептуальная схема построения модели оценки рисков изменения состояния лесного фонда

4) Местоположение данных представляет собой нерегулярную пространственную сетку;

5) Объем данных достаточно велик и составляет несколько десятков тысяч записей за несколько лет.

С учетом указанных характерных свойств имеющихся данных для оценки рисков изменения состояния лесного фонда Ханты-Мансийского автономного округа – Югры предлагается создание модели на основе сочетания методов машинного обучения и методов геоинформационного анализа. Использование методов машинного обучения обусловлено большими объемами разнородных данных, для обработки которых использование альтернативных методов может быть малоэффективным или слишком требовательным к ресурсам, как техническим (оборудование), так и временным (длительное время обработки). Для учета пространственных характеристик объектов исследования традиционно применяются методы ГИС-анализа (пространственного анализа) с использованием современных средств геоинформационных систем. Разработанная в ходе исследований концептуальная схема построения модели оценки рисков изменения состояния лесного фонда представлена на рисунке 1.

Препроцессинг данных

На этапе предварительной обработки данных, импортированных из информационной системы «Экономадзор», проведен ряд процедур по исправлению некорректных данных, удалению дублирующихся записей, приведению способа написания географических координат в единую систему в виде десятичной записи, добавлению отсутствующей информации о местоположении объекта или инцидента.

На следующем этапе все записи поставлены в соответствии с их пространственным расположением и отнесены к одному из 402 лесных участков.

Категорийные признаки (принадлежность участка к определенной компании, тип и вид леса и др.) преобразованы в вектора с использованием техники One Hot Encoding [13].

В результате предварительной обработки и преобразования исходной информации подготовлен набор данных для построения модели рисков загрязнения территории лесного фонда нефтью и нефтепродуктами. Размерность массива составила 402 записи с 83 признаками (каждая запись содержит информацию об инцидентах по годам, начиная с 2010 года). На этапе предварительной обработки данных наряду с языком Python 3, использованы открытые библиотеки NumPy, Pandas, scikit-learn [14–16].

Реализация модели

Решение поставленной задачи исследования сводится к реализации модели для прогнозирования возникновения риска на определенной территории Ханты-Мансийского автономного округа – Югры.

Задача прогнозирования сводится к выбору такой математической модели (оптимальной модели), при которой данные, полученные в ходе прогнозирования, как можно больше соответствуют реальным значениям. Нахождение оптимальной модели является крайне затруднительной задачей и зависит от исследуемого объекта и набора исходных данных.Примитивным решением возникшей задачи является перебор всех доступных математических моделей, но так как количество этих моделей не ограничено сверху, то данный подход на практике является неприменимым.

В качестве модели для прогнозирования используются многослойные нейронные сети [17–20]. Архитектура нейросетевой модели для прогнозирования загрязнений имеет 83 нейрона на входном слое (по количеству признаков для каждого лицензионного участка), три скрытых слоя, в каждом из которых в качестве функции активации используется сигмоид, и выходной слой, возвращающий вещественное число в интервале от 0 до 1. Это число, характеризующее степень негативного воздействия на лесной фонд в границах лицензионного участка, будем трактовать как оценку экологического риска изменения состояния лесного фонда по аналогии с понятием отнесения производственных объектов, оказывающих негативное воздействие на окружающую среду, к определенной категории риска. Таким образом, в качестве входных данных используются показатели, представленные в таблице 1, а также инциденты, произошедшие в 2014–2017 гг. В качестве метки для прогнозирования использовались аварии за 2018 год. Архитектура нейронной сети представлена на рисунке 2.

В качестве ответа получаем число в диапазоне от 0 до 1, которое позволяет категорировать степень негативного воздействия на лесной фонд на пять категорий риска в зависимости от значения:

$$\begin{cases} y = 0 – \textit{низкий риск}, \\ 0 < y \le 0{,}21 – \textit{умеренный риск}, \\ 0{,}21 < y \le 0{,}439 – \textit{средний риск}, \\ 0{,}439 < y \le 0{,}889 – \textit{значительный риск}, \\ 0{,}889 < y \le 1 – \textit{высокий риск}. \end{cases}$$

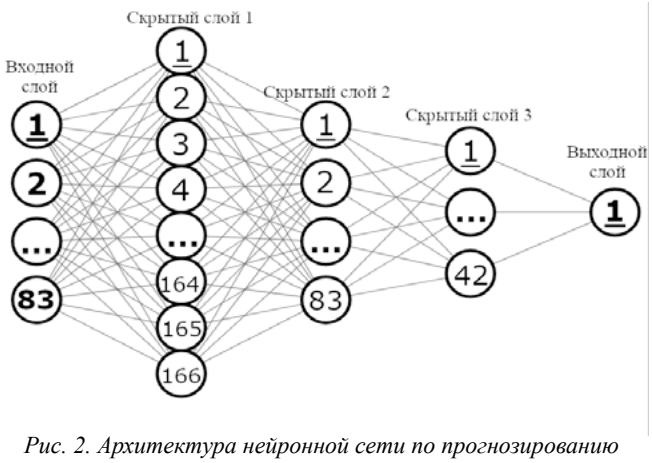


Рис. 2. Архитектура нейронной сети по прогнозированию нефтяных загрязнений

Для оценки качества модели применен метод кросс-валидации [21], и набор данных разбит на 10 блоков. Результаты оценки представлены в таблице 2.

Итоговая оценка качества реализованной нейросетевой модели, полученная как среднее от всех промежуточных оценок, составила 94,3 %, что говорит о ее достаточно высоком качестве и ее практической применимости для решения данного класса задач.

Визуализация

С использованием реализованной модели составлена тепловая карта с прогнозами аварий, связанных с нефтеразливами для лицензионных участков на 2019 год (рис. 3). Для визуализации используется NextGIS Web, предназначенная для хранения и организации многопользовательского доступа к геоданным.

Выводы

В рамках проекта описан подход к реализации контрольно-надзорной деятельности при осуществлении экологического надзора на примере Ханты-Мансийского автономного округа – Югры. Подход основан на применении имеющихся данных о ранее случившихся инцидентах, проведении препроцессинга сведений для формирования подходящей выборке данных и использовании многослойных нейронных сетей для реализации модели прогнозирования. С использованием реализованной модели построена тепловая карта, которая в зависимости от степени риска отображает лицензионные участки различным оттенком. Участки, имеющие наибольшую степень риска, должны проверяться с большей частотой.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-45-860003 «Модель оценки рисков состояния лесного фонда на территории Ханты-Мансийского автономного округа – Югры».

Таблица 2

Результаты оценки качества нейросетевой модели по прогнозированию лесных вырубок

Номер блока набора данных	1	2	3	4	5	6	7	8	9	10
Рассчитанная оценка, %	93,7	94,1	94,3	93,5	95,6	94,6	94,8	94,7	93,9	94,2

Рис. 3. Тепловая карта прогноза рисков загрязнения земель лесного фонда на 2019 год

Работа выполнена при поддержке Научного Фонда ЮГУ № 13-01-20/25.

Список литературы

1. Сведения о произошедших экологических инцидентах на территории Российской Федерации. Дата обращения: 10.01.2018.https://www.greenpeace.org/russia/Global/russia/report/Arctic-oil/Report_Russia_Spills_Oil.pdf.

2. Сведения о произошедших нефтяных разливах на территории Российской Федерации. Дата обращения: 10.01.2018. https://www.greenpeace.org/russia/Global/russia/report/oil-spills-may-18.pdf.

3. Бельская Е.Н., Бразговка О.В., Сугак Е.В. Методика расчета экологических рисков // Современные проблемы науки и образования. – 2014. – № 6.; URL: http://www.science-education.ru/ru/article/view?id=15755 (дата обращения: 01.11.2018).

4. Акинина Н.В. и др. Теория и практика применения нечетких сетей Петри для мониторинга экологических рисков // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2017. – №. 41. – С. 4–11. DOI: 10.17223/19988605/41/1.

5. Болбаков Р.Г. Принятие решений в геоинформационных системах // Государственный советник. – 2015. – №. 4. – С. 11–15.

6. Гвоздкова И.А. Эколого-ориентированные математические модели управления рисками и обеспечением безопасности // Вестник университета. – 2016. – №. 9. – С. 200–207.

7. Государственная информационная система «Эконадзор» Службы по контролю и надзору в сфере охраны окружающей среды, объектов животного мира и лесных отношений Ханты-Мансийского автономного округа – Югры. Дата обращения 01.02.2019. https://econadzor.admhmao.ru/

8. Baruque B. et al. A forecasting solution to the oil spill problem based on a hybrid intelligent system //Information Sciences. – 2010. – Т. 180. – №. 10. – С. 2029–2043.

9. Oliver T.H., Roy D. B. The pitfalls of ecological forecasting //Biological Journal of the Linnean Society. – 2015. – Т. 115. – №. 3. – С. 767–778.

10. Petchey O.L. et al. The ecological forecast horizon, and examples of its uses and determinants // Ecology Letters. – 2015. – Т. 18. – №. 7. – С. 597-611.

11. Dietze M.C. Prediction in ecology: a first□ principles framework //Ecological Applications. – 2017. – Т. 27. – №. 7. – С. 2048-2060.

12. Pennekamp F. et al. The practice of prediction: What can ecologists learn from applied, ecology-related fields? // Ecological complexity. – 2017. – Т. 32. – С. 156–167.

13. Kaggle. Using Categorical Data with One Hot Encoding.Датаобращения: 02.02.2019.https://www.kaggle.com/dansbecker/using-categorical-data-with-one-hot-encoding.

14. Van Der Walt S., Colbert S. C., Varoquaux G. The NumPy array: a structure for efficient numerical computation //Computing in Science & Engineering. – 2011. – Т. 13. – №. 2. – С. 22.

15. McKinney W. Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. – " O'Reilly Media, Inc.", 2012.

16. Pedregosa F. et al. Scikit-learn: Machine learning in Python //Journal of machine learning research. – 2011. – Т. 12. – №. Oct. – С. 2825-2830.

17. Rojas R. Neural networks: a systematic introduction. – Springer Science & Business Media, 2013.

18. Haykin S. Neural networks. – New York : Prentice hall, 1994. – Т. 2.

19. Riedmiller M., Braun H. A Direct Adaptive Method for Faster Backpropagation Learning: the RPROP Algorithm // International Conference of Neural Networks. San Francisco, USA, 1993. Vol. 1, pp. 586-591.

20. Zhang G., Patuwo B.E., Hu M.Y. Forecasting with artificial neural networks:: The state of the art // International journal of forecasting. – 1998. – Т. 14. – №. 1. – С. 35–62.

21. Hornik K., Stinchcombe M., White H. Multilayer feedforward networks are universal approximators // Neural networks. – 1989. – Т. 2. – №. 5. – С. 359–366.

22. Friedman J., Hastie T., Tibshirani R. The elements of statistical learning. – New York : Springer series in statistics, 2001. – Т. 1. – №. 10.

УДК 004.056.5

А.С. Боровский, Я.А. Боровская

ОБЕСПЕЧЕНИЕ КОМПЛЕКСНОЙ БЕЗОПАСНОСТИ КАК ОСНОВА УСТОЙЧИВОГО СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО РАЗВИТИЯ БИЗНЕСА

Аннотация

В настоящее время потребность в области теории безопасности имеет, прежде всего, прикладной характер. Вместе с тем без базовых фундаментальных знаний теории комплексной безопасности в целом нельзя в дальнейшем обеспечить устойчивое развитие регионального предприятия. Исходя из этого, сформулирована фундаментальная задача теории комплексной безопасности бизнеса – решить проблему устойчивого и безопасного развития регионального предприятия. В статье представлены подходы к возможному решению конкретной фундаментальной задачи, которая состоит в первоначальной разработке методологических, математических структур и теорий, выражающих комплексную безопасность

Ключевые слова: региональная экономика, стратегия бизнеса, устойчивое развитие предприятия, комплексная безопасность бизнеса, угрозы и риски бизнеса, принципы построения комплексной системы безопасности, методы оценки комплексной безопасности, экспертные системы.

Актуальность исследования

Безопасность предприятий (ее можно рассматривать в контексте региональной безопасности), в существенной степени зависит от уровня и потенциала развития регионов и очень тесно связана с национальной безопасностью страны в целом и, по сути, является ее важнейшим структурным компонентом. В современных условиях роль и значение региональной безопасности в обеспечении устойчивого развития России, а также обеспечении политической и социальной стабильности общества в условиях глобализации и резко возросшей в последние годы напряженности в отношениях Российской Федерации и стран Запада существенно возрастает. В связи с этим вопросам обеспечения безопасности региональных предприятий должно уделяться повышенное внимание, как в повседневной, так и в перспективной деятельности региональных органов власти. Чем сильнее региональные предприятия а, также стабильнее процессы их экономического развития, тем выше степень защищенности регионов.

Проблема комплексной безопасности региональных предприятий определяется, с одной стороны, их сложной многокомпонентной структурой, которые можно рассматривать как подсистемы региона и ее экономики с многими внутренними и внешними связями, а с другой – особой важностью решаемых ими задач, когда нарушение какой-либо безопасности (экономической, информационной, производственной, технической, экологической и т.д.) функционирования может привести к невосполнимому ущербу и катастрофическим последствиям. Важность и актуальность данной проблемы подтверждается Стратегией научно-технологического развития Российской Федерации-противодействие техногенным, биогенным, социокультурным угрозам, терроризму и идеологическому экстремизму, а также киберугрозам и иным источникам опасности.

Несмотря на широкое развитие и распространение различных видов безопасности (экономическая, информационная, экологическая, промышленная и т.д.), очень часто учитывается какой-то один вид, чаще всего экономическая безопасность, поскольку любые решения принимаются с точки зрения соотнесения вложенных средств и ожидаемого эффекта. С другой стороны, безопасность предприятия это всегда комплексная безопасность. Отдельные виды безопасности составляют единый комплекс, дополняют друг друга и обеспечивают базовые требования к системе защиты бизнеса. Разработка системы комплексной безопасности осуществляется только в рамках одного вида без привлечения соответствующих теоретических научных результатов, что в конечном итоге может привести к нарушению безопасности бизнеса.

Поэтому, учитывая степень важности и повсеместное распространение отдельных видов безопасности, необходимо развивать методологический и теоретический подход к разработке комплексной безопасности региональных предприятий.

Проблема исследования вопросов обеспечения безопасности бизнеса усложняется неопределенностью его функционирования в условиях непрерывно меняющихся внешних и внутренних угроз. Для повышения степени корректности постановки задач по обеспечению безопасности бизнеса необходимо повышать экспертные знания о системе комплексной безопасности в непрерывно изменяющихся условиях ее функционирования. В связи с тем, что процесс разработки и оценки комплексной безопасности предприятия требует знания экспертов, которые отражают неопределенность, неточность, неполноту, неоднозначность данной предметной области исследования то и вопросы, касающиеся теоретических и методологических основ разработки комплексной безопасности предприятия в условиях неопределенности остаются мало исследованными. Результатом этих исследований должны быть новые подходы, методы направленные на обеспечение устойчивого развития предприятия, предотвращения угроз безопасности предприятия, защита законных интересов компании.

Анализ современного состояния исследования

Область исследования, к которой относится данный проект, принадлежит к теории комплексной безопасности. Однако в настоящее время практически нет исследований, которые бы полностью отражали исследования всех видов безопасности в комплексе и их взаимодействие. В настоящее время известны исследования многих отечественных и зарубежных ученых, результаты которых нашли отражение в различных монографиях, книгах, статьях, посвященных проблемам разработки и оценки систем безопасности, касающихся отдельно: информационной безопасности, экономической безопасности, экологической безопасности, технической безопасности, промышленной безопасности и т.д. Так как у авторов имеются значительные теоретические и практические наработки в области технической защиты информации, относящейся к информационной безопасности рассмотрим современное состояние исследований в данной области науки с этой точки зрения. За последние годы в области разработок технической защиты накоплены большие знания и опыт. Первой публикацией, доступной широкому кругу читателей и посвященной разработке и оценке технической защиты информации, стала известная книга М. Гарсия «Проектирование и оценка систем физической защиты». В ее основу положены разработки Национальной лаборатории «Сандия» (США). На сайте Сандийских лабораторий открыт раздел, посвященный обзору методик оценки уязвимости в различных отраслях. В книге детально рассматриваются подходы к проектированию, анализу технических средств защиты в целом и отдельно составляющих ее элементов – подсистем физической защиты: подсистеме обнаружения; подсистеме задержки; подсистеме реагирования. Автор подробно останавливается на особенностях проектирования каждой из подсистем. Отдельная глава посвящена вопросам проведения анализа/оценки системы технической защиты – детально рассмотрена компьютерная модель EASI с примерами ее работы. В 2007 г., в связи с огромной популярностью книги «Проектирование и оценка систем физической защиты» М. Гарсия выпускает второе, доработанное издание книги, в которой была добавлена глава по физической защите материалов при транспортировании, а также переработаны отдельные главы первого издания. В 2005 г. М. Гарсия издает книгу «Оценка уязвимости систем физической защиты» – книга является одним из первых изданий, где подробно расписана методика оценки уязвимости объектов. Много внимания уделено практической стороне при проведении оценки уязвимости объектов. Достаточно подробно описываются вопросы оценки эффективности технических систем защиты, разработки модели угроз, модели нарушителя, все это иллюстрируется различными диаграммами и схемами. Один из разделов книги посвящен вопросам подготовки отчетов по проведению оценки уязвимости объекта. Также имеется много справочного материала: описание типовых угроз, целей защиты, временные и вероятностные характеристики нарушителя и инженерно-технических средств охраны. Джеймс Ф. Бродер выпускает третье переработанное издание своей книги «Анализ риска и исследование безопасности». Автор достаточно подробно описывает процедуры, необходимые для оценки безопасности. В первой части книги рассматривается вопрос анализа рисков, включая разъяснение (обоснование) базисных понятий: риск, уязвимость, угроза. Приводится общая математическая модель оценки безопасности и примеры формул ее количественной оценки. Вторая часть книги посвящена обзору возможных чрезвычайных ситуаций и планированию действий при возникновении чрезвычайных ситуаций, разработке антикризисных планов. Из отечественных работ, посвященных вопросам категорирования объектов и анализа систем защиты, пожалуй, наиболее фундаментальной является монография, авторов Бояринцев А.В., Бражник А.Н., Зуев А.Г. «Проблемы антитерроризма: Категорирование и анализ уязвимости объектов», написанная сотрудниками предприятия «ИСТА – Системс» (г. Санкт-Петербург) в 2006 году. В монографии большое внимание уделяется рассмотрению методологии обоснования требований к техническим средствам защиты, основанной на категорировании и анализе уязвимости объектов. Достаточно подробно изложены методические подходы к оценке эффективности системы технической защиты информации. Рассмотрены подходы к определению количественных и качественных требований к системе технической защиты информации. Впервые дана общая характеристика и классификация программных средств оценки эффективности системы технической защиты информации, как отечественных разработок, так и зарубежных. В монографии Оленин Ю.А. «Системы и средства управления физической защитой объектов» автором обобщен опыт системного подхода к созданию комплексов технических средств физической защиты, с единых системных позиций рассматриваются пути повышения эффективности управления системами физической защиты объектов, в основе которого лежит формируемая средствами обнаружения и системой контроля и управления доступом сигнальная информация. Рассмотрены методы графоаналитической формализации процедур обработки сигнальной информации, основанные на теории сигнатурного анализа траекторий движения нарушителя. Вопросам повышения эффективности охранно-пожарной сигнализации, созданию систем охраны от хищений для не режимных объектов, разработке систем и средств управления защитой объектов посвящены монографии: Моторный И.Д. «Современный терроризм и оценка диверсионно – террористической уязвимости гражданских объектов»; Шепитько Г.Е. «Проблемы охранной безопасности объектов». Вопросам комплексной безопасности различных предприятий посвящены

статьи и монография А.П. Стерхова. Так в статье «Анализ устойчивости предприятия с системных позиций обеспечения комплексной безопасности бизнеса» на основе системного подхода к обеспечению безопасности инновационного бизнеса, а также на базе сформированной концептуальной модели комплексной безопасности исследуются вопросы взаимосвязи и взаимодействия его устойчивости и безопасности. Сформулированы определения устойчивости бизнеса, его статической и динамической устойчивости. С позиций указанного подхода выделяется несколько системообразующих факторов для обеспечения устойчивости предприятия. В статье «Комплексная безопасность бизнеса и основные проблемы ее обеспечения» сформулировано определение комплексной безопасности, определен перечень основных угроз и рисков для бизнеса, дана их классификация. Даны главные составляющие комплексной безопасности бизнеса. В монографии «Основы теории комплексной безопасности промышленных предприятий инновационного типа» с системных позиций рассматриваются вопросы обеспечения комплексной безопасности и устойчивости промышленных предприятий инновационного типа. Показывается, что методология обеспечения безопасности и развития предприятий должна строится на основе системного подхода. В статье авторов Власенко М.Н., Шедько Ю.Н. «Экономическая безопасность и ее экологическая составляющая – значимый фактор устойчивого развития предприятий в условиях современного рынка» исследуется роль и место экологии как составной части экономической безопасности предприятия, ее влияние на устойчивое развитие участников рыночных отношений. Классифицированы основные группы угроз для бизнеса, содержащие в себе экологическую составляющую. Сформулированы основные задачи, которые можно решить с помощью экологического аудита. Предложены основные меры противодействия негативным проявлениям и угрозам предприятию в сфере экологии как значимой подсистеме экономической безопасности организации. В статье «Анализ подходов к оценке региональных процессов формирования социально-экономической безопасности» авторов Чичканов В.П., Беляевская-Плотник Л.А. исследуются вопросы обоснования учета социальной компоненты при оценке уровня социально-экономической безопасности региона. Проанализированы пороговые значения отдельных индикаторов социально-экономической безопасности. Систематизированы основные методологические подходы к моделированию и прогнозированию социально-экономической безопасности региона. В развитии теории устойчивости в настоящее время нет единого подхода к методологии понимания самой теории устойчивости в бизнесе. Однако можно выделить несколько источников которые интенсивно развивают данную тему. Так в источнике автора С.Н. Бобылев «Индикаторы устойчивого развития: региональное измерение. Пособие по региональной экономической политике» исследуются вопросы разработки и использования индикаторов устойчивого развития в российских регионах. Рассматриваются методические подходы к разработке индикаторов, международный и российский опыт в этой области. Анализируются возможности и ограничения использования в регионах систем показателей и агрегированных индикаторов устойчивости. На примере конкретных российских регионов показан конкретный позитивный опыт использования индикаторов. В работе Урсул А.Д., Лось В.А. «Стратегия перехода России на модель устойчивого развития: проблемы и перспективы» изложены концептуальные основы устойчивого развития цивилизации, истоки и причины обострения социально-экологической ситуации в России, проблемы экологической и демографической безопасности, возможные сценарии экоразвития России. Особое внимание уделено ноосферной ориентации и перспективам устойчивого развития. Как видно из анализа приведенных выше источников, вопросы теоретической проработки анализа, синтеза, путей повышения эффективности отдельных видов безопасности исследуются достаточно интенсивно, и получены существенные научные результаты, выразившиеся в новых методиках, например ранжирования угроз, анализа уязвимости предприятий и т.д. Однако до сих пор не рассматривался подход комплексно к разработке системы безопасности бизнеса, к учетом его устойчивого развития исходя из области неопределенности, неточности и неполноты информации предметной области исследования.

Выводы

Одной из важнейших задач решаемых руководителем предприятия является задача эффективного и устойчивого развития бизнеса с целью получения прибыли. Однако данная задача в условиях постоянного воздействия различного вида угроз постоянно находится в зоне риска. С целью уменьшения составляющих риска необходимо учитывать влияние различных составляющих комплексной безопасности бизнеса. Именно комплексную безопасность можно рассматривать, как системообразующий фактор устойчивого развития бизнеса, а устойчивость интерпретировать как баланс между опасностью и угрозами. Для достижения этого баланса необходимо решить ряд задач, и на основе их решения можно будет разработать новые методы теоретического исследования комплексной безопасности, а именно [1–16]:

- метод обоснования показателя качества функционирования комплексной системы безопасности для устойчивого развития бизнеса в условиях неопределенности, с необходимостью установления математических зависимостей между степенью выполнения системой безопасности своих функций и полнотой оснащенности предприятия различными видами под-

систем безопасности основу данного метода составляет нечеткий многокритериальный анализ – в дальнейшем на основе установления математических зависимостей можно будет прогнозировать состав комплексной безопасности для бизнеса;

- метод определения защищенности предприятия на основе теории нечетких гиперграфов, позволяющий определять уровни защищенности предприятия, который даст качественную и количественную оценку защищенности;

- метод модифицированных DFD-диаграмм (Data Flow Diagram) как средства представления предприятия на производственно-технологическом и структурно-логическом уровнях, который позволяет учесть в модели комплексной безопасности производственно-технологический процесс, для учета влияния различных видов безопасности друг на друга;

- метод адаптации генетического алгоритма под данную предметную область с целью оптимизации проектных решений, позволяющий спрогнозировать виды подсистем комплексной безопасности и их влияния на устойчивость бизнеса;

- разработка алгоритмов и программ разрабатываемых методов и моделей и проверка их адекватности.

Список литературы

1. Костин В.Н. Обоснование требований к эффективности подсистем физической защиты объектов информатизации / Соловьев Н.А., Тишина Н.А. // Научно-технический вестник Поволжья. 2018. № 4. С. 125–128.

2. Шрейдер М.Ю. Разработка моделей описания обработки информации в задачах управления системой защиты объекта на основе сетей петри / Шрейдер М.Ю., Боровский А.С. // Вопросы кибербезопасности. 2018. № 1(25). С. 46–53.

3. Боровский А.С. Построение модели системы защиты в облачных технологиях на основе многоагентного подхода с использованием автоматной модели / Боровский А.С., Ряполова Е.И. // Вопросы кибербезопасности. 2017. № 4 (22). С. 10–20.

4. Пищухин А.М. Управление предприятием на основе прогноза в ассортиментном пространстве / Пищухин А.М. // Экономика региона. 2017. Т. 13. № 1. С. 216–225.

5. Пищухин А.М. Система компьютерной поддержки формирования стратегии управления производством / Пищухин А.М., Чернов Ф.В. // Фундаментальные исследования. 2017. № 10–3. С. 512–517.

6. Пищухин А.М. Ступенчатое адаптивное управление с эталонной моделью / Пищухин А.М., Ахмедьянова Г.Ф. // Известия Самарского научного центра Российской академии наук. 2017. Т. 19. № 1-1. С. 157–161.

7. Акимов С.С. Применение искусственного интеллекта в прикладных областях / Акимов С.С., Кушнерова И.А. // В сборнике: Ключевые проблемы и передовые разработки в современной науке Сборник научных трудов по материалам I Международной научно-практической конференции. 2017. С. 106–107.

8. Костин В.Н. Синтез оптимального размещения технических средств систем физической защиты критически важных объектов / Костин В.Н. // Информационные технологии – 2017 № 1. С. 41–49.

9. Костин В.Н. Метод оценки утечки конфиденциальной информации о функционировании системы защиты объекта информатизации по информационному критерию / Костин В.Н., Боровский А.С. // Вестник компьютерных и информационных технологий – 2016. № 8. С. 34–43.

10. Костин В.Н. Оценка потенциала опасности нарушителей на основе информационного метода и метода главных компонент / Костин В.Н. // Информационные технологии и вычислительные системы. – 2016. № 3. С. 74–81.

11. Фот Ю.Д. Модель определения уровня защищенности системы отбора персонала организации / Фот Ю.Д., Боровский А.С. // Интеллект. Инновации. Инвестиции. 2016. № 2. С. 117–123.

12. Боровский А.С. Программный комплекс информационной поддержки задачи проектирования системы физической защиты / Боровский А.С., Тарасов А.Д. // Системы управления и информационные технологии. 2016. Т. 66. № 4.1. С. 122–128.

13. Акимов С.С. Разработка модели инвестиционного портфеля / Акимов С.С. // Интеллект. Инновации. Инвестиции. – 2016. – № 5. – С. 6–8.

14. Костин В.Н. Метод оценки глубины прогноза развития (эволюции) характеристик сложных систем на основе энтропийного подхода / В.Н. Костин, Д.В. Даньшин // Информационные технологии – 2015. № 1. С. 61–68.

15. Костин В.Н. Информационно вероятностный метод формирования категорий потенциально опасных объектов / В.Н. Костин, А.К. Пономарев // Вестник компьютерных и информационных технологий – 2015. № 6 с. 34–42.

16. Боровский А.С. Системный анализ обеспечения стабильности эффективного функционирования инновационной и цифровой экономики на основе интеллектуализации системы комплексной безопасности / П.И. Огородников, Г.М. Залозная // Экономика региона. – 2018. – Т.14, вып. 4. – С. 1221–1231.

УДК 004.056.5 669.183.2 658.51.001.57

V.I. Vasilyev, A.D. Kirillova, A.E. Sulavko, S.S. Zhumazhanova

ANALYSIS OF CONFIDENTIAL DATA PROTECTION IN CRITICAL INFORMATION INFRASTRUCTURE AND THE USE OF BIOMETRIC, NEURAL NETWORK AND CRYPTOGRAPHIC ALGORITHMS (STANDARDS REVIEW AND PERSPECTIVES)

Abstract

The paper is devoted to the problem of protecting critical infrastructure objects (CIO) from malicious cyberattacks. The main international and national standards (NIST SP 800-82, NERC CIP, ISA/IEC 62443), as well as the other normative-legal documents in the field of IACS cybersecurity (Russian Federal Law 187-FZ, FSTEC of Russia's Orders No. 31 and 239, GOST R MEC 62443) are considered. The ability to confidently identify subjects is critical to providing fast, safe, and effective business or different types of processes. From this point of view, we consider reliable subjects identification method based on cryptographic key generation using person's biometric features, their storing in protected neural network containers and Russian standards on them.

Keywords: information security, critical infrastructure protection, control systems, cybersecurity, biometry, neural network containers.

I. Introduction

The statistics of last years shows the sharp rise of the number of hackers' attacks on critical infrastructure objects, including Industrial Automation and Control Systems (IACS) in many countries. According to Kaspersky Lab [1] 41.2 % of IACS were subjected to cyber-attacks in the first half of 2018.

Today it is generally recognized that security of IACS does not come to providing information security, i.e. providing the confidentiality of gathering, processing and transferring information. The security of IACS should consist first of all in providing continuity and integrity of technological process by itself that relates to new fast developing direction – cybersecurity of IACS [2, 3]. The cybersecurity is understood here as realization of measures taken to protect IACS against threats to security through accidental circumstances, actions or events, or through deliberate attack.

Taking into account the topicality of mentioned problem, consider below the issues of normative-legal provision of IACS cybersecurity connected primarily with development of corresponding standards. From the general requirements for protection against cyberattacks, society has come to the standardization of specific methods to ensure authorized access. For example, methods of biometric protection, neural network and cryptographic algorithms are already brought to a number of standards separately, but a number of methods of security are born at the intersection of directions and also introduced as standards. A series of similar GOSTs will also be discussed in this article.

II. American and international standards in the field of IACS cybersecurity

NIST SP 800-82 "Guide to Industrial Control Systems (ICS) Security" – developed by the Computer Security Resource Center (CSRC) of the National Institute of Standards and Technology (NIST) in 2011 (Revision 2 – in 2015) [4]. This standard contains:

- practical recommendations connected with design of information security system;
- simplificated models of intruder and IACS vulnerabilities;
- list of typical threats and vulnerabilities of IACS;
- recommendations on creation and implementation of IACS security maintenance program;
- description of IACS and main security subsystems architecture;
- description of traditional information security subsystems (access control and management, identification and authentication, antivirus protection etc.).

NERC Critical Infrastructure Protection (CIP), Cybersecurity – the family of industry standards, developed by the North-American Electric Reliability Corporation) in the early 2000s [5]. The aim of these standard is in providing the reliable protection of automated systems and communication networks for energy sector objects against possible cyberthreats. The family of the standards includes 9 standards NERC CIP-001÷009, including such as:

- CIP-002-1 Critical Cyber Asset Identification;
- CIP-005-1 Electronic Security Perimeter;
- CIP-006-1 Physical Security;
- CIP-007-1 Systems Security Management.

The standards NERC CIP (as well as the standard NIST SP 800-82) represent detailed guides on providing IACS cybersecurity which are successfully applied not only in USA, but in many other countries including Russia.

ANSI/ISA-99 "ISA-99 Security for Industrial Automation and Control Systems" – the family of standards developed by the committee ISA-99 of the International Society of Automation (ISA).In 2007 the standard ISA-99 was approved by the American National Institute of Standards (ANSI) and published as ANSI/

ISA-99, and in 2010 it was republished as ANSI/ISA-62443 series and their modified version IEC 62443 [6].

Two of the most important features regarding the measures included throughout above mentioned standards are called security zones and conduits. A security zone is a group of physical or logical assets that share common security requirements. A conduit is a communication path between two security zones. Conduits control access to zones, resist external attacks or the transfer of malware, protect the integrity and confidentiality of network traffic. Focusing on conduit security functions is typically more effective than having to upgrade every device or computer in a zone to meet the specified security requirements.

IEC 62443 series of standards "Industrial Automation and Control Systems Security" – developed jointly by the ISA-99 committee and the International Electrotechnic Commission (IEC). The goal of the 62443 series is to improve the safety, availability, integrity and confidentiality of components or systems used for industrial automation and control, and to provide criteria for procuring and implementing secure IACS.

The work on creating the standards is under way from 2009. It is planned to release 13 key standards and technical reports divided in 4 groups:

1. General (basic notions, models, terminology);
2. Policies and Procedures (requirements to security management system, rules in the framework of security program);
3. System (technologies of providing IACS security, risk and security level assessment);
4. Component (security requirements at the level of separate subsystems and components of IACS).

Accordingly to IEC 62443 approach, the design of IACS security management system must include the following stages:

- high-level risk assessment;
- design of IACS reference model;
- design of asset model;
- design of reference architecture model;
- design of zone and conduit model;
- detailed risk analysis for each zone;
- determination of current security and target security levels of zones;
- selection of corresponding countermeasures.

The separate standards of IEC 62443 series are translated now in Russian and actively used by the leading Russian companies at the market of IACS security.

III. Normative-legal acts in the field of IACS security in Russia

At the moment, the notions "cyberspace" and "cybersecurity" are absent in the Russian legislation. In 2014 the "Conception of cybersecurity strategy of the Russian Federation" was presented to discussion, but this document was not approved by the government.

At the same time, it is necessary to note that the concrete security requirements to IACS exploited at critically important objects are determined by the Order of Federal Service on Technical and Export Control (FSTEC) of Russia № 31 [7]. The given Order considers 3-level architecture of IACS, including: data input/output level; automatic control level and operator (dispatching) control level. The requirements to IACS level security depend on the security class which in its turn depends on the rate of significance (criticality) of information processed at this level. The specific basic sets of countermeasures for each security class are recommended. The chosen security measures are considered separately for each IACS level with account of peculiarities of each level operation.

The important role in solving the problems of IACS security belongs to the Federal Law "On the security of critical information infrastructure of the Russian Federation" № 187-FZ [8] and the Order of FSTEC of Russia № 239 [9]. The notion of critical information infrastructure (CII) objects covers such groups of objects as information systems, information-telecommunication networks, automated control and management systems of the Russian Federation subjects, functioning of which is critically important for the state. The Order of FSTEC of Russia № 239 contains recommendations concerning providing the security of significant CII objects at different stages of their lifecycle. The list of basic measures of providing security for different CII significance categories in presented.

And, finally, from 2015 the series of national standards GOST R MEC 62443 "Industrial communication networks. Network and system security" began to be edited in Russia in the framework of the policy of national standards system harmonization [10]. At the moment, three standards of this series are translated and adopted in Russia:

- GOST R MEC 62443-1-1-2009. Part 1-1. Terminology, concepts and models;
- GOST R MEC 62443-2-1-2015. Part 2-1. Establishing an industrial automation and control system security programme;
- GOST R MEC 62443-3-3-2016. Part 3-3. System security requirements and security levels.

The continuation of these standards series translation and edition as completing the ISA work under the planned versions of IEC 62443 standards. Obviously, in this case the specialists working in the sphere of IACS cybersecurity will acquire a powerful tool for achieving the settled goals.

IV. Standards and Methodf of Confidential data protection using biometric, neural net and cryptographic instruments

Data protection from unauthorized access at the level of storage and transmission, as a rule, is ensured by their encryption. Modern encryption methods are reliable if we use strong keys. There are standards for cryptographic algorithms. When working with official documentation, legally relevant documents in electronic

form, it is allowed to use only algorithms that meet the standards. Russian Federation has its own state standards for encryption algorithms and generation / verification of electronic signatures (ES): GOST R 28147-89 [11], GOST R 34.10-2012 [12]. Standards define the requirements for cryptographic transformation, data encryption, electronic digital signature scheme, processes of digital signature generation and verification.

Durability of keys is not only in the possibility of their selection, but also the loss, transfer to third parties. There are various approaches to generation or production of bit sequences using the subject's biometric data (for example, face characteristics, handwriting or keystroke dynamics). There are a lot of biometric technologies; the primary processing of biometric images for each technology is standardized, for example:

1	GOST R ISO/IEC 19794-5-2013 "Information technology – Biometric data interchange formats – Part 5: Face image data"
2	GOST R ISO/IEC 19794-4:2011 "Information technologies. Biometrics. Biometric data interchange formats. Part 4. Finger image data"
3	GOST R ISO/IEC 19794-6 2014 "Information technology. Biometrics. Biometric data interchange formats. Part 6. Iris image data"
4	GOST R ISO/IEC 19794-11-2015 "Information technologies. Biometrics. Biometric data interchange formats. Part 11. Signature/sign processed dynamic data"
5	GOST R ISO/IEC 19794-10-2010 "Automatic identification. Biometric identification. Biometric data interchange formats. Part 10. Hand geometry silhouette data"

By definition, given in GOST R 52633.0-2006 [13]: a biometrics-code converter converts a vector of "Fuzzy" biometric parameters into a clear and unique key code (password) and responds with a random output code to a random input vector that does not belong to the set images of "Own". By images, that don't belong to the set of "Own", is meant the images of "Stranger" (the image of "All strangers"). Accordingly, the images "Own" and "Stranger" are vectors of biometric parameters (features) characterizing the legal user and the attacker (or an unknown subject).

There are a number of basic approaches for building a biometrics-code converter: fuzzy extractor, neural network biometrics-code converter (NNBCC) based on a network of perceptrons trained according to GOST R 52633.5-2011, using neuron networks based on Pearson's metrics, multidimensional Bayesian functionals, etc. Foreign researchers consider the use of fuzzy extractors [14, 15], the use of which is not yet standardized at the international level. The technology of neural network converters "biometric- code" is being developed and standardized in Russian Federation. The list of Russian standards of the protection technology:

1	GOST R 52633.0-2006 "Information protection. Information protection technology. Requirements for the means of high-reliability biometric authentication"
2	GOST R 52633.1-2009 "Information protection. Information protection technology. Requirements for creation procedures for bases of natural biometric images, intended for high-reliability biometric authentication means testing"
3	GOST R 52633.2-2010 "Information protection. Information protection technology. Requirements for creation procedures for bases of synthetic biometric images, intended for high-reliability biometric authentication means testing"
4	GOST R 52633.4–2012 "Information protection. Information protection technology. Neural network biometric-code converters programming interface"
5	GOST R 52633.5-2011 "Information protection. Information protection technology. The neural net biometry-code converter automatic training"
6	Technical specification "Protection of neural network biometric containers using cryptographic algorithms" (project, public discussion of the first edition started on February 1, 2017 by members of TC 26 "Cryptographic protection of information", public discussion of the second edition started on 01.06.2018)

Full description and analysis of Russian and international standards is provided in [16].

Regardless of the chosen approach in practice, there should be the possibility of presenting and storing subject's templates in a form that does not allow restoring their biometric features (features values or biometric image). This is achieved by storing not the parameters of the distribution laws in their original form, but the weights of the inputs of neurons, each input is associated with a specific biometric feature.

GOST R 52633.5-2011 recommends to use the direct calculation of the modules of weight coefficients in terms of mathematical expectations and standard deviations of the "Own" and "Stranger" biometric parameters. At least 21 sample of "Own" data is required for network training and 64 independent samples of "Stranger" data. It is also necessary that the biometric data have a distribution law that is close to normal, to control this, the Pearson criterion is used, as well as its modifications [17]. The advantage of biometric networks is that the human biometric template is not stored in computer memory, instead of which weights between neurons are stored (there is no effective way to restore the parameters of the distribution of biometric features from the data of the neural network biometric container). If the tables of trained neurons are stored openly, then the attacker has the opportunity to extract from the open neural network container knowledge about the access code and / or about the biometric image "Own". In order to make the storage and transmission of biometric information in safe mode, it is necessary to use cryptographic protection mechanisms, for example, by encryption of the parameters of neurons at

the outputs of other neurons to protect the tables of neural network functionals from analysis in order to restore the subject's templates. This principle of protection is called a protected neural network (biometric) container.

During authentication, one or more "Own" biometric images are presented, each of them is converted into a vector of features values, then the vectors are combined into a single features vector (according to the concatenation principle). The vector is fed to the input NNBCC. A bit sequence is being generated. This sequence can be a password or a secret key for asymmetric encryption (an ES key), depending on the procedure being performed. If a key is generated, it is necessary to check its correctness. For this purpose, arbitrary data is encrypted on the generated private key and decrypted using the public key. Protection in this way does not interfere with the "biometrics – code" neural network converter to give a stable "Own" code at the output upon presentation of examples of the "Own" image. When the data of the "Strange" image are fed to the inputs of the protected neural network converter, then the generated data is different, and the output code becomes unstable (random).

According to the results of a series of experiments on user's handwritten signature data and face parameters, the Bayes-Pearson-Hemming network shows the best result in the reliability of authentication and key length [18, 19].

Conclusion

The last years are characterized by the sharp escalation of the problem of providing IACS cybersecurity. The important role in solving this problem belongs to development of national and international, as well as the other normative-legal documents in the field of IACS cybersecurity.

The normative-legal documents have to consider new methods for information protection making accustomed mechanisms of information protection more effective and safe using cryptographic and biometric methods.

Acknowledgment

The reported study was funded by RFBR according to the research project № 17-48-020095.

References

1. Threat landscape for industrial automation systems in the second half of 2017. – URL: https://ics-cert.kaspersky.com/media/KL_ICS_REPORT_H2-2017_FINAL_EN_22032018.pdf

2. Cyber Security for Industrial Automation and Control Systems (IACS). – URL: https://www.hse.gov.uk/foi/internalops/og/og-0086.pdf

3. A summary of Cybersecurity Best Practices, NHTSA, October 2014. – URL: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/812075-cybersecuritybestpractices.pdf

4. Guide to Industrial Control Systems (ICS) Security / NIST Special Publication 800-82. Rev. 2, May 2015. – URL: https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp800-82r2.pdf.

5. Introduction to NERC CIP Version 5 / Steven Parker, 06.01.2014. – URL: https://www.powermag.com/introduction -to-nerc-cip-version-5/

6. IEC 62443: Evolution of the ISA 99 / Security and Industry CERT, 25.08.2015. – URL: https://www.certsi.es/en/blog/iec62443-evolution -of-isa99.

7. On approval of requirements to provision of information security in automated systems of production and technological processes control at critically important objects, potentially dangerous objects, and the objects representing higher danger to the human life and health and environment / Order of FSTEC of Russia No. 31 dated of 14.03.2014. – URL: https://fstec.ru/normotvorcheskaya/akty/53-prikazy/868-prikaz-fstek-rossii-ot

8. On the security of critical information infrastructure of the Russian Federation / Federal Law № 187-FZ dated of 26.07.2017. – URL: https://www.consultant.ru/document/cons_doc_LAW_220885/

9. On approval of Requirements to providing security of significant objects of critical information infrastructure of the Russian Federation / Order of FSTEC of Russia No. 239 dated of 23.12.2017. – URL: https://fstec.ru/normotvorcheskaya/akty/53-prikazy/868-prikaz-fstek-rossii-ot

10. Chuchaev S.V. "The issues of standardization in IACS IS" In: *Association of Information Security Services Chiefs*. – URL: https://new.groteck.ru/images/catalog/63227/4653c4db0ef4361331f0ee4930c4175c.pdf

11. GOST R 28147-89 Information processing systems. Cryptographic protection. Cryptographic Algorithm. – Intr. 1990-07-01. Moscow: USSR State Committee on standards; M.: IPC Publishing House of Standards. 1990: pp. 28. (In. Russ.).

12. GOST R 34.10-2012 Information technology. Cryptographic data security. Generation and verification processes of electronic digital signature. – Intr. 2013-01-01. Moscow: Federal Agency on Technical Regulating and Metrology; M.: Standartinform. 2012: pp. 20. (In. Russ.).

13. GOST R 52633.0-2006 Information protection. Information protection technology. Requirements to the means of high-reliability biometric authentication. – Intr. 2007-04-01. Moscow: Federal Agency on Technical Regulating and Metrology; M.: Standartinform. 2007: pp. 25. (In. Russ.).

14. Ramirez-Ruiz, J., Pfeiffer, C., Nolazco-Flores, J. Cryptographic Keys Generation Using FingerCodes. Advances in Artificial Intelligence – IBERAMIASBIA 2006 (LNCS 4140): 178–187.

15. Feng, Hao, Anderson, R. and Daugman, J. Crypto with Biometrics Effectively, IEEE Transactions on Computers, 55(9), September 2006.

16. Ivanov, A.I. Neyrosetevaya biometriya dlya oblakov Rossiyskiye standarty dlya zashchity tsifrovykh prav grazhdan [Neural Network Biometrics for Clouds.

Russian Standards for the Protection of Digital Rights of Citizens]. *Access Control Systems (ACS)*: 134–139. (In Russ.)

17. Ivanov, A.I., Gazin, A.I., Vyatchanin, S.E., Perfilov, K.A. Sravneniye moshchnosti khi-kvadrat kriteriya i kriteriya Kramera – fon Mizesa dlya malykh testovykh vyborok biometricheskikh dannykh [Comparison of the power of the Chi-square criterion and the Cramer-von Mises criterion for small test samples of biometric data]. *Reliability and quality of complex systems*, 2016, 2(14): 21–28. (in Russ.)

18. Sulavko, A.E., Eremenko, A.V. and Zhumazhanova, S.S., Buraya, E.V. Generatsiya klyuchevykh posledovatel'nostey i verifikatsiya sub'yektov na osnove dvumernogo izobrazheniya litsa [Key sequence generation and verification of subjects based on two-dimensional face image]. *Automation of Control Processes*, 2017, 1 (47): 58–66. (in Russ.)

19. Lozhnikov, P.S., Sulavko, A.E., Eremenko, A.V., Volkov, D.A. Eksperimental'naya Otsenka Nadezhnosti Verifikatsii Podpisi Setyami Kvadratichnykh Form, Nechetkimi Ekstraktorami [Experimental Evaluation of Reliability of Signature Verification by Quadratic Form Networks, Fuzzy Extractors and Perceptrons]. *Information and Control Systems,* 2016, 5 (84): 73–85. (In Russ.)

УДК 004.04 004.9

Ю.И. Молородов, Г.Б. Загорулько

АРХИТЕКТУРА И ОСОБЕННОСТИ ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ ДОСТУПА К УДАЛЕННЫМ ХРАНИЛИЩАМ ДАННЫХ

Аннотация

В научном сообществе стоит проблема разработки информационно–аналитических ресурсов, которые бы обеспечивали систематизацию накопленных знаний и фактического материала в областях знаний (ОЗ) связанных с охраной окружающей среды. Современным подходом, является технология разработки интеллектуальных научных Интернет–ресурсов (ИНИР), которая основывается на онтологическом подходе. Важной для районов, расположенных в непосредственной близости от нефтепромысловых объектов с интенсивной нефтедобычей. добычи нефти в Западно-Сибирском нефтегазоносном бассейне, является ОЗ, связанная с загрязнением окружающей среды, в том числе и от горящих газовых факелов. Эта ОЗ является неотъемлемой составляющей науки об окружающей среде, которая в значительной степени базируется на огромных массивах данных, хранящихся, как правило, в специально организованных хранилищах. Ключевая задача информатизации наук об окружающей среде состоит в организации удаленного доступа и дальнейшей коллективной работы с этими хранилищами. С развитием Интернет-технологий появилась возможность по-новому организовывать хранилища данных и доступ к ним. Теперь пользователю необходим удобный инструментарий и средства для доступа и возможности первичного анализа таких данных. Следовательно, по-прежнему актуальна проблема разработки инфраструктуры, которая позволит объединить информационные системы, созданные на основе технологии ИНИР.

Ключевые слова: интеллектуальный научный Интернет-ресурс, онтология, сервис, внешнее хранилище данных, система доступа к данным.

Введение

Инфраструктура мегаполисов чрезвычайно насыщена производствами потенциально опасных объектов, что приводит к техногенным и природным катастрофам, в том числе и с выбросом аэрозолей. Аэрозольные загрязнения наиболее динамичны и представляют собой непосредственную угрозу здоровью и окружающей среде. Именно эти обстоятельства сделали актуальными решения задачи обеспечения мониторинга атмосферы крупных промышленных центров, оценки последствий и защиты от техногенных катастроф и аварий, пожаров и стихийных бедствий. Указанный круг проблем активно изучают коллективы специалистов, работающие в различных областях наук об окружающей среде. Эти коллективы успешно работают в области физики, химии, гидро- и газодинамики, биологии, экологии и т.д. Из этого следует необходимость междисциплинарных подходов в науках об окружающей среде, а присущее любой науке стремление к пониманию явлений и их прогнозированию приводит к широкому использованию математического моделирования в качестве вычислительного аппарата.

Накопленные большие объемы фактического материала привели к проблеме их систематизации и предоставления к ним эффективного содержательного доступа широкому кругу пользователей. Для этого разрабатываются информационные системы (ИС), в основу которых закладываются разные подходы и принципы, отражающие специфику представляемых ими областей знаний (ОЗ) и решаемых ими задач [1, 2]. К создаваемым ИС предъявляются большие требования. Они должны учитывать неоднородность, распределенность, продолжающийся рост объемов, как экспериментальных данных, так и неструктурированной информации в виде публикаций, отчетов, изображений и других типов. Создание таких систем – важная, актуальная и сложная задача. В своем развитии они прошли несколько этапов, основанных на технологиях развития средств и методик для разработки порталов знаний (ПЗ), интеллектуальных научных интернет-ресурсов (ИНИР), интеллектуальных информационно-аналитических интернет-ресурсов (ИИАИР). Подобный подход была успешно использован в таких областях, как активная сейсмология [3], теплофизические свойства веществ [2], поддержка принятия решений, исследования в энергетике. Изначально ориентируясь на слабоформализованные ОЗ, на представление семантических зависимостей понятий и систематизацию разнородной информации, она дополнялась средствами обработки этой информации и решения типичных для рассматриваемой ОЗ задач. Чтобы полноценно представлять области знаний, в которых имеются большие массивы числовых экспериментальных данных, хранящихся в распределенных источниках, технология должна предоставлять средства управления этими данными. Возросший поток информации привел к необходимости использования новых способов ее хранения, представления, формализации, систематизации и автоматической обработки. Появились способы создания баз знаний (Knowledge Base – KB), позволяющих использовать их для различных практических целей [6–7]. Среди них – технологии Semantic Web (SW), содержащие гипертекстовые страницы с дополнительной разметкой. Часто в ней содержатся сведения о семантике элемен-

тов страницы. Важным компонентом SW является понятие онтологии, описывающее смысл семантической разметки.

Как правило, под онтологией понимают систему понятий некоторой предметной области. Она представляется как набор сущностей, соединенных различными отношениями. Создание, основанного на развитой онтологической структуре, интеллектуального научного Интернет-ресурса или портала по определенным тематикам, позволит обеспечить эффективный доступ к информации и ее последующую обработку [8]. Онтология, как основа ресурса, позволяет, помимо структуризации данных, производить первоначальную верификацию данных, исходя из заданных в ней правил [9]. Это помогает улучшить качество вносимой информации, и, в ряде случаев, исключить противоречивые факты.

Онтологии во многом похожи на тезаурусы и таксономии, но на самом деле шире их, поскольку предоставляют дополнительные средства для описания структуры описываемых данных. В центре большинства онтологий находятся *классы*, каждый из которых может иметь *подклассы*, представляющие собой более точные понятия, чем исходный класс. Все классы онтологии выстраиваются в одну или несколько *иерархий* и описывают понятия предметной области. При этом классы могут содержать *атрибуты*, которые описывают свойства и внутреннюю структуру понятий, лежащих в основе классов. Все подклассы *наследуют* атрибуты родительских классов. Каждый атрибут класса помимо названия имеет *тип значения*, *разрешенные значения*, *число значений*. Тип значения атрибута описывает, какие типы значений может содержать атрибут, например, строку или целое число. Существует также ограничение значения атрибута, состоящее в том, что он может принимать только определенные классы или экземпляры определенных классов.

Онтологии могут быть использованы везде, где требуется обработка данных, учитывающая их семантику. В силу изначальной ориентированности языка OWL на машинную обработку, правильное применение онтологий может, с одной стороны, существенно упростить и, с другой стороны, открыть новые возможности в разработке приложений, решающих задачи автоматизированной обработки и доступа к данным.

Технология разработки интеллектуальных научных Интернет-ресурсов

ИНИР представляет собой систему с Web-интерфейсом, которая содержит систематизированную информацию, относящуюся к области знаний (ОЗ), ориентированной на атмосферные аэрозоли. Она предоставляет содержательный доступ к этой информации, к методам ее обработки и методам решения задач, принятым в данной области знаний, а также к относящимся к ней Интернет-ресурсам. Основным компонентом ИНИР является ***онтология ОЗ***. На её основе осуществляется систематизации информации конкретной ОЗ и функционирование ИНИР.

Технология разработки предоставляет:

- методику построения онтологии,
- набор базовых онтологий,
- оболочку ИНИР,
- средства спецификации пользовательского интерфейса,
- редактор данных,

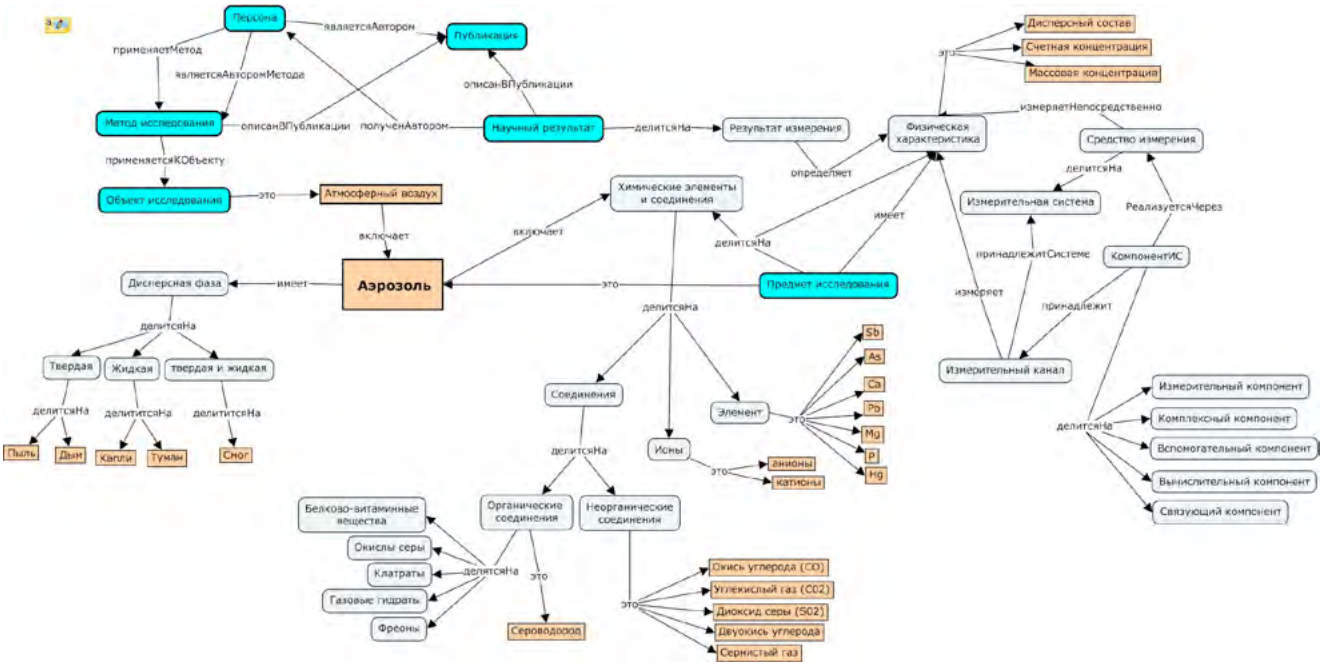


Рис. 1. Онтология аэрозоля и средств измерений

• набор сервисов, обеспечивающих функциональность ресурса.

Для разработки онтологии используются средства CmapTools, редактор Protégé и методика, предложенная в ИСИ СО РАН технология разработки ИНИР. Методика ориентирована на использование базовых онтологий и паттернов онтологического проектирования. При построении онтологии использовалась структура понятий, основанная на последовательности ***Аэрозоль => Измерения => Исследователь => Публикации.***

Онтология измерительных систем

При измерении любой физической величины важно знать не только сам результат измерения, но и его пространственно-временную привязку, условия, в которых производились измерения, характеристики измерительных приборов и т.п. В том или ином виде любая система, производящая измерения, хранит такие метаданные наряду с результатами измерений.

Процесс измерения физических величин хорошо изучен и регламентирован в рамках науки метрологии. Основные ее положения изложены в стандартах и нормативных документах, поэтому на этапе моделирования предметной области мы опираемся на них. ГОСТ Р 8.596-2002. Метрологическое обеспечение измерительных систем [ГОСТ]{GOST_8.596} определяет измерительную систему, как совокупность компонентов (измерительных, связующих, вычислительных), образующих измерительные каналы. Измерительный канал – это логическая сущность, объединяющая весь комплекс измерений и преобразований по получению результата измерения заданной физической величины.

На основе данного метрологического стандарта была разработана онтология измерительных систем. Общая схема этой онтологии представлена на рисунке 2.

На рисунках 3-4 представлен пользовательский Web-интерфейс ИНИР «Аэрозоли в атмосфере» [6] в двух видах: табличное представление и графическое.

В левой области страницы показан верхний уровень иерархии понятий онтологии. В центральной части находится описание конкретного объекта измерения, являющегося экземпляром выбранного в онтологии понятия. Этот объект представляет «Свойства объекта» – Аэрозоля и его связи в основе которых лежат «Химические элементы и соединения». Помимо метаинформации, описывающей собственно объекты (Анион, Двуокись углерода, Диоксид серы и др.) и содержащейся в контенте данного ИНИР, состояние этого объекта в виде дисперсной фазы: Дым, Капли, Пыль, Смог, Туман.

Важными свойствами являются измеряемые характеристики объекта измерения. Они обладают по крайней мере двумя параметрами: наименование измеряемой характеристики (концентрация какого-либо вида аэрозоля) и его размерность (например, мг/м^3). Значения массовой концентрации субмикронной фракции атмосферных аэрозолей в виде экспериментальных числовых рядов данных, хранятся во внешней системе хранения данных. Там же может храниться и информация, например, о пространственно-временной изменчивости концентраций органического и неорганического углерода, диоксида азота. Таким образом у пользователей ИНИР есть потребность не только просматривать внешние данные непосредственно в ИНИР, но и использовать их для решения задач ОЗ ресурса.

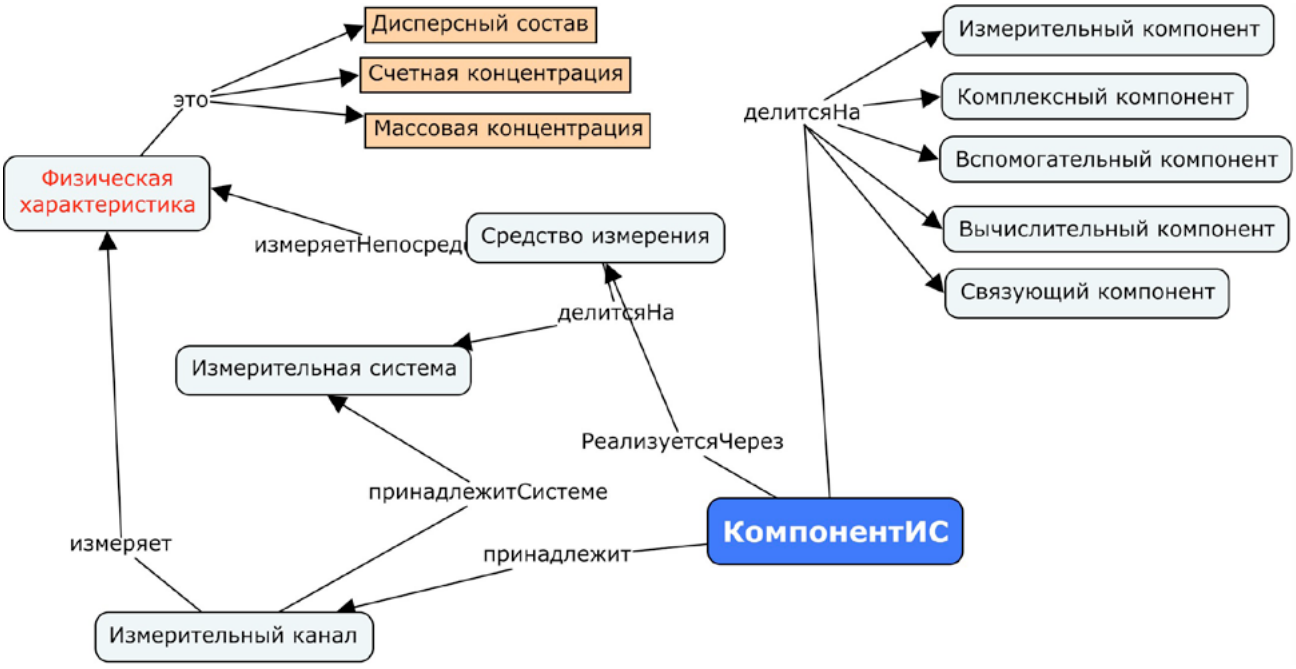


Рис. 2. Онтология измерительных систем

Рис. 3. Онтология атмосферного аэрозоля (табличное представление)

Рис. 4. Онтология атмосферного аэрозоля (графическое представление)

Поэтому были созданы специальные средства, позволяющие работать с данными из внешних источников.

Архитектура системы доступа к внешним данным

Для доступа к внешним данным была разработана система, предоставляющая пользователям ИНИР следующие функциональные возможности:

1. Организация взаимодействия с внешними источниками данных. Это могут быть сторонние базы данных (БД), либо БД, созданные разработчиками конкретных ИНИР.
2. Описание информационных объектов со значениями из внешних БД.
3. Импорт значений свойств указанных объектов из внешних источников.
4. Визуализация значений свойств объектов в виде таблиц или графиков.
5. Запуск сервисов анализа импортированных данных.
6. Использование импортированных данных для решения задач.

Основным компонентом данной системы является сервис загрузки данных – *Загрузчик*, осуществляющий непосредственное взаимодействие с внешними хранилищами. Для подключения к системе конкретных БД служит *Панель администратора*, имеющая пользовательский Web-интерфейс, который позволяет зарегистрировать новые источники данных и сформировать шаблоны запросов для доступа к ним. Загрузчик имеет собственную БД, которая содержит адреса зарегистрированных в нем БД и информацию, требующуюся для построения запроса к конкретному ресурсу. Здесь могут использоваться как шаблоны SQL запросов к реляционным БД, так и иные форматы запросов (REST API, SOAP, SPARQL и другие) к внешним ресурсам. Для построения конкретных запросов Загрузчику передаются необходимые параметры, которые извлекаются из онтологии ИНИР.

Сервисы для работы с данными позволяют показывать их пользователю, выполнять их анализ или использовать для решения задач ОЗ ресурса.

Для организации взаимодействия ИНИР с Загрузчиком и Сервисами для работы с данными был разработан специальный плагин – *Менеджер*. Данный плагин предназначен для извлечения из онтологии ИНИР параметров, необходимых Загрузчику для построения запроса к внешней БД. Менеджер передает Загрузчику параметры, получает от него идентификатор запроса, который затем передает необходимому Сервису.

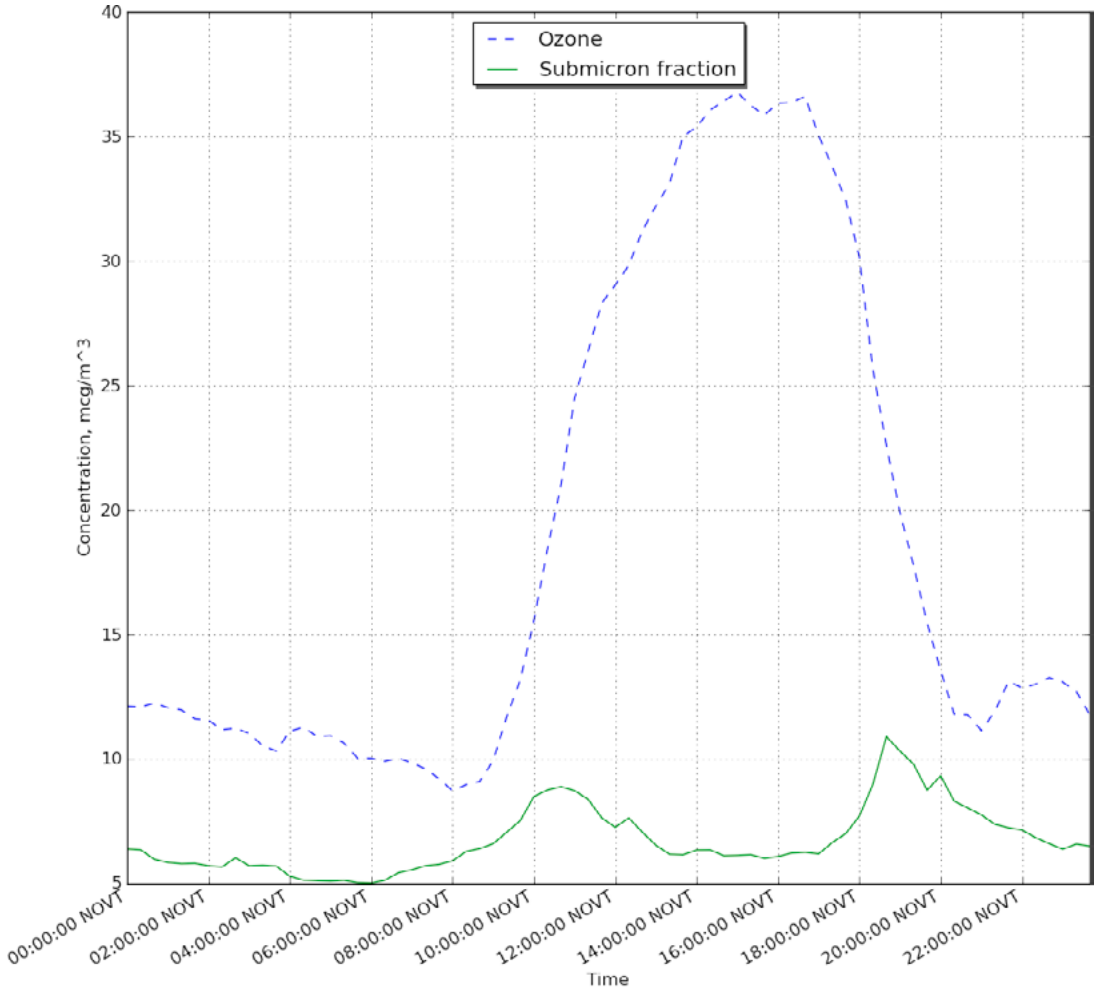
Рассмотрим схему функционирования системы доступа к внешним данным. Для того чтобы иметь возможность использовать в ИНИР внешние данные, необходимо предварительно через Панель администратора зарегистрировать в Загрузчике БД и шаблон запроса, делающего выборку необходимых данных. При этом каждой тройке (шаблон, строка подключения, тип БД) присваивается уникальный идентификатор, который сообщается инженеру знаний. Инженер знаний должен определить в онтологии ОЗ класс объектов, свойства которых будут принимать значений из внешней БД, и связать с этим классом полученный идентификатор. Кроме того, он должен позаботиться о том, чтобы связать свойства таких объектов с параметрами шаблона запроса к внешней БД (их именами и порядком следования в шаблоне).

При отображении в ИНИР свойств объектов, принадлежащих таким классам, появляется атрибут «Значение» с гиперссылкой «>>>». Переход по этой ссылке инициирует работу Менеджера, которому передается имя отображаемого объекта. По имени объекта Менеджер делает запрос к онтологии ИНИР, результатом которого будут связанный с классом данного объекта идентификатор и набор параметров шаблона запроса с их именами, значениями и порядком вхождения в запрос. Идентификатор и набор параметров передаются Менеджером Загрузчику. Загрузчик записывает эту информацию во временную кэш-память и также присваивает каждому такому новому запросу уникальный идентификатор. В зависимости от того, в какой момент работы ИНИР и для каких целей вызывался Менеджер, он передает идентификатор запроса определенному Сервису, способному обратиться к Загрузчику, и предъявив идентификатор запроса, получить и обработать данные, полученные от Загрузчика.

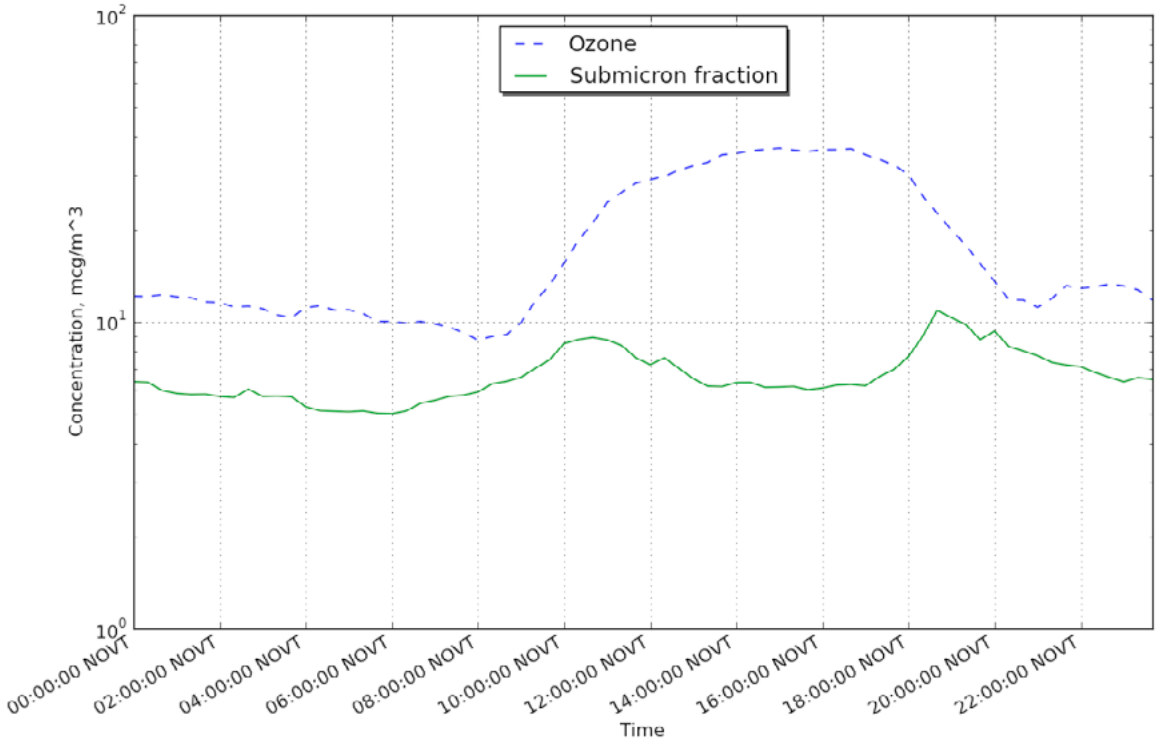
Для ресурса «Атмосферный аэрозоль» был разработан сервис визуализации данных (Визуализатор), который позволяет отображать зависимости некоторых теплофизических параметров, полученных в разное время разными авторами. Результат работы визуализатора показан на примете изменения синоптические циклов и их сезонной изменчивости. На 5 (a, b) представлены синоптические циклы и сезонная изменчивость массовой концентрации субмикронной фракции атмосферного аэрозоля (АА) и озона за 4 месячных серии наблюдений, соответствующие зимнему (январь-февраль), весеннему (апрель-май), летнему (июнь-июль) и осеннему (сентябрь-октябрь) сезонам в линейном рисунке 5 (a) и логарифмическом рисунке 5 (b) масштабах. На них четко виден существенный рост концентрации озона после восхода Солнца и падение до уровня исходного после захода Солнца. В свою очередь, для субмикронной фракции АА наблюдается значительный кратковременный рост концентрации после восхода Солнца и перед его заходом!

Выводы

В рамках технологии создания интеллектуальных научных интернет-ресурсов была разработана система доступа к данным из внешних источников. Использование технологии и средств Semantic Web



a)



b)

Рис. 5 (a, b). Визуализация данных из внешних источников

позволило максимально упростить установление связи между объектами контента ИНИР и значениями их свойств, хранящимися во внешних БД.

В основе разработки оболочки ИНИР был использован сервис-ориентированный подход. Согласно этому подходу, вся функциональность ИНИР реализуется с помощью сервисов – локальных либо распределённых, слабо связанных, заменяемых компонентов, оснащённых стандартизированными интерфейсами для взаимодействия по стандартизированным протоколам. Такой подход позволяет разработчикам ресурса создавать различные сервисы для обработки информации, хранящейся как в контенте ИНИР, так и во внешних хранилищах, а также использовать сервисы сторонних разработчиков

В качестве протокола общения между компонентами системы была использована библиотека Protobuf компании Google. Это позволило централизованно управлять протоколом общения компонент, написанных на разных языках и оптимизировать трафик, за счёт сжатия передаваемых данных. Сервисы загрузки и визуализации написаны на фреймворке django языка Python с использованием Javascript библиотек React и Plotly. Это позволило построить мощный клиентский интерфейс пользователя и получать динамический масштабируемый график. Менеджер реализован на языке PHP с использованием пакета EasyPHP для формирования запросов к онтологии.

Благодарности

Исследования выполнены при частичной поддержке грантов РФФИ № 16-07-00569, № 18-07-01457, Инт.П СО РАН № АААА-А18-118022190008-8 и ПФНИ № АААА-А17-117120670141-7

Список литературы

1. Загорулько Ю.А., Загорулько Г.Б., Боровикова О.И. Технология создания тематических интеллектуальных научных интернет-ресурсов, базирующаяся на онтологии // Программная инженерия. 2016. № 2. С. 51–60.

2. Л. Брагинская, В. Ковалевский, А. Григорюк, Г. Загорулько. Онтологический подход к информационному обеспечению исследований в активной сейсмологии. Вторая российско-тихоокеанская конференция по компьютерным технологиям и приложениям (RPC-2017). Цифровая библиотека IEEE Xplore, 2017, 27–29.

3. Загорулько Г.Б., Молородов Ю.И., Федотов А.М. Систематизация знаний по теплофизическим свойствам веществ // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2014. – Т.12. – № 3. – С.48–56.

4. Мониторинг качества атмосферного воздуха для оценки воздействия на здоровье человека. Копенгаген: Региональные публикации ВОЗ // Европейская серия. 2001. № 5. С. 293.

5. ГОСТ 17.2.1.03-84. Охрана природы. Атмосфера. Термины и определения контроля загрязнения.

6. Guariano N., Giaretta P.Ontologies and Knowledge Bases. Towards a Terminalogical Clarification //Towards Very Large Knowledge Bases.1995-N.J.I.Mars (ed.) IOS Press, Amsterdam.

7. Thomas R. Gruber. Towards Principles for the Design of Ontologies Used for Knowledge Sharing // International Workshop on Formal Ontology. 1993. March, Padova, Italy.

8. Н.М. Боргест Критериальный анализ предметной области – ключевая проблема в онтологии проектирования // Информационные технологии и системы [Электронный ресурс]: тр. Шестой Между-нар. науч. конф., Банное, Россия, 1–5 марта 2017 г. (ИТиС–2017). – С. 28–30.

9. Пименова Н.В. «Терминология предметной области, особенности ее формирования и использования для семантической классификации документов.» http://www.sei.irk.ru/sei34/main_programma.htm

10. Rogushina J., Gladun A. Ontological Approach to Domain Knowledge Representation for Informational Retrieval in Multiagent Systems // International Jornal “Information Theories & Applications”, 2006. – V.13, N.4. – P. 354–362.

11. Protégé. A free, open-source ontology editor and framework for building intelligent systems – URL: https://protege.stanford.edu / (last visited 22.05.2018).

УДК 519.816

Д.Р. Богданова, В.А. Котельников, Н.И. Юсупова

МОДЕЛИ И МЕТОДЫ ОЦЕНКИ ЭФФЕКТИВНОСТИ ПРОЦЕССА ОКАЗАНИЯ УСЛУГ В СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМАХ

Аннотация

В работе рассматривается система оказания услуг населению по оплате платежей как социально-экономическая система имеющая два уровня клиентов – население и провайдеры, обеспечивающая проведение услуг по оплате платежей. Анализ особенностей системы позволил выявить факторы, необходимые при оценке эффективности процесса оказания услуг по оплате платежей с выделением их параметров управления и показателей эффективности. Функциональные модели процесса оказания услуг на основе применения системной методологии показывают особенности структуры оказываемых услуг. Предложены модели и методы оценки эффективности процесса оказания услуг на основе использования нечеткой логики, позволяющие учесть качественные и количественные характеристики. Для оценки характеристик используется пенташкала оценок и характеристики представляются в виде нечетких Т-чисел. Для множества качественных характеристик набор функций принадлежности задается стандартно по выбранной пенташкале. Для классификации множества количественных характеристик экспертам необходимо определить пятерку нечетких Т-чисел которые определяют набор функций принадлежности для каждой характеристики, т е задать узловые точки трапециевидного нечеткого числа. В исследовании приводятся результаты применения предлагаемого подхода на примере системы моментальных платежей в рамках Башкирской платежной сети.

Ключевые слова: социально-экономические системы, процесс оказания услуг, оценка эффективности, нечеткая логика, система моментальных платежей.

Введение

Оценка эффективности в социально-экономических системах является сложной задачей. Для ее решения требуется формализации внутренних закономерностей системы и четкое представление ее параметров. Социально-экономические системы – это совокупность взаимосвязанных социально-экономических субъектов и отношений производства и распределения товаров и услуг. Сфера оказания услуг может быть рассмотрена как социально-экономическая система, в которой оценка эффективности услуг зависит от субъективной составляющей – клиента.

В статье рассматриваются вопросы разработки моделей и методов оценки эффективности процесса оказания услуг с использованием нечеткой логики. В качестве реального примера выбран процесс оказания услуг по оплате платежей населением в системе моментальных платежей.

Степень проработанности темы исследования

Вопросами эффективности сферы услуг занимаются многие авторы. Н. Лукашевич, Д. Гаранин и др. [1] рассматривают проблему эффективности процесса оказания финансовых услуг для малого бизнеса с применением оценки финансовых рисков этого процесса. Некоторые исследователи в частности А. Винклер, Б. Хорват и соавторы [2] обратили внимание на эффективность принятия решений в сфере услуг общественного транспорта. В их работах рассматриваются примеры того, как интеллектуальные когнитивные информационные коммуникации могут поддерживать эти решения. А. Шреста и др. [3] рассматривают эффективность поддержки принятия решений менеджерами ИТ-услуг. Важным вопросом является вопрос об эффективности оказания услуг в электронных платежах. Такие авторы как С. Катамадзе, Н. Топуриа и др. [4], М. Паскует и С. Гербаих с соавторами [5], Д. Пал и др. [6], рассматривают новые технологии, которые стимулируют инновации в оплатах и повышают эффективность мгновенных платежей. В. Яокума и др. [7] исследовали предпочтения клиентов в платежных системах. В их работах оценивается эффективность оказания услуг по оплате на основе оценки степени удовлетворенности клиентов при использовании услуг электронных платежей. В работах [8–17] авторы формализуют неопределенность в условиях функционирования социально-экономических систем с помощью инструментов нечеткой логики при решении экономических задач.

Анализ особенностей объекта исследования

В системе моментальных платежей (СМП) предоставляется услуга в виде доступного и удобного способа оплаты счетов за потребленные товары и услуги. СМП выступает посредником между населением с одной стороны и провайдерами услуг и продавцами товаров – с другой. У каждой из сторон, для которой система предлагает свои услуги, есть требования, которые определяют параметры системы [18].

Выделим в СМП в качестве управляемых параметров для населения ассортимент услуг, скорость оказания услуг, наличие тех. поддержки, % комиссии, минимальная сумма платежа. Для провайдеров в качестве таких параметров могут быть скорость под-

ключения услуги, % вознаграждения, требования к шифрованию, разрешение конфликтных ситуаций, быстрота передачи данных («толстый» канал связи). В качестве показателей эффективности процесса оказания услуг можно выделить для населения снижение времени проведения платежа, минимальной суммы платежа, % комиссии, работа call-центра; для провайдеров – время подключения услуги, снижение % вознаграждения, увеличение канала связи, эффективная работа call-центра, быстрое устранение неполадок [19].

Для разработки моделей оценки эффективности при оказании услуг по оплате платежей необходимо определить все этапы процесса оказания услуг, входы и выходы системы. Построенные на основе методологии системного моделирования функциональные модели процесса оказания услуг СМП представлены на рисунках 1 и 2 [20].

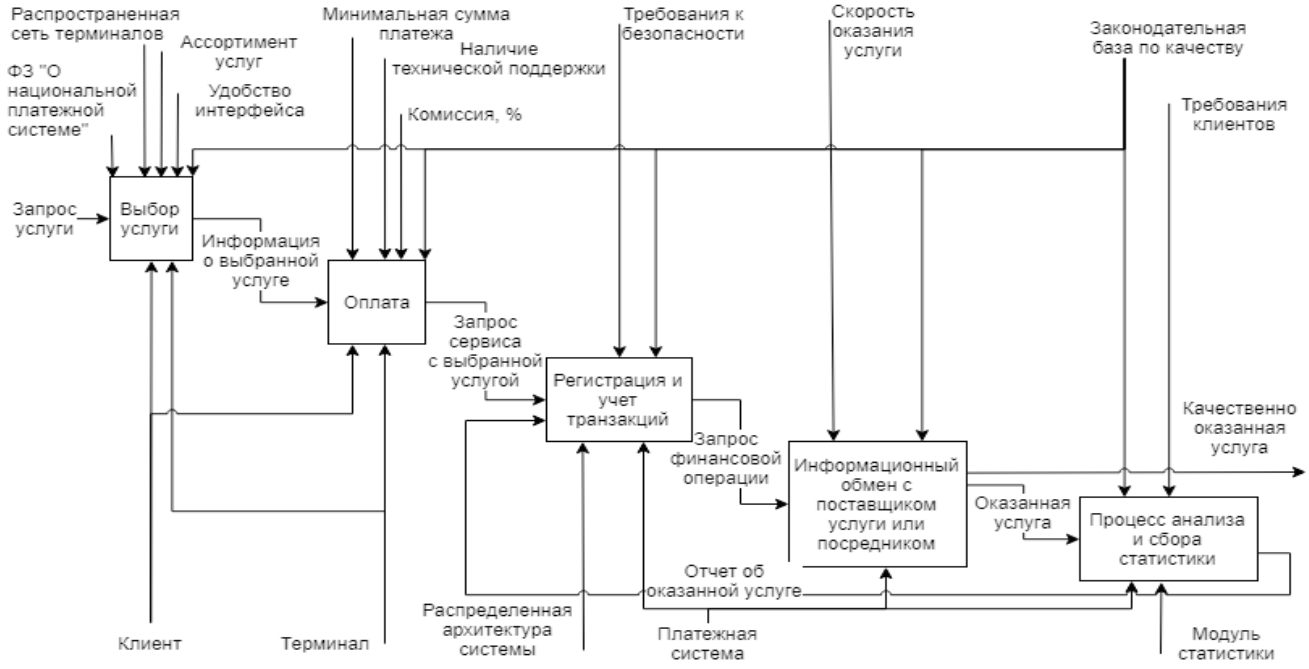


Рис. 1. Функциональная модель процесса «Оказание услуг населению»

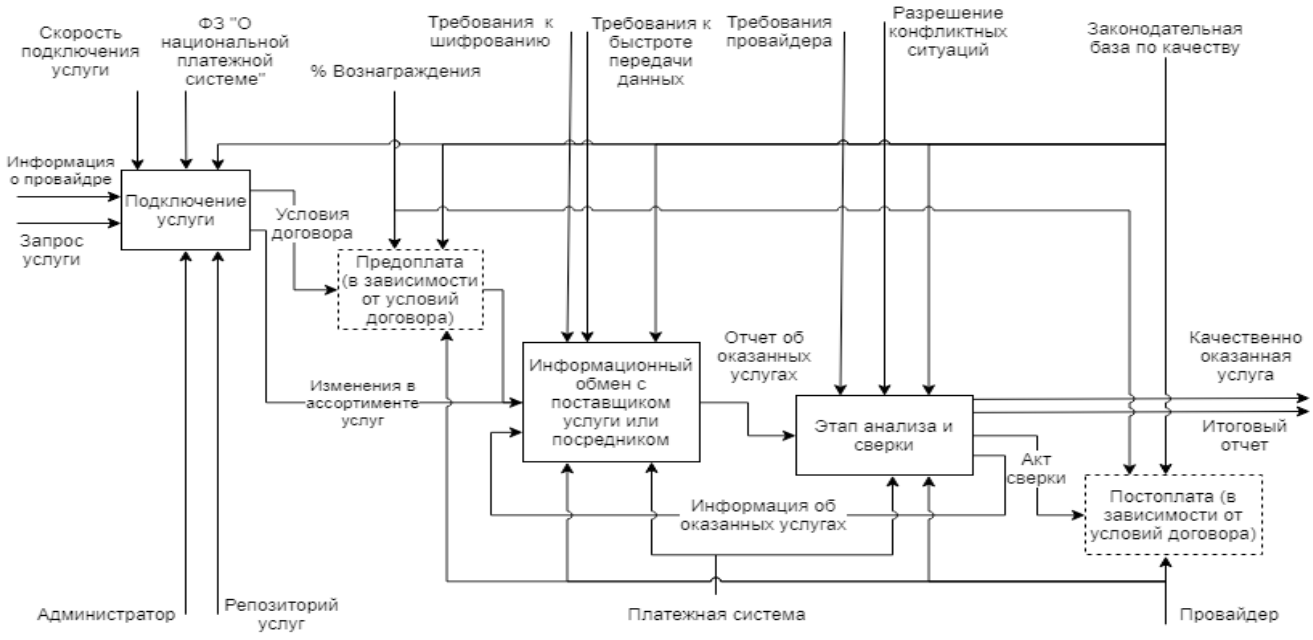


Рис. 2. Функциональная модель процесса «Оказание услуг провайдерам»

Данные модели позволяют отразить влияние формализованных параметров на разных этапах процесса оказания услуг, что дает наглядное представление внутренних свойств системы.

Разработка метода оценки эффективности

Структура услуг в рассматриваемой системе разная и параметры внутри процессов существенно отличаются, это накладывает на разрабатываемые модели и методы требования по универсальности относительно особенностей оцениваемых услуг. Показатели эффективности процесса оказания услуг по своей природе, как правило, качественные т.к. вытекают из оценок степени удовлетворённости клиентов [21]. Потребителям услуг на интуитивном уровне понятна их качественная оценка. Для разработки моделей и методов оценки эффективности процесса оказания услуг предлагается использовать нечеткую логику.

Пусть процесс оказания услуг характеризуется значениями показателей: $X=\{X_1, X_2, \ldots, X_n\}$, $X = X^v \cup X^q$, где X^v– множество количественных показателей оценки услуг; X^q – множество качественных показателей оценки услуг; N – количество показателей. Рассматривается задача определения интегрального показателя эффективности процесса оказания услуг $Q(X)$.

Для решения указанной задачи предлагается метод оценки эффективности процесса оказания услуг, основные этапы которого представлены на рисунке 3.

Каждой характеристике X_i поставим в соответствие уровень значимости r_i. Систему оценок значимостей $(r_1,\ldots,r_N)$ целесообразно пронормировать. Для оценки уровня r_i используются следующие правила:

1) Если система предпочтений характеристик отсутствует тогда:

$$r_i = 1/N, \text{ для любых } i \qquad (1)$$

2) Если имеются приоритеты в структуре характеристик, то характеристики X_i сравниваются попарно по отношению к их значимости. Система парных сравнений приводит к обратно симметричной матрице приоритетов. Элементами этой матрицы $a_{i,j}$ является значимость характеристики X_i относительно характеристики X_j, оцениваемая по шкале от 1 до 9, предложенной Т. Саати [22], где «1» – равная важность; «3» – умеренное превосходство одного над другим; «5» – существенное превосходство одного над другим; «7» –значительное превосходство одного над другим; «9» – очень сильное превосходство одного над другим; «2», «4», «6», «8» – соответствующие промежуточные значения.

Если при сравнении одной характеристики X_i с другой X_j получено $a_{i,j} = b$, то при сравнении второй характеристики с первой получаем $a_{j,i} = 1/b$. Значимость характеристики X_i определяется оценкой соответствующего ей элемента собственного нормализованного вектора матрицы приоритетов

$$A^*_j = (1^*a_{ij}^*\ldots^*a_{iN})^{1/N}, \text{ где } i=1,\ldots,N,\ j=1,\ldots,N,\ rj = \frac{A_j}{\sum_{i=1}^{N} A_i}. \qquad (2)$$

Пенташкала оценки характеристик оказываемых услуг имеет следующие значения «очень низкий», «низкий», «средний», «высокий», «очень высокий» [9]. Лингвистическая классификация характеристик

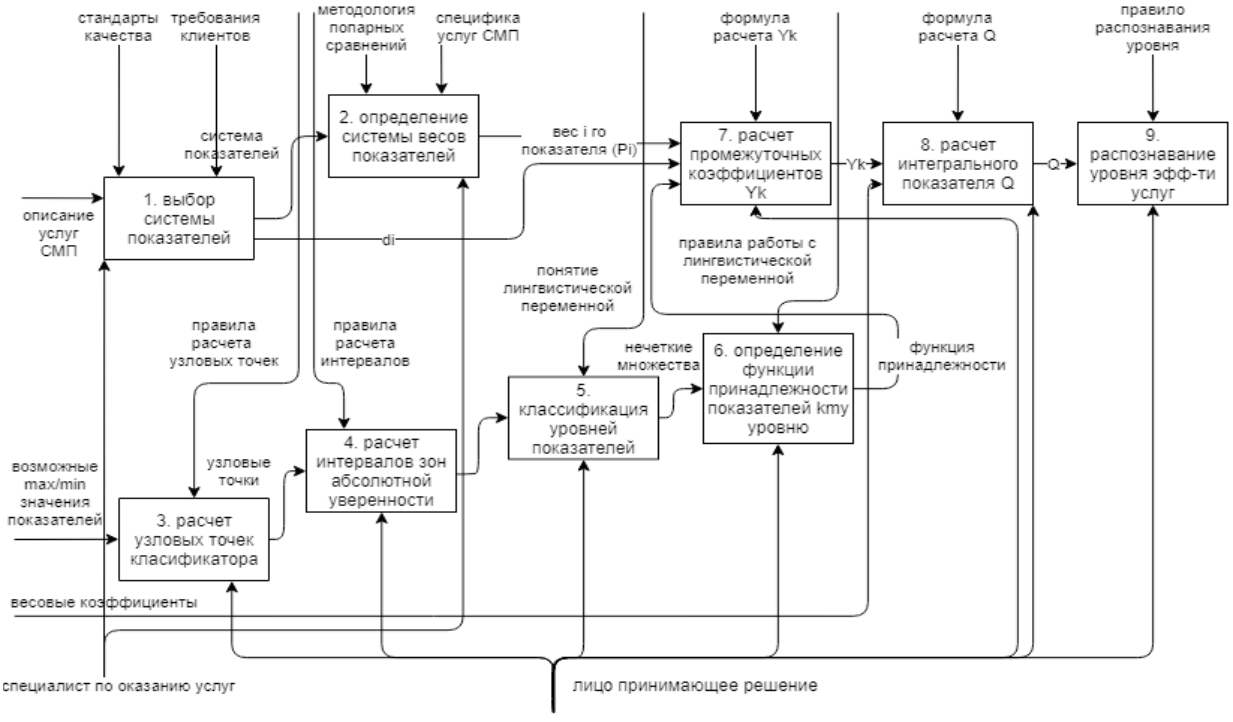


Рис. 3. Этапы метода оценки эффективности процесса оказания услуг

X_i проводится ЛПР в диалоге со специалистом по знаниям в сфере оказания услуг. Весовые коэффициенты для уровней значимости пенташкалы p_k, $k=1,...,5$ оцениваются экспертным путем.

Для множества качественных характеристик X^q набор функций принадлежности задается стандартно по выбранной пенташкале. Для классификации множества количественных характеристик X^v экспертам необходимо определить пятерку нечетких Т-чисел $b_k=(c_1,c_2,c_3,c_4)$, $k=1,...,5$ которые определяют набор функций принадлежности $\lambda_{k,i}$, $k=1,...,5$, для каждой характеристики $X_i \in X^v$, где c_1,c_2,c_3,c_4 – это узловые точки трапециевидного нечеткого числа. Всему набору функций принадлежности $\lambda_{k,i}$ по каждому показателю X_i отвечает система Т-чисел. Тогда промежуточные коэффициенты имеют вид:

$$Y_k = \frac{\sum_{i=1}^{N} \delta_i r_i \lambda_{ki}}{\sum_{i=1}^{N} \delta_i r_i} \le 1, \text{ где } k=1,..,5 \quad (3)$$

$$\delta_i = \begin{cases} 1, \text{ если рост } X_i \text{ сопровождается ростом Q} \\ -1, \text{если рост } X_i \text{ сопровождается падением Q} \end{cases}.$$

Формула расчета интегрального показателя эффективности процесса оказания услуг тогда имеет вид:

$$Q(X) = \sum_{k=1}^{5} p_k Y_k \quad (4)$$

В соответствии с методологией предложенной в [9] для принятия решения об уровне эффективности оказания услуг предлагается правило классификации представленное в таблице 1.

Результаты оценки уровня эффективности процесса оказания услуг обеспечивают информационную поддержку принятия решений ЛПР в качестве которого в этой системе могут быть менеджер, администратор, системный аналитик и т. д.

Результаты практического применения

Апробация предложенного метода проведена на примере процесса оказания услуг по оплате платежей населением в Башкирской платежной сети. В результате анализа процесса оказания услуг в указанной системе были выделены 9 характеристик оказания услуг:

- для населения – X_1 – ассортимент услуг; X_2 – скорость оказания услуг; X_3 – комиссия;
- для провайдеров – X_4 – объем транзакций; X_5 – количество узлов; X_6 – скорость передачи данных; X_7 – вознаграждение для провайдеров; X_8 – шифрование; X_9 – скорость подключения услуги.

Для качественных характеристик экспертами построены наборы функций принадлежности нечетких чисел. Для количественных характеристик были построены нечеткие пятиуровневые классификаторы на трапециевидных числах и заданы узловые точки пенташкалы на основе экспертного опроса. В рассматриваемом примере в СМП отсутствует приоритет характеристик. Весовые коэффициенты пенташкалы выбраны экспертами следующим образом $p=(0,1;0,3;0,5;0,7;0,9)$. Характеристики процесса оказания услуг и параметры функции принадлежности приведены в таблице 2.

Таблица 1

Классификация уровней эффективности оказания услуг

Уровень эффективности	Интервал значений	Классификация уровня эффективности	Степень оценочной уверенности (функция принадлежности)
1	$0 \le Q \le 0.15$	очень низкий	1
2	$0.15 < Q < 0.25$	очень низкий	$\mu_1 = 10 * (0.25 - Q)$
3		низкий	$1 - \mu_1 = \mu_2$
4	$0.25 \le Q \le 0.35$	низкий	1
5	$0.35 < Q < 0.45$	низкий	$\mu_2 = 10 * (0.45 - Q)$
6		средний	$1 - \mu_2 = \mu_3$
7	$0.45 \le Q \le 0.55$	средний	1
8	$0.55 < Q < 0.65$	средний	$\mu_3 = 10 * (0.65 - Q)$
9		высокий	$1 - \mu_3 = \mu_4$
10	$0.65 \le Q \le 0.75$	высокий	1
11	$0.75 < Q < 0.85$	высокий	$\mu_4 = 10 * (0.85 - Q)$
12		очень высокий	$1 - \mu_4 = \mu_5$
13	$0.85 \le Q \le 1.0$	очень высокий	1

Таблица 2

Характеристики процесса и параметры функции принадлежности

Хар-ка	Критерий разбиения по уровням				
	очень низкий	**низкий**	**средний**	**высокий**	**очень высокий**
$\mathbf{X}_1$	{0; 0; 6; 11}	{6; 11; 20; 25}	{20; 25; 37; 43}	{37; 43; 57; 63}	{57; 63; 70; 70}
$\mathbf{X}_2$	{0; 0; 17,84; 20,34}	{17,84; 20,34; 25,34; 27,84}	{25,34; 27,84; 35,34; 40,34}	{35,34; 40,34; 50,34; 55,34}	{50,34; 55,34; 60,34; 60,34}
X_3	{0;0;1;2}	{1;2;3;4}	{3;4;5;6}	{5;6;7;8}	{7;8;9;10}
$\mathbf{X}_4$	{0; 0; 5476174; 5685359}	{5476174; 5685359; 6103727; 6312912}	{6103727; 6312912; 6715543; 6908990}	{6715543; 6908990; 7295885; 7489332}	{7295885; 7489332; 7682779; 7682779}
$\mathbf{X}_5$	{0; 0; 5,8; 6,7}	{5,8; 6,7; 8,3; 9,2}	{8,3; 9,2; 10,8; 11,7}	{10,8; 11,7; 13,3; 14,2}	{13,3; 14,2; 15; 15}
$\mathbf{X}_6$	{0; 0; 55,8; 71,5}	{55,8; 71,5; 102,7; 118,4}	{102,7; 118,4; 140,2; 146,4}	{140,2; 146,4; 158,8; 165}	{158,8; 165; 171,2; 171,2}
$\mathbf{X}_7$	{0; 0; 0,2; 0,4}	{0,2; 0,4; 0,63; 0,67}	{0,63; 0,67; 0,77; 0,83}	{0,77; 0,83; 0,93; 0,97}	{0,93; 0,97; 1; 1}
X_8	Отсутствие шифрования	Применение стандартного протокола HTTPS с использованием криптографических протоколов SSL или TLS	Применение симметричного шифрования	Применение асимметричного шифрования	Применение асимметричного шифрования с системой контроля ключей
$\mathbf{X}_9$	{0; 0; 3,2; 5,3}	{3,2; 5,3; 9,7; 11,8}	{9,7; 11,8; 16,7; 19,3}	{16,7; 19,3; 24,7; 27,3}	{24,7; 27,3; 30; 30}

На основание приведенных данных и в соответствии с правилом принятия решений по оценке уровня эффективности процесса оказания услуг можно считать, что в данном случае уровень эффективности оказываемых услуг соответствует значению «высокий», т.к. показатель Q(X)=0,68 попадает десятый интервал (табл. 1).

Выводы

Рассмотрена система оказания услуг по оплате платежей населением как пример социально-экономической системы. Анализ особенностей системы выявил структуру процесса оказания услуг, позволил формализовать требования клиентов системы и оценить управляемые параметры и показатели эффективности.

Для данной системы предложены модели и методы оценки эффективности процесса оказания услуг в основе которых лежат инструменты нечёткой логики. Разработанные модели и методы являются универсальными для сферы оказания услуг, в которой многие параметры оценки несут в себе субъективную составляющую удовлетворенности клиента. Данный подход, в отличие, от классических методов сочетает в себе учет как количественных, так и качественных показателей.

Апробация предложенного подхода для СМП в рамках Башкирской платежной сети показала применимость разработанных моделей и методов. На основе оценки эффективности оказания услуг на данном предприятии, можно утверждать, что она является «высокой» в соответствии с выбранной шкалой.

Список литературы

1. Лукашевич Н.С., Гаранин Д.А. Аналитическая система поддержки принятия решений по кредитованию малого бизнеса // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Экономические науки, # 5 (251), 2016, с. 83–94.

2. Винклер А., Хорват Б., Интеллектуальные технологии поддержки принятия решений в общественном и индивидуальном транспорте // Интеллектуальные Технологии Решений, том. 11, # 4, 2017, с. 441–449.

3. Шреста А., Катер-Стил А, Толеман М. Инновационная поддержка принятия решений для управления ИТ-услугами // Журнал систем принятия решений, том. 25, #S1, 2016, с. 486–499.

4. Катамадзе С., Топурия Н., Бизнес-модель и техническое описание системы мгновенных платежей // Компьютерные науки и телекоммуникации, вып. 1 (53), #S1, 2018, с. 486–499.

5. Паскует М., Гербаих С., Мгновенная оплата в сравнении с оплатой через смартфон//Материалы Третьей международной конференции 2017 года по мобильным и безопасным услугам (Mobisecserv), февраль 2017 г., ISBN 978-1-5090-3633-2

6. Пал Д., Хетхават П., Чен Т. и Чжан Ю. Мобильные платежи на мировых рынках с использованием

биометрии и облака// Международный журнал систем связи, том. 30 (14) 2017.

7. Якума В., Кума П. и Окаи Э.С., Демографическое влияние на услуги электронных платежей // Международный журнал исследований электронного бизнеса, том. 13 (1), 2017, с. 43–63.

8. Хавас А. Социальные и бизнес-инновации: возможны ли общие подходы к измерению? // Форсайт, вып. 10, # 2, 2016, с. 58-80.

9. Недосекин А.О., Фролов С.Н. Лингвистический анализ гистограмм экономических факторов // Вестник воронежского государственного университета. Серия: экономика и управление, Воронежский государственный университет, Воронеж, № 2, 2008, с. 48–55.

10. Карпова Н.А. Применение методов нечеткой логики при оценке и прогнозировании финансовой устойчивости консолидированных групп компаний // Интернет-журнал «Науковедение» Том 7, #5, 2015.

11. Кандеа К., Филип Ф.Г. На пути к интеллектуальным платформам поддержки принятия совместных решений // Исследования по информатике и управлению, том. 25 (2), 2016, с. 143-152, ISSN 1220-1766.

12. Соломон К., Социально-системный подход к институтам: примеры из западной экономической истории // Журнал институциональных исследований. 8, #3, 2016, с. 6–34.

13. Дягель О.Ю., Горленко О.А., Дягель А.В. К вопросу о совершенствовании методики оценки риска финансовой несостоятельности на основе теории нечетких множеств // Аудит и финансовый анализ: ООО Издательство «ДСМ Пресс», Москва, # 6: 2015 с. 207–215.

14. Солодовников И.В, Рогозин О.В., Шуруев О.В, Экспертная система оценки эффективности обучения на основе математического аппарата нечеткой логики/ Качество. Инновации. Образование № 1 (17) Европейский центр по качеству, Москва, 2006 – с. 19–22.

15. Максименко О.И. Оценка интеллектуальных лингвистических систем методом нечеткой логики // Вестник московского государственного областного университета. Серия: лингвистика Московский государственный областной университет, Москва № 6–2, 2011 – с. 44–47.

16. Котляров И.Д. Сущность услуг как экономического блага // Вестник Омского университета, Серия «Экономика» № 3, 2012, с. 79–86.

17. Киселев С.В., Шакиров М.М. Содержание, характеристика и классификация телекоммуникационных услуг как объекта исследования // Вестник Казанского технологического университета Гуманитарные и социально-экономические проломы: Казанский технологический университет, 2013 с. 311–316

18. Котельников В.А, Богданова Д.Р., Яхина Э.И. Управление качеством оказания услуг финансового посредничества в системе моментальных платежей // Материалы 3-ей Международной конференции «Информационные технологии для интеллектуальной поддержки принятия решений»: сб. ст. – Т. 2., Уфа, 2015. с. 153–163.

19. Котельников В.А., Юсупова Н.И., Карамова И.И. Оценка надежности распределенной системы моментальных платежей // Материалы 3-ей Международной конференции «Интеллектуальные технологии обработки информации и управления»: сб. ст. – Т. 1., Уфа, 2015. с. 45–51.

20. Котельников В.А., Богданова Д.Р., Салахутдинова И.И. Проектирование программного модуля оценки качества услуг системы моментальных платежей // Материалы 4-ой Международной конференции «Информационные технологии для интеллектуальной поддержки принятия решений»: сб. ст. – Т. 1. Уфа, 2016. с. 105–114.

21. Котельников В.А., Юсупова Н.И., Богданова Д.Р. Поддержка принятия решений при управлении качеством оказания услуг финансового посредничества в системе моментальных платежей на основе технологий искусственного интеллекта / Информационные технологии и системы: 5-ой Международной научной конференции: сб. ст. – Банное, Россия, 2016. с. 103–115.

22. Коробов В.Б., Тутыгин А.Г. Преимущества и недостатки метода анализа иерархий // Известия РГПУ им. А. И. Герцена. Естественные и точные науки. #122, 2010, с. 108–115.

УДК 519.816 57.087

Г.Б. Загорулько

МЕТОДОЛОГИЯ РАЗРАБОТКИ ИНТЕЛЛЕКТУАЛЬНЫХ СППР И ЕЕ ПРИМЕНЕНИЕ ДЛЯ ЗАДАЧ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

Аннотация

В статье рассматривается методология разработки интеллектуальных систем поддержки принятия решений (ИСППР) для слабоформализованных предметных областей. Проводится анализ современного состояния в области средств и методологий разработки ИСППР, анализируются имеющиеся подходы. Предлагаемая методология базируется на авторской концепции комплексной поддержки разработки ИСППР, осуществляемой на трех уровнях – концептуальном, информационном и компонентном. Средством концептуальной поддержки является онтология поддержки принятия решений (ППР), информационной – информационно-аналитический интернет-ресурс (ИАИР) по ППР, компонентной – репозитарий методов принятия решений. Методология использует онтологический и сервис-ориентированный подходы, технологии Semantic Web, подходы быстрого прототипирования и гибкой разработки программного обеспечения, технологию разработки ИАИР. Описываемые в статье подходы и программные средства могут быть использованы разработчиками на всех этапах создания ИСППР от проектирования до непосредственной реализации. Возможности предлагаемой методологии демонстрируются на примере решения задач медицинской диагностики, в частности, разработки ИСППР для диагностики, профилактики и лечения элементозов. Описывается пошаговая методика построения ИСППР: постановка цели и задач, выбор методов, доработка базовой онтологии медицинской диагностики, построение на ее основе онтологии элементозов, создание ИАИР по элементозам, подключение сервиса, реализующего метод рассуждений на основе экспертных правил и сервиса, позволяющего пользователю вводить данные в систему и просматривать результат.

Ключевые слова: интеллектуальные системы поддержки принятия решений, методология разработки ИСППР, онтологический подход, сервис-ориентированный подход, Semantic Web, информационно-аналитические интернет-ресурсы, ИСППР медицинской диагностики.

Введение

Разработка интеллектуальных систем поддержки принятия решений (ИСППР) является сложной актуальной задачей. Для ее решения предложено много подходов, как в России, так и за рубежом.

В работе [1] дан ретроспективный обзор методологий разработки СППР. В работе [2] подчеркивается, что не существует наилучшего стандартного подхода или методологии для разработки СППР. Автор выделяет три основных подхода, указывая их плюсы и минусы: 1) подход, обеспечивающий полный жизненный цикл системы (SDLC); 2) быстрое прототипирование и 3) подход конечного пользователя, при котором ЛПР может сам выбирать средства для решения своих задач. Второй и третий подходы позволяют быстро получить рабочую систему. Наиболее распространенным, универсальным и практически значимым автор считает подход быстрого прототипирования, который позволяет быстро, оценить возможности полученной системы, при необходимости провести модификацию. Такой подход способствует объединению усилий лица, принимающего решения (ЛПР) и аналитика.

В последнее время много говорится о целесообразности использования методологии гибкой разработки программного обеспечения (Agile software development) для создания СППР [3]. Такие принципы этой методологии, как использование итеративной разработки, динамическое формирование требований, тесное взаимодействия разработчиков различного профиля, перекликаются с принципами быстрого прототипирования.

О методологии создания медицинских СППР упоминается в [4]. Авторы считают перспективными как разработку специализированных систем, так и использование гибких, легких сервисов, позиционируемых как отдельные приложения и интегрируемых между собой.

Об унифицированном подходе к разработке программного обеспечения СППР говорится в [5]. Данный подход подразумевает использование каркасной архитектуры программного обеспечения и следующих принципов проектирования: принцип свободы от субъективизма разработчиков, позволяющий системе быть максимально гибкой и подстраиваться под нужды и потребности ЛПР; принцип инвариантности по отношению к предметной области; принцип множественности методов, при котором СППР обеспечивает легкую и прозрачную интеграцию требуемых методов, советует пользователю, какой метод лучше применим и почему.

Популярной является методология, когда СППР строится для решения определенного класса задач, содержит множество методов и позволяет выбирать методы для решения конкретных задач. Описываемая в работе [6] экспертная СППР (ЭСППР) содержит около 50 методов принятия решений, а также базу

знаний, представленную в виде последовательности вопросов об элементах задачи принятия решения и их различных реализациях (ответах). ЭСППР является web-приложением, что делает ее доступной широкому кругу пользователей и позволяет привлекать экспертов для оценки альтернативных вариантов решения. Данная ЭСППР ориентирована на решение экономических задач.

В настоящее время широко признанным является мнение, что современная ИСППР должна быть информационной системой с развитыми средствами анализа и обработки хранящихся в них данных [7]. СППР, предоставляющие доступ через интернет, описаны в работах [8, 9].

Описываемая в статье методология реализует большинство подходов и принципов, перечисленных выше. Предлагаемые ею средства позволяют получать быстрые прототипы СППР, легко осуществлять их модификацию за счет использования оболочки информационно-аналитических интернет-ресурсов (ИАИР) и репозитария методов поддержки принятия решений (ППР). Онтология ППР предоставляет концептуальный базис и позволяет организовать взаимодействие всех типов разработчиков (экспертов, инженеров знаний, программистов). Методы ППР, входящие в репозитарий, реализованы в виде сервисов, что обеспечивает масштабируемость репозитария и позволяет легко интегрировать методы между собой. Разработанная в коллективе ИСИ СО РАН технология создания ИАИР [10] предоставляет оболочку ресурса, который будет являться каркасом будущей ИСППР, и к которому будут подключаться необходимые сервисы, обеспечивающие её функциональность.

Предлагаемые методологией средства для разработки тематических ИСППР не зависят от её предметной области (ПО), однако для облегчения представления знаний об этой области разработчики могут воспользоваться набором базовых онтологий и паттернов [11, 12], методикой построения онтологии ПО. Созданный на основе этой онтологии и оболочки ресурс будет являться информационной системой, систематизирующей знания о своей ПО и предоставляющей к ним содержательный доступ. Подключенные к ресурсу сервисы из репозитария сделают его полноценной ИСППР и позволят решать поставленные перед ней задачи.

В статье описывается данная методология, реализованные в ней подходы и её применение для решения задач медицинской диагностики.

Методология разработки ИСППР для СФПО

В основе описываемой методологии лежит концепция комплексной поддержки, которая сопровождает процесс разработки ИСППР от проектирования до непосредственной реализации и предполагает максимальное использование готовых методических

Рис. 1. Этапы создания ИСППР

и программных решений, как авторских, так и созданных сторонними разработчиками [13]. Комплексная поддержка осуществляется на трех уровнях – концептуальном, информационном и компонентном. Средством концептуальной поддержки является онтология поддержки принятия решений (ППР), информационной – информационно-аналитический интернет-ресурс по ППР (ИАИР ППР), компонентной – репозитарий методов принятия решений.

Методология основывается также на онтологическом [14] и сервис-ориентированном подходах [15], на технологиях Semantic Web [16, 17], технологии создания ИАИР [10].

В разработке ИСППР можно выделить те же этапы, что и в разработке любых интеллектуальных систем: концептуализация, формализация, реализация отдельных компонентов ИСППР, реализация ИСППР в целом. На рисунке 1 показаны этапы разработки ИСППР и подробно расписаны выполняемые на этих этапах действия.

Предлагаемые методика и программные средства упрощают и упорядочивают процесс разработки интеллектуальных СППР. Согласно предлагаемому подходу, основным компонентом разрабатываемой интеллектуальной СППР, её каркасом, на который будут нанизаны интерфейсные и функциональные компоненты, является тематический ИАИР выбранной ПО.

Построение тематического ИАИР осуществляется средствами технологии разработки ИАИР и, фактически, сводится к созданию его системы знаний, т.е. включает разработку онтологии и наполнение контента ресурса. Так как онтология составляет основу системы знаний ИАИР, важнейшим компонентом данной технологии является методика построения онтологий, поддержанная инструментальными средствами (редакторами онтологий и данных), а также представительным набором базовых онтологий и паттернов, на основе которых эксперты строят онтологии и тезаурусы тематических ИАИР.

К тематическому ИАИР в виде сервисов подключаются элементы пользовательского интерфейса, выбранные методы решения задач и, при необходимости, сервисы доступа к данным из внешних хранилищ [18].

ИАИР ППР, обеспечивающий информационную поддержку разработчиков, снабжает их информацией, позволяет выделить цели, задачи создаваемой интеллектуальной СППР, определить и опробовать необходимые для их решения методы.

Онтология выбранной ПО может строиться с использованием любого специализированного редактора, например, Protégé [19]. Она разрабатывается на основе предоставляемых методологией базовых онтологий путем их достройки и конкретизации. Для использования онтологии в ресурсе-оболочке ИСППР необходимо сохранить ее в синтаксисе OWL-Turtle и задать параметры визуализации ее элементов. Задание параметров делается на основе паттернов представления, предлагаемых пользователям технологией разработки ИАИР. Наполнение контента тематического ресурса может осуществляться средствами встроенного в оболочку редактора данных.

В случае возникновения затруднений при выборе необходимых методов разработчики могут воспользоваться интегрированной в ИАИР ППР экспертной системой (ЭС), которая в соответствии с задачами, которые должна решать ИСППР, предлагает наиболее подходящие методы и средства [20]. Например, для решения классических задач принятия решений – многокритериального выбора из ограниченного множества альтернатив, необходимы сервисы для построения анкет-опросников, проведения анкетирования и обработки его результатов. Для задач диагностики будут предложены сервисы для задания симптомов диагностируемого заболевания и машины вывода, работающие на экспертных правилах.

Выбранные для решения задач методы при необходимости настраиваются на предметную область и класс решаемых задач, прописываются протоколы и адреса для получения внешних данных. Создается сценарий решения задачи, включающий вызов сервисов, решающих подзадачи, с указанием входных и выходных данных. В соответствии со сценарием решается поставленная задача.

Применение методологии для разработки ИСППР «Элементозы»

Описанная выше методология была применена для разработки ИСППР для диагностики элементозов – заболеваний, связанных с дисбалансом в организме человека микро и макроэлементов. Основная задача данной ИСППР заключается в предоставлении доступа к систематизированной информации об элементозах, о симптомах этих заболеваний, о факторах, которые могут их вызывать, о заболеваниях, к которым приводят элементозы. Рассматриваемая система должна позволять провести диагностику для конкретного пациента, проведя его опрос и обследование и введя данные о наблюдаемых у него симптомах, ощущениях, факторах, известных заболеваниях.

Согласно методике, создание ИСППР начинается с разработки онтологии элементозов. Эта онтология строится с использованием базовой онтологии медицинской диагностики. Такие понятия этой онтологии, как «Показатель», «Средство лечения», «Схема лечения» конкретизируются с учетом специфики элементозов и использованием содержательных паттернов онтологического моделирования, описывающих заболевания, показатели, способы обследования и лечения пациентов. На рисунке 2 показаны паттерн для описания понятия «Заболевание» (а) и пример его использования для описания элементоза «Дефицит магния» (б). Онтология элементозов разработана средствами редактора Protege. На ее основе построен тематический ИАИР.

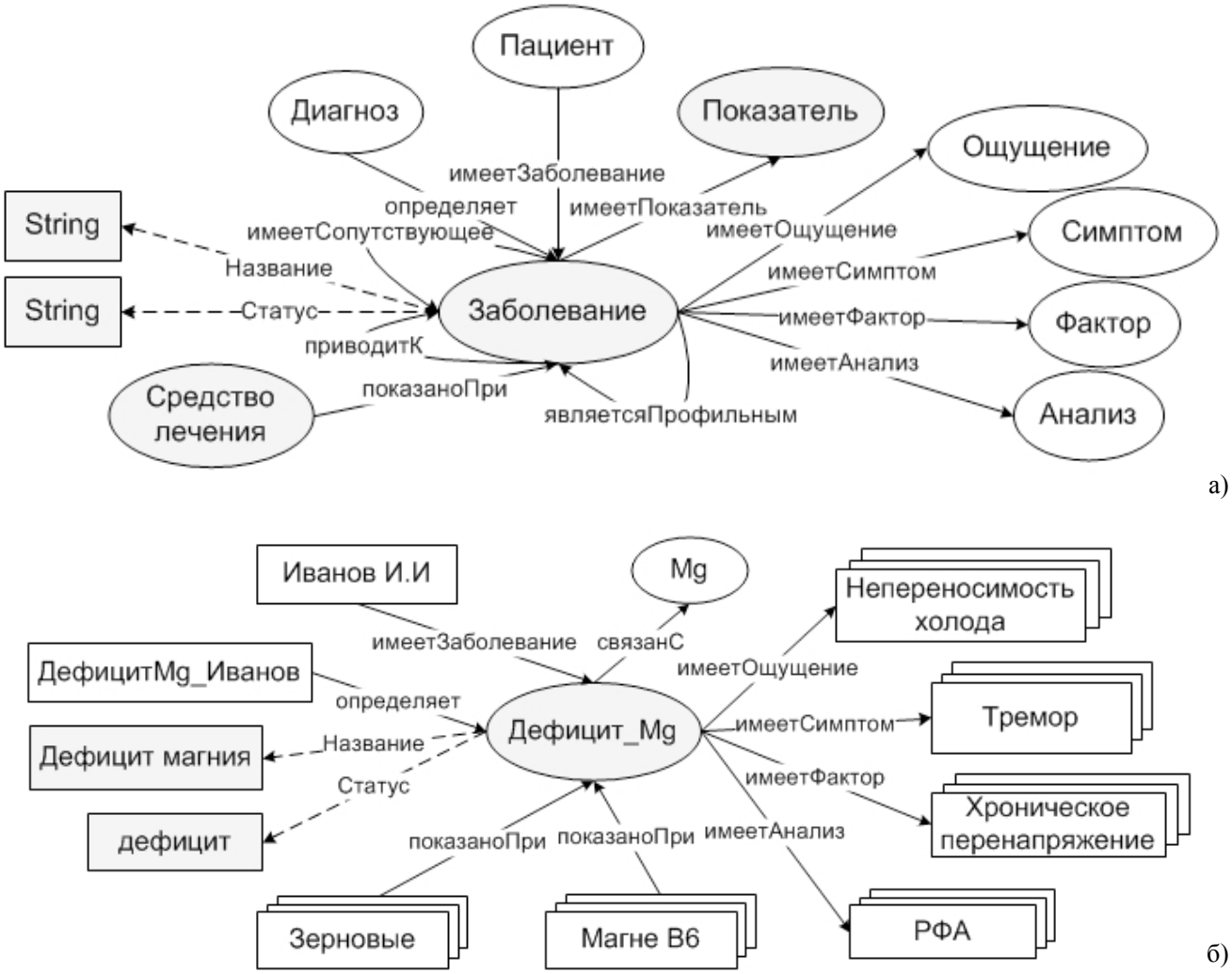


Рис. 2. а) Содержательный паттерн «Заболевание», б) пример использования

Для решения задачи диагностики был выбран метод рассуждений на основе экспертных правил. Разработана модель задачи, включающая описание предметной области, правил вывода, данных о конкретном пациенте. Для этого был использован технологический комплекс Semp-Tao [21]. Заметим, что язык представления знаний технологического комплекса Semp-Tao и язык OWL имеют сопоставимые выразительные возможности, позволяющие построить конвертор из одного представления в другое.

Для систематизации всей информации, относящейся к решаемой задаче, в контент ИАИР добавлен информационный объект «Диагностика элементозов». Отображающая его страница показана на рисунке 3. Объекты, с которыми связана данная задача, представлены на этой странице гиперссылками. Решающие задачу методы связываются с ней отношениями с атрибутами, в качестве которых выступают ссылки на инструкцию по его применению, реализующие метод сервисы, модель задачи.

Непосредственное решение задачи диагностики элементозов, т.е. ввод данных о конкретном пациенте, получение диагноза и объяснения результатов, может быть выполнено двумя способами. Пользователь может воспользоваться либо сервисом, предоставляющим удаленный доступ к комплексу Semp-Tao, либо сервисом, предоставляющим интерфейс для ввода данных о пациенте. При необходимости воспользоваться внешними данными, например, базой данных о пациентах, базой продуктов питания или лекарственных средств, к ИСППР подключается подсистема доступа к данным из внешних хранилищ [18].

Выводы

Описанная в статье методология разработки ИСППР имеет следующие особенности:

- нацеленность на слабоформализованные ПО;
- комплексность поддержки на всех уровнях и этапах разработки ИСППР;
- нацеленность на максимальное использование готовых разработок (как своих, так и сторонних);
- использование технологий Semantic Web и сервис-ориентированного подхода, авторской технологии разработки ИАИР;

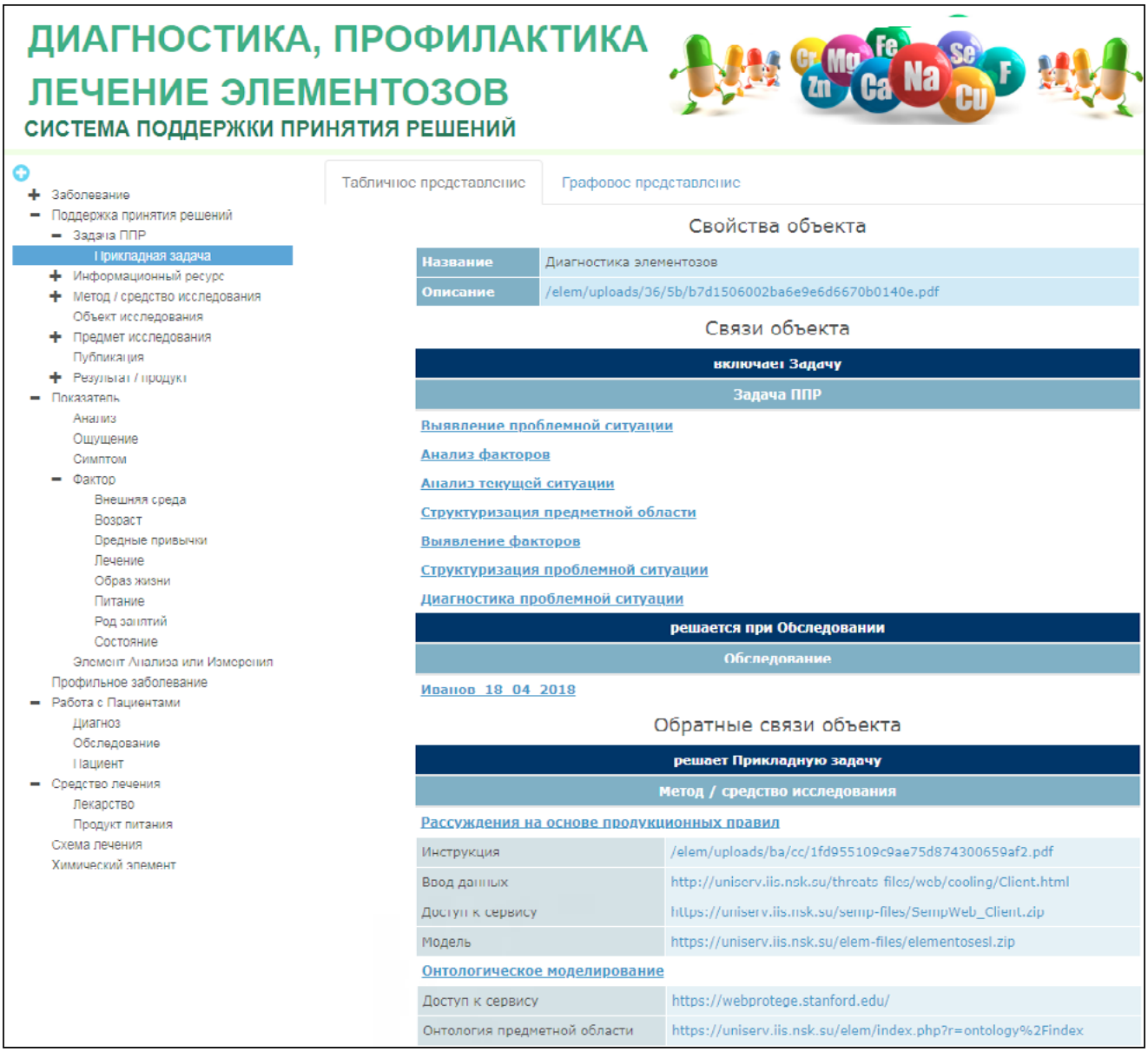


Рис. 3. Задача Диагностика элементозов

• использование оболочки ИАИР в качестве каркаса будущей ИСППР;

• открытость и масштабируемость предлагаемых средств (легко расширять как области применения, так и набор используемых методов);

• удобство и низкий порог входа для использования предлагаемых средств.

Проведенный анализ существующих подходов и средств, предлагающих решение задачи разработки ИСППР, позволяет сделать вывод, что предлагаемые методологией средства, обладая перечисленными выше особенностями, по отдельности не уступают имеющимся аналогичным разработкам, а в совокупности обеспечивают уникальную функциональность, позволяющую быстро строить и легко использовать на практике ИСППР в различных слабоформализованных областях.

Благодарности

Работа выполнена при частичной финансовой поддержке РФФИ (проект № 19-07-00762).

Список литературы

1. Garaibeh N.K., A. Al-Raddadi, 2014. Survey of DSS Development Methodologies. Encyclopedia of Business Analytics and Optimization. https://www.researchgate.net/publication/267625331_Survey_of_DSS_Development_Methodologies, Date Views 06.02.2019.

2. Power D.J., 2000. Decision Support Systems Hyperbook. Cedar Falls, IA. Date Views 07.02.2019 http://dssresources.com/dssbook/contents.html.

3. Garaibeh N.K., 2012. DSS Development and Agile Methods: Towards a newFramework for Software Development Methodology. International Journal of

Machine Learning and Computing, Vol. 2, No. 4, Date Views 06.02.2019 https:// pdfs.semanticscholar.org/d2be/d1bd3a1eb792a774145826177f2cf5dbb803.pdf.

4. Ефименко И.В., В.Ф. Хорошевский, 2017. Интеллектуальные системы поддержки принятия решений в медицине: ретроспективный обзор состояния исследований и разработок и перспективы. Материалы конференции ОСТИС, Минск, Белоруссия, февраль 2017, № 7, с. 251–260.

5. Судаков В.А., В.П. Осипов, 2014. Унификация разработки программного обеспечения прикладных СППР. Материалы Всероссийского совещания по проблемам управления, Москва, 2014, с. 8855–8863.

6. Кравченко Т.К., 2013. Развитие экспертной системы поддержки принятия решений. Искусственный интеллект и принятие решений, № 4: 72–80.

7. Халин В.Г., Г.В. Чернова, А.В. Юрков, 2015. Методологические аспекты создания и функционирования систем поддержки принятия решений. Экономический анализ: теория и практика, № 7 (406): 20–33.

8. Edwards M.P., T. Lippeveld, U. Khatri, D. Kunaka, M. Mwebaze, R. Tohouri, 2018. Building a Web-Based Decision Support System. MEASURE Evaluation. Date Views 06.02.2019 https://www.measureevaluation.org/resources/publications/fs-18-288.

9. Zhang D., X. Chen, H. Yao, 2015. Development of a Prototype Web-Based Decision Support System for Watershed Management. Water, № 7: 780–793.

10. Zagorulko Y., G. Zagorulko, 2015. Ontology-Based Technology for Development of Intelligent Scientific Internet Resources. In the Proceedings of the International Conference SoMet, Naples, Italy, September 2015, pp. 227–241.

11. Gangemi A. and V. Presutti V., 2009. Ontology Design Patterns. In Handbook on Ontologies, Eds., Staab, S. and R. Studer. Berlin: Springer Verlag, pp:221–243.

12. Zagorulko Y., G. Zagorulko, O. Borovikova, 2018. Implementation of Content Patterns in the Methodology of the Development of Ontologies for Scientific Subject Domains. Artificial Intelligence. RCAI. Communications in Computer and Information Science, vol 934: 260–272.

13. Загорулько Г.Б., Ю.А. Загорулько, 2016. Подход к организации комплексной поддержки процесса разработки интеллектуальных СППР в слабоформализованных предметных областях. Материалы конференции ОСТИС, Минск, Белоруссия, февраль 2016, с. 61–64.

14. Kopaygorodsky A., 2018. Ontology-Based Decision Support System for Forecasting of Energy Infrastructure Development. In the Proceedings of the Russian-Pacific Conference, Vladivostok, Russia, September 2017, pp. 1–5.

15. Erl T., 2005. Service-Oriented Architecture: Concepts, Technology, and Design. Prentice Hall PTR, Upper Saddle River, NJ, USA, pp: 792.

16. Hitzler P., V. Krötzsch, S. Rudolph 2009. Foundations of Semantic Web Technologies. Chapman & Hall/CRC, pp: 455.

17. Загорулько Ю.А., И.Р. Ахмадеева, А.С. Серый, В.К. Шестаков, 2016. Построение тематических интеллектуальных научных интернет-ресурсов средствами Semantic Web. Материалы конференции по искусственному интеллекту, Смоленск, Россия, октябрь 2016, Т 2, с. 47–55.

18. Molorodov Yu.I., G.B. Zagorulko, K.E.Vishnev, 2018. Tools for integrating intelligent scientific internet resources with distributed data sources. Eurasian Journal of Mathematical and Computer Applications, 6(3): 45–52.

19. Protégé. A free, open-source ontology editor and framework for building intelligent systems. Date Views 06.02.2019 http://protege.stanford.edu/

20. Загорулько Г.Б., 2018. Методические аспекты разработки интеллектуальных СППР в слабоформализованных предметных областях на основе информационно-аналитических ресурсов. Материалы конференции по искусственному интеллекту, Москва, Россия, сентябрь 2018, Т 1, с. 36–43.

21. Zagorulko Yu.A., I.G. Popov, 1998. Knowledge representation language based on the integration of production rules, frames and a subdefinite model. Joint Bulletin of the Novosibirsk Computing Center and Institute of Informatics Systems. Series: Computer Science, 8: 81–100.

УДК 519.816 510.6

Г.Г. Куликов, В.В. Антонов, А.Р. Фахруллина, Л.Е. Родионова

ПРОГРАММНО-АНАЛИТИЧЕСКИЙ КОМПЛЕКС ОПЕРИРОВАНИЯ С МНОЖЕСТВОМ РЕАЛЬНЫХ И ВИРТУАЛЬНЫХ ОБЪЕКТОВ ПО ПРАВИЛАМ ДЕКАРТОВО ЗАМКНУТОЙ ЛОГИКИ В ИНФОРМАЦИОННОМ ПРОСТРАНСТВЕ ПРЕДМЕТНОЙ ОБЛАСТИ

Аннотация

В статье исследуется проблема формализации и идентификации информационных объектов (реальных и виртуальных) и прослеживаемости их отношений в исследуемой предметной области по правилам мономорфизма и полиморфизма математической теории категорий множеств. На основе этих условий разработаны модели и методы построения программного аналитического комплекса, в соответствии с принципами стандарта ISO 15288. Показано, что структура такого программного аналитического комплекса удовлетворяет условиям декартово замкнутой логики, что значительно повышает круг решаемых задач. Это достигается путем возможного проведения эффективного реинжиниринга программного обеспечения без изменения его базового кода в реальном времени. Предложено формализованное описание отношений между виртуальными и реальными объектами в символьной форме по правилам хранилища данных. Применен метод построения иерархий Н. Хомского и метод семантических дифференциалов Ч. Осгуда. В качестве примера модифицирована структура распределенного программного аналитического комплекса управления кадровым резервом промышленного предприятия, формируемым совместно с образовательными учреждениями высшего образования.

Ключевые слова: программный аналитический комплекс, системная инженерия, принцип семантической трансформации и семантические дифференциалы, теория категорий множеств, реальные и виртуальные объекты, декартово замкнутая логика.

Введение

При решении проблем формализации знаний и описания отношений на основе теории категорий, а так же проектированием и реализацией архитектуры программного обеспечения в исследовании применены результаты трудов отечественных авторов, в том числе: в области теоретических основ и математических методов проектирования крупномасштабных информационных систем Ковалева С.П. [4], в области фрактального подхода к структурированию знаний, интеллектуальных систем и систем поддержки принятия решений Массель Л.В. [5], в области моделирование взаимодействия и самоорганизации в современных интерфейсах, сетях и средах, а так же математического и программного обеспечения моделей интеграционных процессов разнородных данных, системного проектирования и разработки – Бельтюков А.П., Антонов В.В., Куликов Г.Г. [6–8], в области моделирования задач интеллектуального управления образовательными процессами в высшей школе – Юсупова Н.И. [10], *и зарубежных* – в области моделирования и методов управления качеством на основе приложений искусственного интеллекта – Kovács G.L. [11], в области применения топосных моделей и теории категорий – Голдблатт Р. [12], в области формального и структурированного моделирования – Дж. Захман, в области методов динамического программирования – Salas D.F., Powell W.B [13].

В результате проведенного анализа определилось проблемное поле исследования – модели, методы и алгоритмы проектирования программных аналитических комплексов для организации взаимодействия программ и программных систем и инструментарии их представления – ряд мета языков: естественного языка, мета языка теории множеств и категорий множеств.

Определена структура моделей – вербальных, математических, с различными логическими ограничениями вносимыми формальными графо-аналитическими языками высокоуровневого программирования и условиями их кросс-платформенной реализации в средах web-порталов, интранет, интернет и др.

Решается подпроблема формализации задачи синонимии при идентифицируемости и прослеживаемости информационного объекта на каждом уровне соответствующего описания.

Подходы к описанию информационных объектов для моделирования программного аналитического комплекса

На сегодняшний день, особый интерес уделяется программным аналитическим комплексам с элементами интеллектуального управления, способных генерировать не только регламентные сведения и отчеты, но и обеспечивать возможность глубокого аналитического анализа лицами принимающими решения (ЛПР).

Таким образом, становится актуальным проектирование и разработка программных аналитических комплексов (ПАК), которые могут быть использованы большими компаниями, решающими задачи

управления персоналом, кадровым резервом, показатели которых, определяют в большей степени критерии эффективности.

Выявляется следующая проблема исследования – формализация и идентификация информационных объектов и поддержание соответствия их реальным объектам исследуемой предметной области.

При этом под информационными объектами понимаются объекты предметной области в информационной аналитической системе, обуславливающие его структуру, атрибуты, ограничения целостности и поведение.

Соответствие прослеживаемости информационных объектов, прослеживаемости реальных объектов, параметров для базы знаний и проектирования программного аналитического комплекса. Для определения реальных объектов, соответствия их информационным объектам и прослеживаемости на абстрактном уровне исследуемых образовательных производственных систем применим понятия и правила стандарта ISO/IEC 15288. Покажем, что положения данного стандарта полностью отвечают понятиям правил декартово замкнутой системы.

В качестве формального аппарата описания информационных объектов для моделирования, обработки и передачи данных и знаний использованы положения теории категорий и теории множеств.

Для дальнейшего рассмотрения математического описания реальных объектов и их информационных процессов введем следующие обозначения: F где – можно обозначить как множество из возможных состояний ПАК;

$q_i \in Q$, $i = 1,...,m$, где Q – конечное число состояний программной системы;

v_i, – информационный процесс, предназначенный для проектирования ПАК в определенный момент времени;

V – множество информационных процессов для построения ПАК.

Переход ПАК из одного состояния в другое показано функцией L отображения $F_0 V \rightarrow F$,

$$L: F_0 V \rightarrow F, \quad (1)$$

т.е. $l(f_i,v_i)$ будет отображать следующее состояние ПАК после выполнения этапа построения v_i и может быть представлено формулой,

$$f_{i+1} = l(f_i,v_i),\ i = 1,...,n. \quad (2)$$

Показано, что состояния ПАК зависят от состояний объектов f_i и f_j заданным множеством морфизмов, $Hom(f_i,f_j)$ образующих класс объектов, для каждой пары (морфизмов), например, $g_q \in Hom(f_i,f_j)$ и $l_q \in Hom(f_j,f_k)$, где определена их композиция $g_q \circ l_q \in Hom(f_j,f_k)$. Т.е. состояния ПАК образуют категорию множеств, состоящих из объектов, связанных категориальными отношениями.

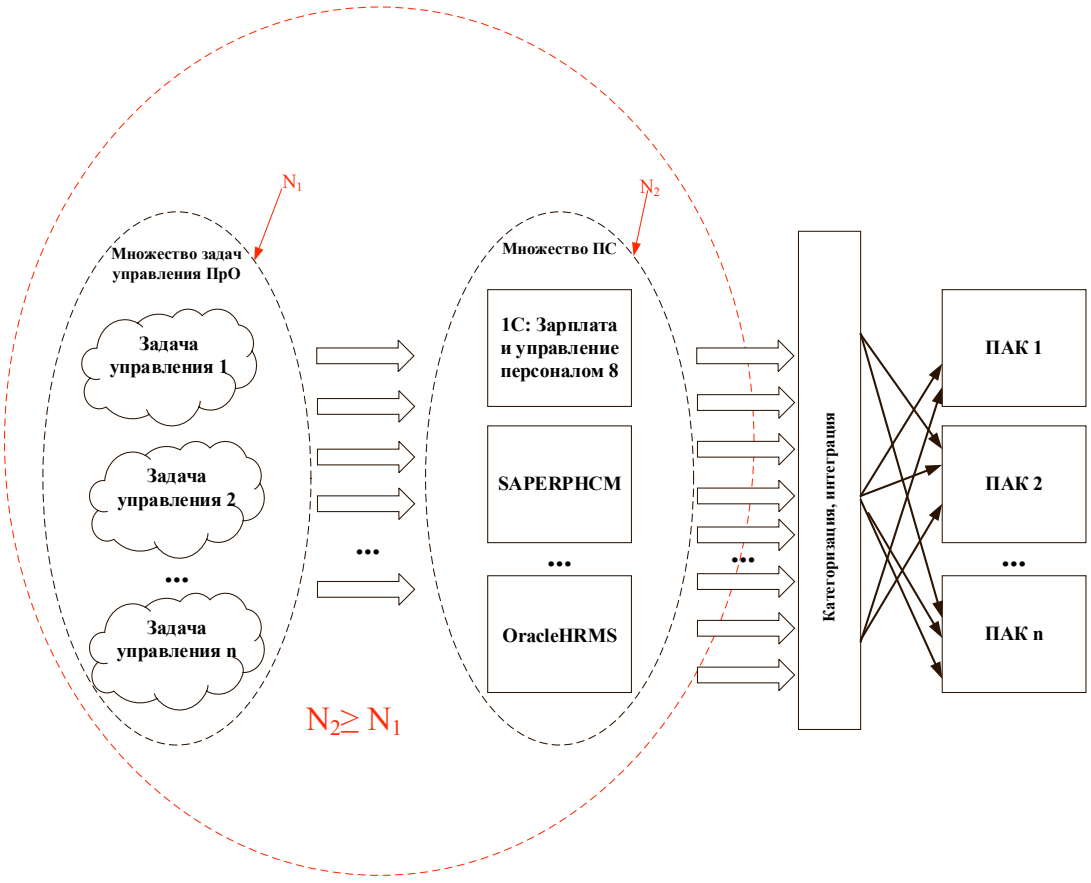


Рис. 1. Отображение модели предметной области в ПАК

Рис. 2. Основные составляющие концепции Индустрия 4.0

Проблема формирования ПАК для определенной предметной области и процесс создания перемещения семантики в синтаксис и в базу данных, являются актуальными (рис. 1).

Структура распределенного программного аналитического комплекса управления кадровым резервом промышленного предприятия

В настоящее время, в соответствии с концепцией Индустрия 4.0 современные предприятия стремятся применять цифровое производство, внедряя умные технологии. Управление происходит на основе знаний, держателями которых являются как университеты, так и предприятия, соединение таких знаний в итоге приводит к созданию образовательно-производственной структуры.

На рисунке 2 представлена концепция Индустрия 4.0, которая разделяется на три составляющих: цифровизацию и интеграцию виртуальных цепочек оценки стоимости, цифровизацию продуктов и услуг, цифровые бизнес-модели и доступ клиентов. Проводя дальнейшую детализацию приходим к совокупности следующих составляющих: мобильные устройства, облачные вычисления, дополненная реальность, персонификация по клиентскому профилю (многоуровневое взаимодействие с клиентом), анализ больших данных, интеллектуальные датчики, 3D–печать, технология определения местонахождения, интернет вещей, кибербезопасность, усовершенствованные интерфейсы взаимодействия между человеком и компьютером. Проводя их кластеризацию и классификацию, приходим к выделению 11 направлений связанных между собой отношениями.

Промышленными предприятиями накоплен достаточно большой объем данных, разнородных и зачастую плохо формализованных. Появилась проблема их использования, т.н. проблема big data.

В широком смысле по big data понимается сбор, хранение, анализ, передача, визуализация, обновление, структурированных, неструктурированных и полуструктурированных данных, характеризующихся большим объемом данных [14]. Для представления знаний об исследуемой предметной области обязательно применение формализованных графо-аналитических метаязыков, формальных логических свойств и правил.

Метод построения иерархий Хомского, на основе категорийного подхода с применением теоретико-множественных моделей, с использованием требований системной инженерии позволяет сформировать границы между семантическими и синтаксическими описаниями процессов исследуемой предметной области и функциональными модулями программного аналитического комплекса [1, 2] (рис. 3).

Таким образом, виртуальный объект – это объект, полученный в результате отношений, достраиваемый до декартово замкнутой логики, которой также не противоречит необходимому разнообразию Эшби [3].

Практическая значимость заключается в проектировании прототипа ПАК, предназначенного для управления процессами формирования и сопровождения кадровым резервом предприятия, состоящий из комплекса 1С, комплекса Prognoz Platform и интеграционного модуля ПАК.

Структура информационного ядра интеллектуального контента представляет собой совокупность информационно-аналитических систем с соответствующими приложениями, базами данных, базами знаний и др., она является открытой, поскольку может постоянно дополняться новыми структурно-связанными системами управления контентом.

Система позволяет использовать различные модули. Появляется еще один шаг к повышению уровню абстракции. Часть информации, которая была в семантике переводится в синтаксис.

Технологически ПАК организован по технологиям Web-портала и обеспечивает функционирование через

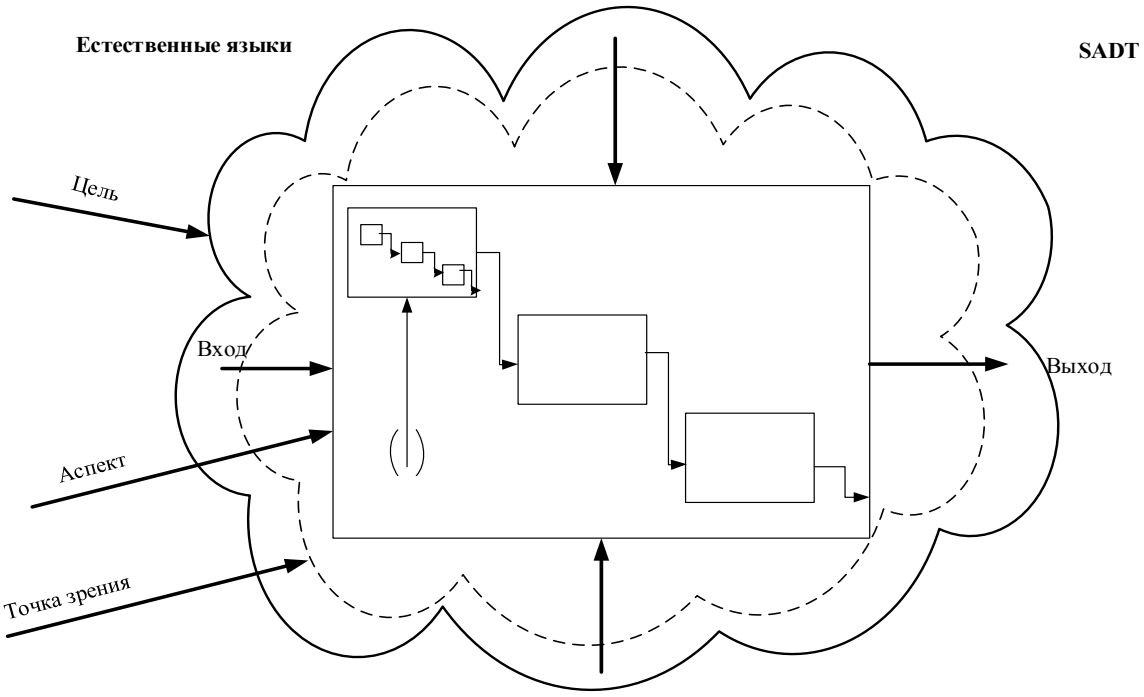


Рис. 3. Границы исследуемой предметной области

Web-браузер. Передача данных происходит в online-режиме. Весь интерфейс ПАК полностью русифицированный (в том числе, сообщения об ошибках).

ПАК обеспечивает разграничение прав доступа к функциям ПАК и защиту от несанкционированного доступа к информации в соответствии с требованиями действующего законодательства в сфере защиты информации [16].

ПАК разработан в соответствии с технологией канонического (оригинального) проектирования с применением высокоуровневых языков программированияC#, Python модули ПАК спроектированы в VisualStudio 2017 [15].

Для разрабатываемых подсистем ПАК обеспечено взаимодействие со смежными и внешними информационными ПАК посредством файлового обмена в формате XML, а также вызовами сервисов внешних систем по SOAP.

Так же в ПАК обеспечены возможности загрузки внешних данных в форматах XLS, CSV.

Фрагмент программного кода реализован в виде конвертора, который обеспечивает интеграцию данных двух программных систем, для создания аналитических отчетов и панелей (рис. 4).

При доработке ПАК предусмотрены возможности ее последующей модернизации по следующим направлениям:

- Расширение функциональности;
- Подключение дополнительных сервисов;
- Расширение перечня форматов обмена данными с внешними источниками.

Структура и состав данных ПАК определяются предметной областью.

В ПАК контролируется корректность только информации, вводимой в автоматическом режиме из сторонних официальных информационных систем, а также проверяется логическая целостность информации в базе данных при выполнении любой прикладной операции.

Пользователи взаимодействуют с ПАК с использованием графического интерфейса.

Все функции ПАК поддерживают русский язык и обеспечивают русскоязычный интерфейс пользователя, соответствующий всем нормам и правилам русского языка.

Серверные части всех подсистем рассчитаны на кросс-платформенное использование.

ПАК управления базами данных всех подсистем обеспечивает транзакционные механизмы, декларативную ссылочную целостность, позволяет наращивать производительность путем увеличения вычислительной мощности без изменений на прикладном уровне. Общие требования к СУБД:

– поддержка реляционной модели базы данных.

– поддержка технологии клиент-сервер.

– поддержка многопроцессорной архитектуры.

– наличие средств создания индексов и кластеров данных.

– автоматическое восстановление базы данных.

– наличие механизма блокировки транзакций.

– наличие встроенных средств контроля целостности баз данных.

– наличие встроенных средств резервного копирования базы данных.

– поддержка импорта и экспорта данных (XML, EDIFACT, DBF, TXT, MS Excel, Word);

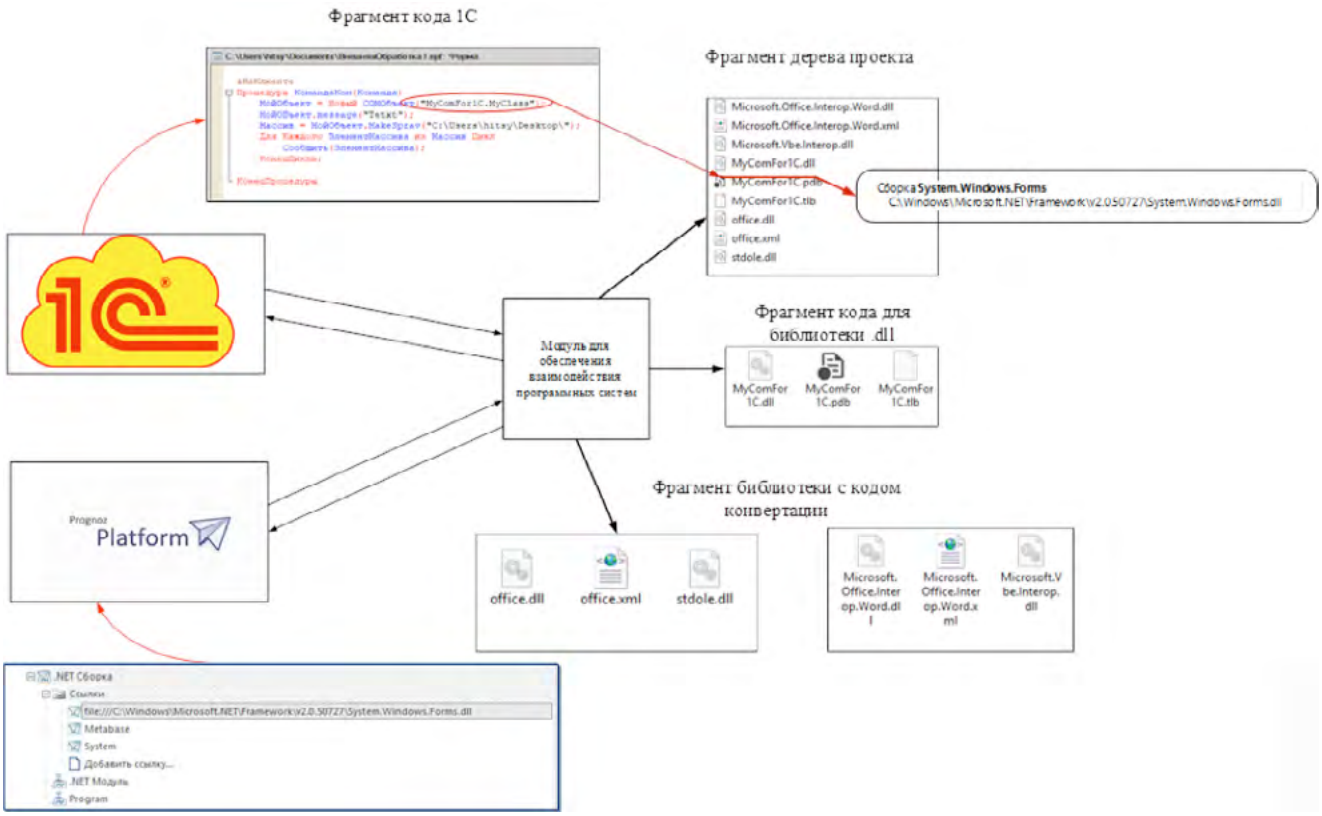


Рис. 4. Пример программной реализации ПАК

– поддержка сетевых протоколов TCP/IP;
– наличие графических средств администрирования;
– возможность контроля доступа к данным;
– централизованное управление пользователями;
– оптимизация запросов;
– наличие механизма встроенных процедур баз данных;
– наличие утилит для переноса данных из внешних систем;
– обеспечение безопасности данных на уровне сервера баз данных.

Выводы

1. Виртуальные объекты являются противоположностью понятий реальных объектов, но они пересекаются своими свойствами, т.е. свойства реальных и виртуальных объектов имеют одинаковую структуру и соответствуют функторным отношениям теории категории.

2. Как правило, виртуальные объекты имеют более широкие потенциальные свойства, по отношению к реальным объектам.

3. Так как реальные и виртуальные объекты имеют отображение в символьной форме, т.е. в информационной области, виртуальные объекты могут быть заменены реальными объектами с сохранением свойств виртуального объекта.

Следует добавить, т.к. ни все свойства виртуальных объектов могут быть реализованы в новом реальном объекте, реализуется процесс, представленный в диссертации достраиванием системы нового виртуального объекта с оставшимися нереализованными свойствами.

4. Можно показать, что отношения между реальными объектами, связаны отношением отрицания (противоположность по Декарту), а виртуальные объекты являются частными определениями в терминологии квадрата Декарта.

Новизна заключается в появлении виртуальных объектов в модели ПАК.

Список литературы

1. Хомский Н. Язык и проблема знания, 1996. Вестник МГУ, 6; 157–185

2. Осгуд Ч., Суси Дж., Танненбаум П., 1972. Приложение методики семантического дифференциалак исследованиям по эстетике и смежным проблемам. Семиотика и искусствометрия, С. 46–48.

3. Куликов Г.Г., Антонов В.В., Фахруллина А.Р., Родионова Л.Е., 2017. Формальное представление модели реализации функций системной инженерии на основе принципа необходимого разнообразия структурных связей. Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника», 17(4): 146–153

4. Ковалев С.П., 2014. Теоретико-категорный подход к проектированию программных систем. Фундамент. иприкл. матем., 19 (3): 111–170

5. Массель Л.В., 2016. Фрактальный подход к структурированию знаний и примеры его применения. Онтология проектирования, 6 (2): 149–161

6. Куликов Г.Г., Бельтюков А.П., Набатов А.Н., Шамиданов Д.Г., 2017. Метод формирования метаструктуры контента предметной области в форме локальной распределённой информационной системы. Сборник Шестой Международной научной конференции. С. 144–148.

7. Антонов В.В., 2015. Метод проектирования адаптивного программного комплекса на основе методологии категорийной формальной модели открытой предметной области. Вестник УГАТУ, 19 (1): 258–263.

8. Куликов Г.Г. В.В. Антонов, Д.В. Антонов, 2011. Теоретические и,прикладные аспекты построения моделей информационных систем. LAP LAMBERT Academic Publishing GmbH& Co.KG, Germany: 134 с.

9. Мартынов В.В., Филосова Е.И., Ширяев О.В., 2014. Организация подготовки и информационная поддержкареализации динамических образовательных программ,учитывающих требования работодателя. Сборник научных трудов Четырнадцатой международной конференции «Управление экономикой: методы, модели, технологии», 1(1): 110–113

10. Миронов В.В., Юсупова Н.И., Шакирова Г.Р., 2011. Ситуационно-ориентированные базы данных:концепция, архитектура, XML-реализация Вестник УГАТУ, 2 (37): 233–244.

11. Kovács G., Yussupova N., Rizvanov D., 2017. Resource management simulation using multi-agent approach and semantic constraints. An International Journal for Engineering and Information Sciences, 12 (1): 45–58.

12. Гольдблатт Р., 1983. Топосы. Категорный анализ логики: 486

13. Salas D.F., Powell W.B., 2018. Benchmarking a scalable approximate dynamic programming algorithm for stochastic control of grid-level energy storage. INFORMS Journal on Computing: 30(1): 106–123.

14.The Big Data Conundrum: How to Define It? [Electronic resource] // MIT Technology Review [Official website]. 3.10.2013.

15. Уэс М. Python и анализ данных [Электронный ресурс] / М. Уэс; пер. с англ. Слинкин А.А. – Электрон.дан. – Москва: ДМК Пресс, 2015. – 482 с. – Режимдоступа: https://e.lanbook.com/book/73074. – Загл. с экрана. (Дата обращения:15.08.2018).

16. Федеральный закон от 27.07.2006 № 149-ФЗ (ред. от 18.12.2018) «Об информации, информационных технологиях и о защите информации».

УДК 004.891.2

Д.А. Ризванов, Н.И. Юсупова

АЛГОРИТМЫ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ ДАННЫХ ПРИ УПРАВЛЕНИИ РЕСУРСАМИ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ И ДИНАМИЧНОСТИ ВНЕШНЕЙ СРЕДЫ

Аннотация

В статье рассматриваются алгоритмы интеллектуальной обработки данных, позволяющие повысить эффективность решения задач управления ресурсами в условиях неопределенности и динамичности внешней среды. В качестве основополагающей информационной технологии, обеспечивающей сбор, хранение, извлечение, обработку и использование данных и знаний для управления ресурсами в условиях неопределенности и семантических ограничений предметной области, выбрана технология обработки данных с использованием многоагентного подхода. Решение задачи в виде плана распределения ресурсов строится в результате ведения переговоров агентов друг с другом. Участвующие в процессе переговоров агенты определяют допустимость и приемлемость для них решения на основе индивидуальной базы знаний. Возможность самообучения агентов многоагентной системы заложена на уровне индивидуальных баз знаний, принадлежащих отдельным агентам. Новые знания, приобретенные в результате ведения переговоров при решении задачи управления ресурсами, а также на основе полученного опыта обновляют базу знаний агентов в виде соответствующих правил и ограничений, которые могут быть использованы в дальнейшем для повышения их эффективности. Апробация алгоритмов проведена на примере решения задачи управления ресурсами при календарном планировании производства.

Ключевые слова: управление ресурсами, многоагентный подход, система поддержки принятия решений, интеллектуальная обработка данных.

Введение

Проблематика управления ресурсов описана и решается в исследованиях ряда авторов. Задачи управления ресурсами как одного из разделов исследования операций исследованы в работах Р. Акоффа, Л. Берталанфи, С. Бира, Р. Беллмана, Дж. Данцига, Л. Канторовича, Т. Купманса и др.

В работе [1] рассматриваются основные подходы к управлению ресурсами при планировании производственных процессов, основанные на применении генетических алгоритмов, алгоритмов оптимизации муравьиной колонии, нейронных сетей и ряда других.

Подход, позволяющий решать задачи планирования и распределения ресурсов как задачу удовлетворения ограничений предложен в [2]. Эта проблема может быть решена либо как проблема временного ограничения, либо как проблема ограничения времени и ресурсов. Авторы предлагают новые механизмы распространения ограничений для задач управления ресурсами для временных ограничений и ограничений совместного использования ресурсов.

Для решения проблемы управления ресурсами и планирования задач при разработке программного обеспечения в работе [3] предлагают использовать оптимизационные алгоритмы муравьиной колонии. Перспективы и преимущества применения многоагентных технологий по сравнению с классическими методами планирования для управления ресурсами в промышленности и здравоохранении подробно обсуждаются и расписаны в [4–8].

В настоящее время на рынке программного обеспечения имеются системы MRP, ERP и MES для решения задач управления ресурсами в производстве [9-11], среди которых наиболее известные были разработаны SAP, Oracle, Baan, PeopleSoft, JD Ed. и т.д. Представленное программное обеспечение позволяет решать широкий спектр задач, связанных с управлением ресурсами. Они могут оперативно построить график производства и контролировать его. Однако задача управления ресурсами при составлении графика производства не полностью решена.

Проведенный анализ отечественных разработок в области IT-систем для календарного производства [17–20], позволяет выделить разработку «Smart Factory» фирмы «Разумные решения» как одно из гибких решений среди множества систем данного класса. Тем не менее, отсутствие или недостаточный учет ряда специфических особенностей предметной области, связанных с учетом предпочтений рабочих, их квалификации и мастерства, износа оборудования, делает задачу календарного планирования актуальной.

Существующие подходы не делают различий между однородными ресурсами и не используют никакой конкретной информации о проблемной среде, которая оказывает существенное влияние на распределение ресурсов и планирование.

При управлении ресурсами в сложных системах в быстро меняющейся среде необходимо предъявлять следующие требования для применения системы поддержки принятия решений:

• построение плана распределения ресурсов должно учитывать индивидуальные свойства, присущие однородным ресурсам, поскольку они существенно

влияют на конечный результат и, следовательно, на эффективность принятия решений;

• система должна быть готова к различным неожиданным ситуациям, которые нарушают существующие планы, и, при необходимости, находить и предлагать решение, поддерживать предоставление вариантов альтернативных решений, устраняющих эти эффекты, с предварительной оценкой. Например, некоторые ресурсы, используемые при изготовлении деталей (станки, оборудование), могут сломаться, и рабочие могут заболеть. Это типичные неожиданные ситуации, которые должны быть решены лицом, принимающим решения.

Задачу управления ресурсами в условиях неопределенности и динамичности внешней среды рассмотрим на примере задачи календарного планирования производства [21].

Описание задачи календарного планирования производства

Рассмотрим задачу календарного планирования производственного участка цеха. Перед началом календарного периода участку задается номенклатурный план изготовления деталей. Технология изготовления каждой детали представляет собой последовательность выполнения определенных технологических операций. На участке имеются различные ресурсы (персонал, инструмент, оборудование), необходимые для изготовления детали. Ресурсы подразделяются по категориям: фрезеровщики, токари, фрезерные станки, токарные станки и т.д. Один ресурс может относиться к нескольким категориям, например, токарь может выполнять работу фрезеровщика. Для выполнения технологической операции необходимы ресурсы различных категорий.

Календарное планирование производственного участка цеха заключается в распределении имеющихся ресурсов для выполнения производственного задания изготовления деталей. Результатом является сформированный календарный план изготовления деталей, включающий в себя информацию о том, какая технологическая операция в какое время, каким рабочим и на каком станке будет выполняться.

Применение интеллектуальных технологий для задачи управления ресурсами при календарном планировании производства

Рассматриваемая задача относится к задачам дискретной оптимизации (в частности, к задачам теории расписаний или календарного планирования). Для классических постановок задач дискретной оптимизации разработано достаточное количество методов (симплекс-метод и др.).

Реальные задачи данного класса имеют NP-сложность и большую размерность. Данное обстоятельство не позволяет получить аналитического решения, поэтому приходится искать приближенные методы, позволяющие получить рациональное решение, или использовать методы искусственного интеллекта.

Интеллектуальные технологии позволяют реализовать моделирование реального объекта в компьютерной среде без жесткой математической модели (в частности, многоагентный подход). Этот вид моделирования позволяет эффективно исследовать реальный объект в компьютерной среде. Такое моделирование также представляет ценность наряду с теми решениями, которые позволяет получить математика. В данной работе делается акцент на использование интеллектуальных технологий при моделировании в компьютерной среде тех процессов, которые либо не получается решить математической постановкой, поскольку не все ограничения удается формализовать, либо математическое решение имеет ограниченное действие.

В соответствии с предложенными методологическими основами решение поставленной задачи предлагается с использованием многоагентного подхода. Онтологические модели, позволяющие представить слабо формализуемые семантические ограничения предметной области, реализованы в базе знаний и служат основой информационного обеспечения задачи календарного планирования производства.

Модель и структура многоагентной системы управления ресурсами для задачи календарного планирования производства

Для реализации в компьютерной среде задачи управления ресурсами сложных систем с учетом семантических ограничений предметной области разработана модель и структура многоагентной системы для управления ресурсами [22, 23]:

$MAS = <A_{рес}, A_{потр_рес}, Acts_{рес}, Acts_{потр_рес}, KB, Constraints, Tasks, F>$, где

$A_{рес} = \{a_{рес1},..., a_{рес N}\}$ – множество агентов-ресурсов;

$A_{потр_рес} = \{ a_{потр_рес1},..., a_{потр_рес M} \}$ – множество агентов-потребителей ресурсов;

$Acts_{рес} = \{acts_{рес1},..., acts_{рес N}\}$ – множество возможных действий агентов-ресурсов;

$Acts_{потр_рес} = \{ acts_{потр_рес1},..., acts_{потр_рес M} \}$ – множество возможных действий агентов-потребителей ресурсов;

$Constraints = \{constraints_1, ..., constraints_L\}$ – множество ограничений, включая семантические, предметной области;

$Tasks = \{tasks_1, ..., tasks_L\}$ – множество задач для каждого агента, которые должны быть решены;

$KB = \{ KB_1, ..., KB_L \}$ – множество баз знаний агентов;

$F = \{ F_1, \ldots, F_S, F_0\}$ – множество критериев оценки производительности для каждого агента, F_o – общий критерий оценки всей системы в целом.

Основные элементы предлагаемой многоагентной системы для управления ресурсами в сложных системах в рамках предлагаемой методологии представлены на рисунке 1.

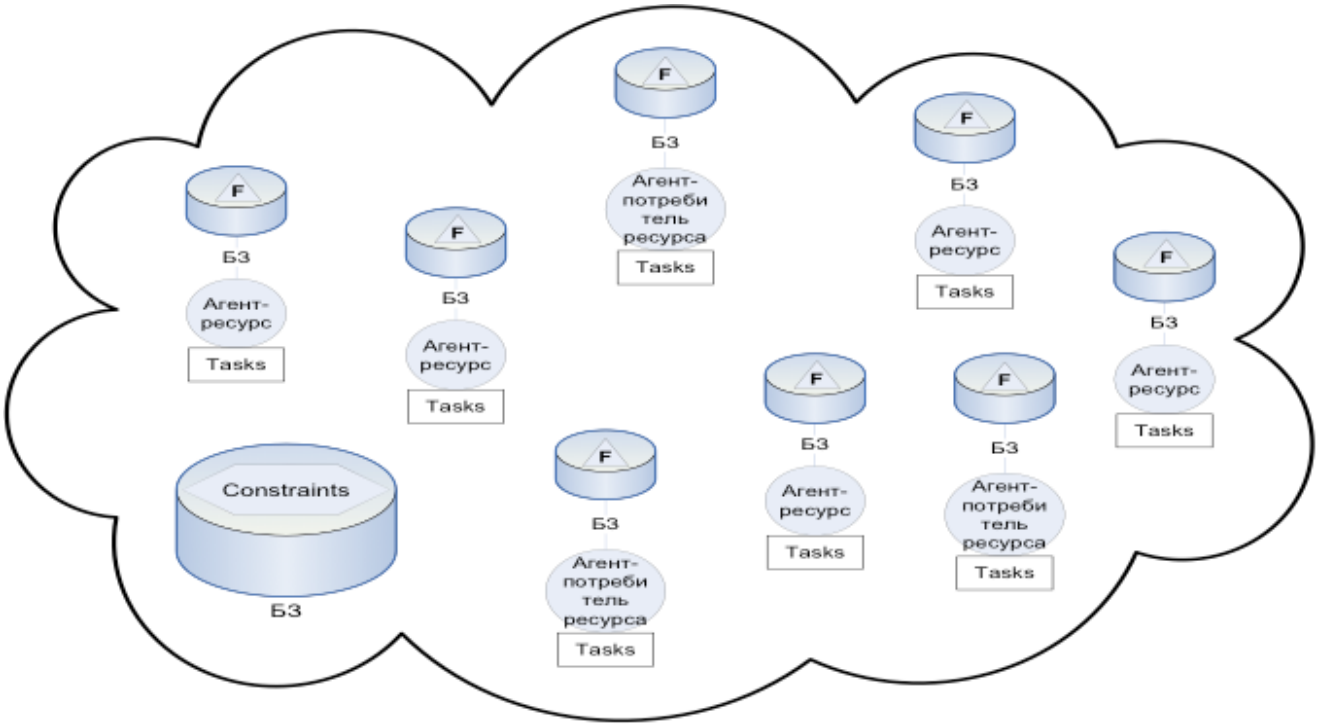


Рис. 1. Состав многоагентной системы для управления ресурсами сложных систем.

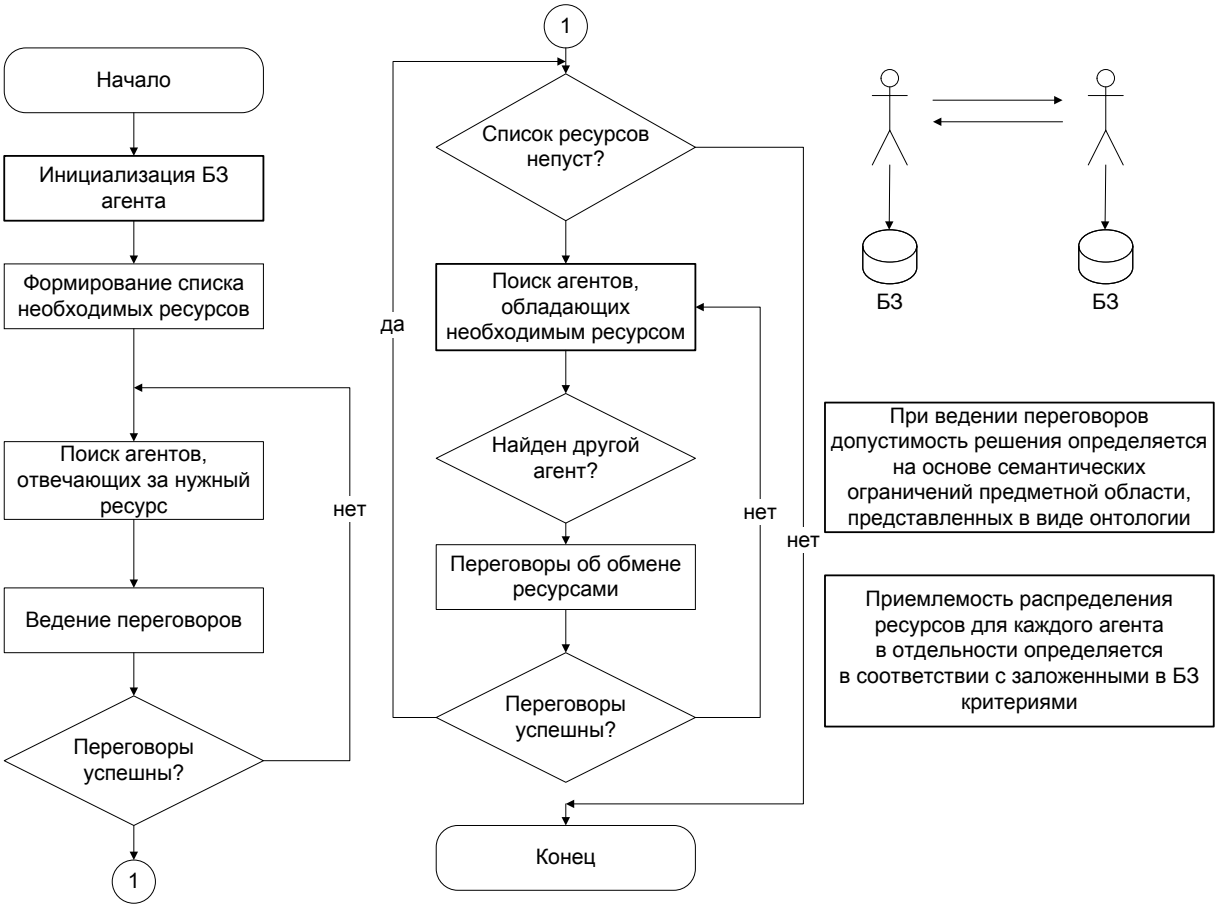


Рис. 2. Обобщенный алгоритм распределения ресурсов

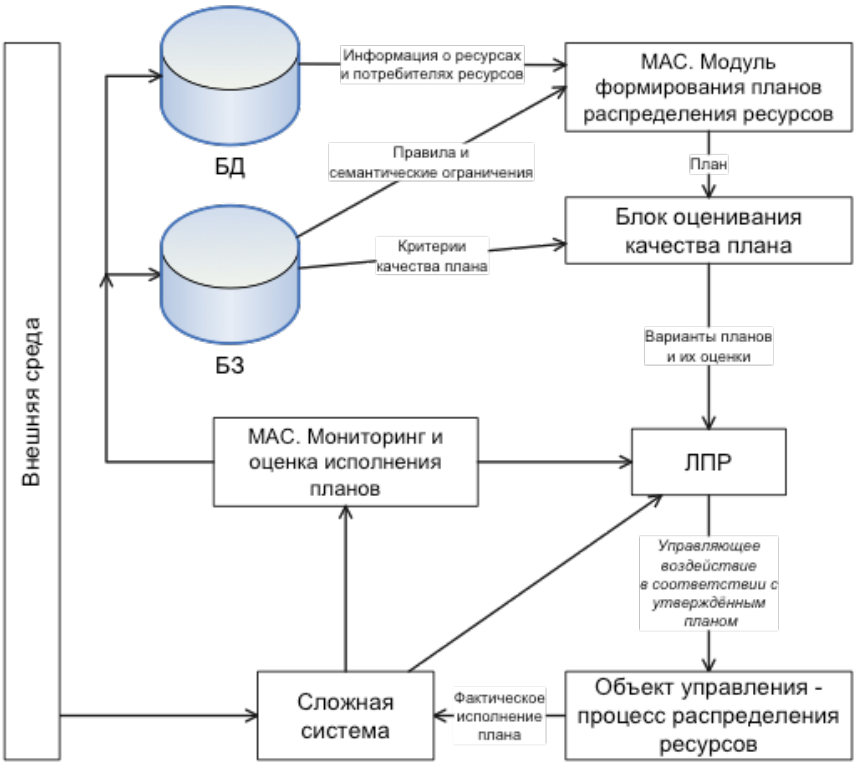


Рис. 3. Схема управления ресурсами в сложных системах

Общая схема решения задачи распределения ресурсов может быть представлена в виде обобщенного алгоритма, представленного на рисунке 2 [24].

Любые сложно формализуемые семантические ограничения предметной области представляются в виде онтологической базы знаний, доступной агентам. Кроме этого база знаний агента используется для описания критериев, отражающих приемлемость распределения ресурсов, и учитывает индивидуальные особенности моделируемых объектов реального мира.

При ведении переговоров агенты определяют допустимость и приемлемость для них решения на основе индивидуальной базы знаний. Таким образом, если меняются критерии и предпочтения у какого-то агента, достаточно внести изменения в его локальную базу знаний. При этом сама процедура ведения переговоров остается без изменений.

С учетом требований к системе поддержки принятия решений при управлении ресурсами в сложных системах, сформулированными в [25], была разработана схема управления ресурсами в сложных системах, представленная на рисунке 3.

В качестве объекта управления выступает процесс распределения ресурсов. Лицо, принимающее решение (ЛПР), оказывает управляющие воздействия на объект управления в соответствии с утвержденным планом распределения ресурсов. Информационная поддержка принятия решений заключается в выработке альтернативных вариантов планов распределения ресурсов с соответствующими оценками и предоставлении их лицу, принимающему решения. За построение этих планов отвечает модуль формирования планов распределения ресурсов. При построении планов распределения ресурсов используется информация о ресурсах и потребителях ресурсов, технологические параметры, правила и семантические ограничения предметной области. Оценивание планов производится с учетом критериев эффективности, которые хранятся в базе знаний СППР.

Блок Мониторинг и оценка исполнения планов позволяет осуществлять мониторинг фактического исполнения плана. Результаты этого мониторинга поступают в базу данных и базу знаний СППР, а также ЛПР.

В соответствии с разработанной методологией поддержки принятия решений при управлении ресурсами в сложных системах [26] для задачи календарного планирования производства в качестве ресурсов будут выступать рабочие, станки, оборудование, инструмент и т.д., а в качестве потребителей ресурсов – деталь, операция.

Глобальная цель потребителей ресурсов и ресурсов – минимизация времени изготовления всех деталей с учетом имеющихся ресурсов. Локальные цели ресурсов – максимизация дохода, повышение квалификации и т.д. Локальные цели потребителей ресурсов – изготовление в первую половину месяца.

Для оценки эффективности поддержки принятия решений при управлении ресурсами для задачи календарного планирования производства были проведены два типа экспериментов [27]. Целью экспери-

ментов являлось показать наличие положительного эффекта выбранных показателей при формировании месячного календарного плана изготовления деталей по сравнению с существующими планами, формируемыми в настоящее время вручную.

В ходе экспериментальных расчетов получены следующие результаты:

- снижение общего времени выполнения плана в среднем составило от 1.2 до 2 раз в зависимости от конкретного набора исходных данных по сравнению с текущим планированием;

- снижение непроизводственных простоев в среднем на 10–30 % по сравнению с текущим планированием.

Выводы

Рассмотрено применение алгоритмов интеллектуальной обработки данных для управления ресурсами в условиях неопределенности и динамичности внешней среды. Предложенные модели и алгоритмы апробированы для решения задачи календарного планирования производства. Результаты апробации позволяют судить о работоспособности и эффективности предлагаемых алгоритмов.

Благодарности

Исследования проводились при частичной поддержке грантами РФФИ 18-07-00193-а, 19-07-00895-а.

Список литературы

1. Szendrői E.: A robust hybrid method for the multimode resource-constrained project scheduling problem. Pollack Periodica 5(3), 175−184 (2010).

2. Fera F., Fruggiero F., Lambiase A., Martino G., Nenni M.E.: Production scheduling approaches for operations management, InTech (2013).

3. Xiao J., Osterweil LJ., Wang Q., Li M.: Dynamic scheduling in systems with complex resource allocation requirements. Department of Computer Science, University of Massachusetts, Amherst, MA 01003 (2009).

4. Artigues C., Huguet M.J., Lopez P.: Generalized disjunctive constraint propagation for solving the job shop problem with time lags. Engineering Applications of Artificial Intelligence 24(2), 220–231 (2011).

5. Monica D., Devi S., Subashini D., Devi R.: Scheduling and resource allocation for employees, software projects. International Journal of Advanced Computational Engineering and Networking 2(5), 26–29 (2014).

6. Kirn S., Herzog O., Lockemann P., Spaniol O.: Multiagent engineering, Theory and applications in enterprises. Springer (2006).

7. Kovács G.L., Gavalcová D.: Some experiments with a knowledge based simulator-scheduler. In: Stelson, K., Oba, F. (eds.) Japan-USA Symposium on Flexible Automation, vol. 2, pp. 1337−1344. Boston, USA (1996).

8. Fiorucci P., Gaetani F.P., Minciardi R., Sacile R., Trasforini E.: Real time optimal resource allocation in natural hazard management. In: iEMSs Conference, pp. 1−6. Osnabrück, Germany (2004).

9. Huang T., Yasuda K. Comprehensive review of literature survey articles on ERP, Business Process Management Journal, Vol. 22, No. 1, 2016. Pp. 2−32.

10. Haddara M., Zach O. ERP systems in SMEs: An extended literature review, International Journal of Information Science, Vol. 2, No. 6, 2012. Pp. 106−116.

11. Гайфуллин Б.Н., Обухов И.А. Автоматизированные системы управления предприятиями стандарта ERP/MRPII. М.: Богород. печатник, 2000. 103 с.

12. Симонова Л.А., Руднев М.П. Интегрированное информационное обеспечение процесса управления технологическими маршрутами в рамках ERP-системы. М.: Академия, 2005. 284 с.

13. О'Лири Д. ERP системы. Современное планирование и управление ресурсами предприятия. М.: Вершина, 2004. 272 с.

14. Фирсов М.В. Концепция создания ERP-систем. М.: ТЕИС, 2004. 93 с.

15. Фролов Е.Б., Загидуллин Р.Р. MES-системы. MES-системы, как они есть или эволюция систем планирования производства. Часть I, Часть II. URL: http://erpnews.ru/doc2592.html, http://erpnews.ru/doc2593.html. (дата обращения 03.08.2018).

16. Загидуллин Р.Р. Управление машиностроительным производством с помощью систем MES, APS, ERP. Старый Оскол: ТНТ, 2011. 372 с.

17. Характеристика и функции MES-системы ФОБОС. URL: http://www.fobos-mes.ru/fobos-system/MES-system-characteristic-and-functions.html (дата обращения 03.08.2018).

18. Гольфстрим. В фокусе – производство. Производственное планирование. URL: http://gulfstream-mrp.ru/functions/planning/ (дата обращения 03.08.2018).

19. MES-система для машиностроения PolyPlan. URL: http://www.fobos-mes.ru/sistema-polyplan/mes-sistema-dlya-mashinostroeniya-polyplan.html (дата обращения 03.08.2018).

20. Скобелев П.О. Интеллектуальные системы управления ресурсами в реальном времени: принципы разработки, опыт промышленных внедрений и перспективы развития // Приложение к журналу «Информационные технологии», 2013. № 1. С. 1–32.

21. Kovacs G., Yussupova N., Rizvanov D.: Resource management simulation using multi-agent approach and semantic constraints. Pollack Periodica 12(1), 45–58 (2017).

22. N. Yussupova, D. Rizvanov Decision-Making Support in Resource Management in Manufacturing Scheduling // Proceedings of 18th IFAC Conference on Technology, Culture and International Stability (TECIS 2018). Volume 51, Issue 30, Pp. 544–547.

23. Ризванов Д.А., Юсупова Н.И. Модели обработки данных при управлении ресурсами в сложных системах в условиях неопределенности и семантических ограничений предметной области // Информа-

ционные технологии интеллектуальной поддержки принятия решений (ITIDS'2018). Труды VI Всероссийской конференции (с приглашением зарубежных ученых). 2018. С. 30-35.

24. Ризванов Д.А. Алгоритмы управления ресурсами в сложных системах с применением многоагентных технологий // Вестник УГАТУ. 2013. Т. 17, № 5 (58). – С. 117–123.

25. Ризванов Д.А., Юсупова Н.И. Математическое и программное обеспечение для информационной поддержки принятия решений при управлении ресурсами сложных систем // Фундаментальные исследования. 2015. № 10, ч. 2. С. 294–298.

26. Ризванов Д.А. Методологические основы поддержки принятия решений при управлении ресурсами в сложных системах / Вестник Ижевского государственного технического университета. Ижевск: Изд-во ИжГТУ. 2018 г., том 21, № 4, с. 200–207.

27. Ризванов Д.А., Юсупова Н.И. Применение интеллектуальных технологий управления ресурсами при календарном планировании производства / Интеллектуальные системы в производстве. Ижевск: Изд-во ИжГТУ. 2018 г., том 16, № 4, с. 130–137.

УДК 622.276(075.8)

М.Б. Балдунов

МЕТОДИЧЕСКИЕ ОСНОВЫ КОЛИЧЕСТВЕННОГО АНАЛИЗА СИСТЕМ ЗАВОДНЕНИЯ НЕФТЯНЫХ МЕСТОРОЖДЕНИЙ

Аннотация

Заводнение – наиболее интенсивный и экономически эффективный способ воздействия на нефтяной пласт. Он позволяет значительно уменьшить число нефтяных скважин и резко повысить их дебиты. Однако непосредственное использование традиционных сеточных моделей не дают понимания процессов перемещения нефти вытесняющим агентом и не позволяют оценить их количественно и поэтому мало полезны для решения ряда актуальных проблем повышения эффективности разработки месторождений. Напротив, модели линий тока являются источниками уникальной информации о направленности и интенсивности фильтрационных потоков в нефтяном резервуаре. Эта информация может быть использована для количественного анализа эффективности систем разработки. Использование знаний об интенсивности дренирования позволяет выполнить прямой расчет коэффициента охвата вытеснением, оценить потенциал добывающих скважин, оценить эффективность и сбалансированность систем заводнения, выполнить анализ возможных источников обводнения продукции скважин и решить другие задачи. В отличие от традиционных сеточных методов, метод линий тока способен дать не только последовательность изменений полей насыщенности во времени, но и визуализировать структуру потоков флюидов и дать немедленный «снимок» всей структуры потока и взаимодействия скважин. Метод линий тока позволяют получить еще и уникальную количественную информацию о потоках и взаимодействиях скважин. Ее использование позволяет проводить количественный анализ эффективности сложных реальных систем разработки. Можно сказать, что метод линий тока это мощное дополнительное средство к традиционным методам используемых в управлении разработкой и добычей.

Ключевые слова: заводнение, метод линий тока, нагнетательные скважины, добывающие скважины, коэффициент влияния, коэффициент зависимости, дренируемый поровой объём, время полета.

Введение

Современный этап развития отечественной нефтедобывающей отрасли характеризуется весьма высокими уровнями добычи нефти, определяемыми крупными и крупнейшими месторождениями. Боль-шинство этих месторождений вступили в зрелую стадию разработки и постепенно снижают свой потенциал. Структура запасов нефти этих месторождений постоянно ухудшается, фонд высокообводненных и бездействующих скважин на месторождениях не уменьшается и даже растет.

В этих условиях особую актуальность приобретают вопросы эффективной разработки месторождений, совершенствования традиционных и применения новых технологий воздействия на залежи.

Заводнение – наиболее интенсивный и экономически эффективный способ воздействия на нефтяной пласт. Позволяет значительно уменьшить число нефтяных скважин и резко повысить их дебиты. Этот способ воздействия на пласт при разработке нефтяных месторождений, при котором поддержание и восстановление пластового давления и баланса энергии осуществляются закачкой воды. Обеспечивает высокие темпы добычи нефти и (в благоприятных условиях) сравнительно высокую степень извлечения нефти (до 70 %) [1].

Однако непосредственное использование тради-ционных сеточных моделей нефтяных резервуаров для анализа эффективности систем разработки затруднено. Они не дают понимания процессов перемещения нефти вытесняющим агентом и не позволяют оценить их количественно и поэтому мало полезны для решения ряда актуальных проблем повышения эффективности разработки месторождений.

Модели линий тока являются источниками уни-кальной информации о направленности и интенсивности фильтрационных потоков в нефтяном резервуаре. Эта информация может быть использована для количественного анализа эффективности систем разработки. Такой количественный анализ абсолютно необходим для регулирования и оптимизации систем разработки как с традиционными вертикальными, наклонно-ориентированными и горизонтальными скважинами, так и с многозабойными скважинами. Использование знаний об интенсивности дренирования позволяет выполнить прямой расчет коэффициента охвата вытеснением, оценить потенциал добывающих скважин, оценить эффективность и сбалансированность систем заводнения, выполнить анализ возможных источников обводнения продукции скважин и решить другие задачи [2].

В отличие от традиционных сеточных методов, в методе линий тока флюид движется вдоль естественной сетки линий тока, а не между блоками обычной сетки как в стандартном методе [3]. Метод линий тока способен дать не только последовательность изменений полей насыщенности во времени, но и визуализировать структуру потоков флюидов и дать

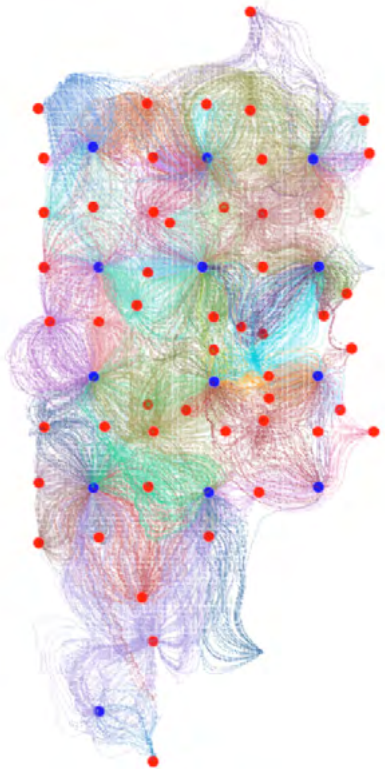


Рис. 1. Линии тока между добывающими и нагнетательными скважинами [4]

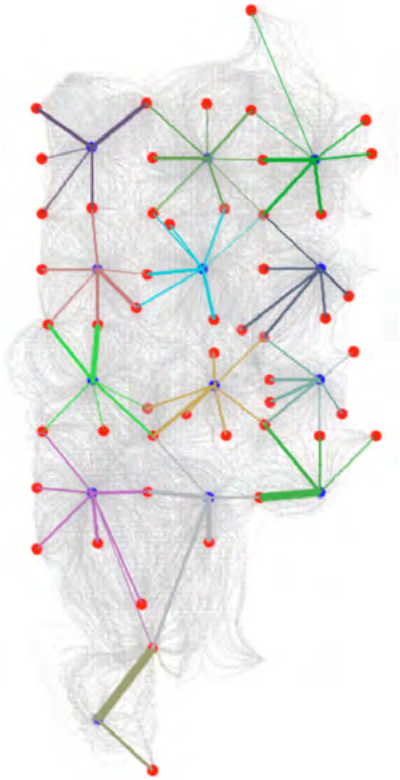


Рис. 2. Карта потоков [4]

немедленный «снимок» всей структуры потока и взаимодействия скважин.

Однако метод линий тока позволяют получить еще и уникальную количественную информацию о потоках и взаимодействиях скважин. Ее использование позволяет проводить количественный анализ эффективности сложных реальных систем разработки.

Возможность получения количественной информации о взаимосвязях нагнетательных и нефтяных скважин по потокам флюидов обусловлена линиями тока и их пучками, направленными от источников к стокам процесса фильтрации. В качестве таких источников и стоков будем рассматривать скважины нагнетательные *i* (Injectors) и добывающие *p* (Producers), а также контур питания *e* (Edge) и поровое пространство резервуара *f* (Field).

Flux Pattern Map (FPmap) – это технология визуализации Streamsim для отображения и количественного определения связности между парами нагнетательных и добывающих скважин [4].

Красные точки обозначают открытые добывающие скважины, а синие точки – открытые нагнетательные, в определенный момент времени. Линии тока, рассчитанные с помощью симулятора линий тока – 3DSL с использованием базовой трехмерной сетки с пространственными геологическими свойствами и (историческими) скважинами, соединяют нагнетательные и добывающие (рис. 1).

Чтобы лучше представить, какие скважины подключены, нагнетающие и добывающие соединены линией, в которой цвет, толщина и метка используются для характеристики отдельных соединений. Результатом является запатентованная карта потоков FPmap. Здесь все соединения добывающие и нагнетательные, окрашены в один и тот же цвет, что позволяет легко идентифицировать узоры (рис. 2).

Величина притока флюида *j* к стоку «*b*» по пучку линий тока, приходящему от источника «*a*», может быть рассчитана по формуле

$$Q_{a,b}^{(j)} = \sum_{k=1}^{n_{ab}} q_{k,j}^{a,b} \qquad (1)$$

Здесь n_{ab} – число линий тока от источника *a* к стоку *b*, $q_{k,j}^{b,a}$ – приток фазы *j* к стоку *b* от источника *a* по *k*–й линии тока, м^3/сут; *j*=1 – вода, *j*=2 – нефть, *j*=3 – газ.

Для описания структуры потоков флюидов используется **коэффициент влияния (WAF – Well Allocation Factor)** нагнетательной скважины *i* на добывающую скважину p по флюиду *j*. WAF изменяется во времени, так как линии тока обновляются из-за изменения дебита скважины, закрытия/открытия скважин и/или изменения распределения внутренней жидкости.

$$WAF^{(j)}{}_{pi} = \frac{Q_{pi}^{(j)}}{q_i} \qquad (2)$$

где $Q_{pi}^{(j)}$ – приток флюидов, м^3/сут q_i – приемистость нагнетательной скважины *i*, м^3/сут.

Коэффициентом влияния нагнетательной скважины *i* на добывающую скважину *p* по жидкости будем

называть сумму коэффициентов влияния по каждому из флюидов:

$$WAF_{pi} = \sum_{j=1}^{3} WAF^{(j)}_{pi}, \quad (3)$$

где j=1 – вода, j=2 – нефть, j=3 – газ.

Коэффициентом зависимости добывающей скважины p от нагнетательной скважины i по флюиду j будем называть отношение притока этого флюида в скважину p от скважины i к дебиту скважины p по жидкости:

$$WAF^{(j)}_{ip} = \frac{Q^{(j)}_{ip}}{q_p} \quad (4)$$

Линии тока и их пучки позволяют разделить нефтяной резервуар на зоны динамического дренирования каждой из скважин. В каждой такой зоне могут быть рассчитаны как общий **дренируемый поровый объем (WPV – Well Pore Volume)**, так и текущие объемы воды, остаточные запасы нефти и газа:

$$WPV^{(j)}_{p} = \sum_{u=1}^{N_p} (v_u \cdot s^{(j)}_u) \text{ и} \quad (5)$$

$$WPV_p = \sum_{j=1}^{3} WPV^{(j)}_{p}, \quad (6)$$

где WPV_p – объем порового пространства, дренируемый скважиной p, $WPV^{(j)}_p$ – объем воды (j=1), нефти (j=2), газа (j=3) в зоне дренирования, p – номер скважины.

Время полета (TOF – Time of Flight) – время переноса частицы флюида по линии тока от ее начала в текущую точку – третий важнейший результат метода линий тока. Определение этого параметра основано на нахождении поля скоростей фильтрации. По известному TOF нетрудно вычислить «время дренирования» (DRT) – время необходимое для преодоления частицей расстояния от текущей точки до конца линии тока:

$$DRT = TOF_{end} - TOF \quad (7)$$

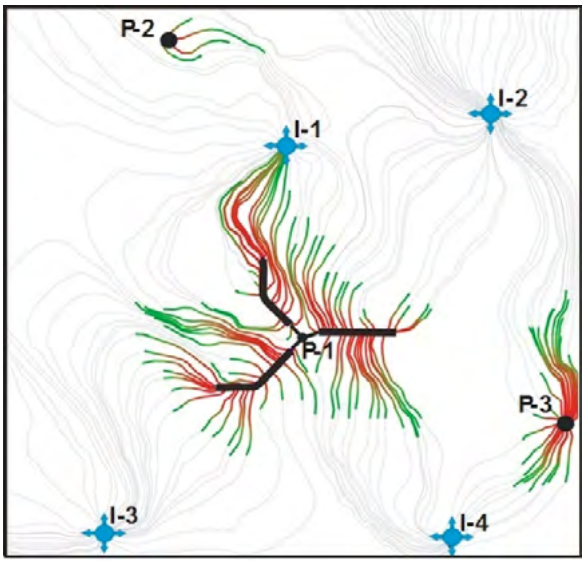


Рис. 3. Зоны дренирования [2]

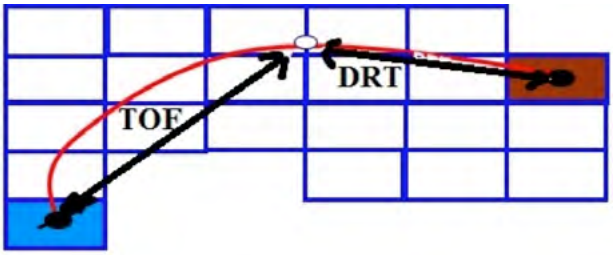


Рис. 4. Расчет скоростей и времен дренирования [4]

Здесь TOF_{end} – время полета вдоль всей линии тока. Параметр DRT позволяет оценить время прорыва воды в добывающую скважину, рассчитать фронт и кривую обводнения.

Метод линий тока – это, вероятно, единственный, метод расчета величин потоков между скважинами и зон дренирования, времен полета и дренирования. Эти результаты могут стать основой для анализа эффективности систем разработки. Целью такого количественного анализа является оценка эффективности каждой из добывающих и нагнетательных скважин и оценка воздействия системы скважин на залежь в целом.

Эффективность использования воды, закачиваемой в i-тую нагнетательную скважину для компенсации отбора пластовых флюидов n добывающими скважинами, можно оценить с помощью WAF – коэффициентов:

$$Ioffset_i = \sum_{p=1}^{n} WAF_{pi}, \quad (8)$$

где n – число добывающих скважин, гидродинамически связанных с i-той нагнетательной скважиной.

Параметр *Ioffset* показывает долю общего объема закачиваемого в скважину вытесняющего агента, притекающую по линиям тока к n добывающим скважинам от i-той нагнетательной скважины, от общего объема закачиваемого вытесняющего агента в эту скважину. Окончательное суждение о величине компенсации отбора закачкой может быть сделано после решения уравнения материального баланса для зоны дренирования.

Классификация нагнетательных скважин по величине параметра *Ioffset* позволяет:

1. Выбирать нагнетательные скважины для проведения дополнительных исследований и ремонтных работ.
2. Обосновывать нецелесообразность продолжения закачки в отдельные нагнетательные скважины.
3. Обосновывать необходимость приобщения и изоляции пластов, применения MIE-технологий.

Выбирать местоположения проектных нагнетательных скважин и направления разбуривания дополнительных стволов скважин, или обосновывать перевод отдельных нефтяных скважин под закачку. В качестве примера на рисунке 5 приведена классификация нагнетательных скважин по эффективности участия в процессах компенсации отбора закачкой.

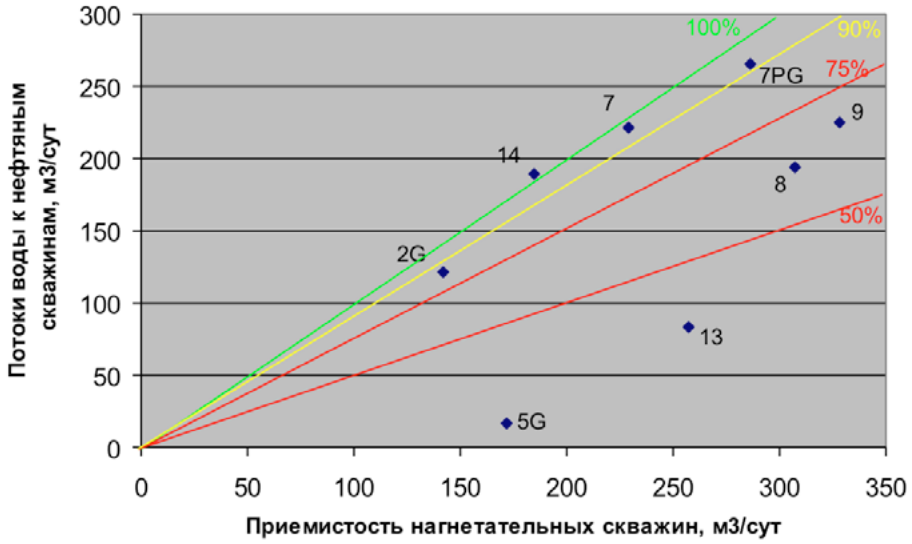


Рис. 5. Классификация нагнетательных скважин по эффективности компенсации отбора [2]

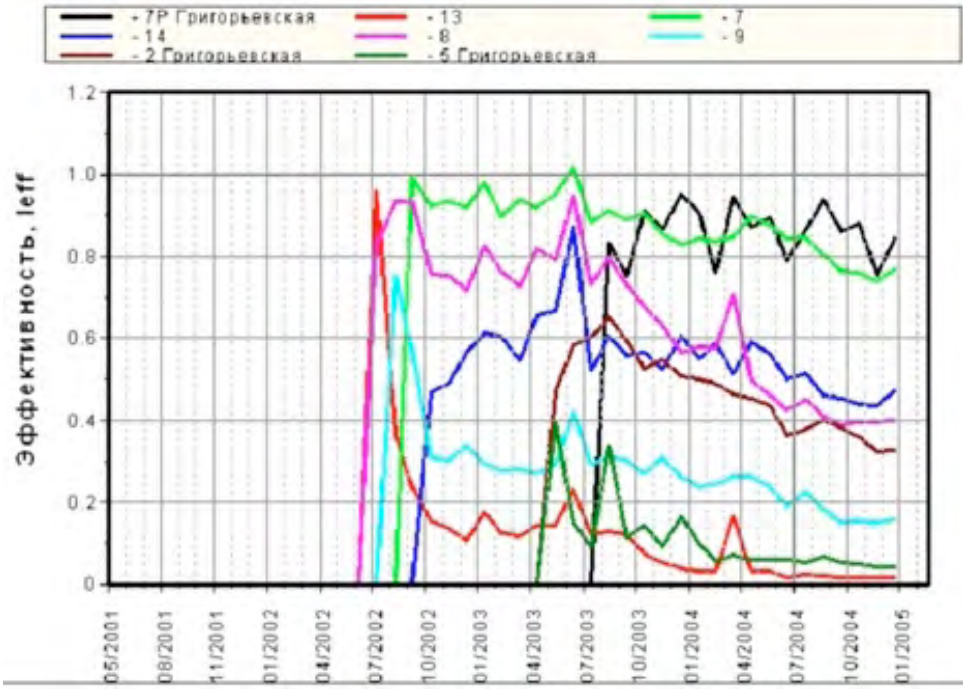


Рис. 6. Динамика эффективности нагнетательных скважин [2]

Эффективность влияния *i*-той нагнетательной скважины на добычу нефти всеми гидродинамически связанными с ней нефтяными скважинами можно рассчитать с использованием *WAF* – коэффициентов:

$$Ieff_i = \sum_{p=1}^{n} WAF^{(2)}{}_{pi} \qquad (9)$$

Параметр *Ieff* показывает: сколько нефти добыто *n* добывающими скважинами в результате закачки единицы объема воды (газа) в *i*-тую нагнетательную скважину.

Классификация нагнетательных скважин по величине параметра *Ieff* является основой для решения следующих задач:

1. Подбор технологических режимов работы нагнетательных скважин с целью:

• увеличения дебита добывающих скважин по нефти: максимизации добычи нефти, вместо максимизации объемов нагнетания воды,

• перераспределения объемов закачек в пользу высокоэффективных нагнетательных скважин;

2. Обоснование необходимости приобщения или изоляции пластов, использования MIE-технологий;

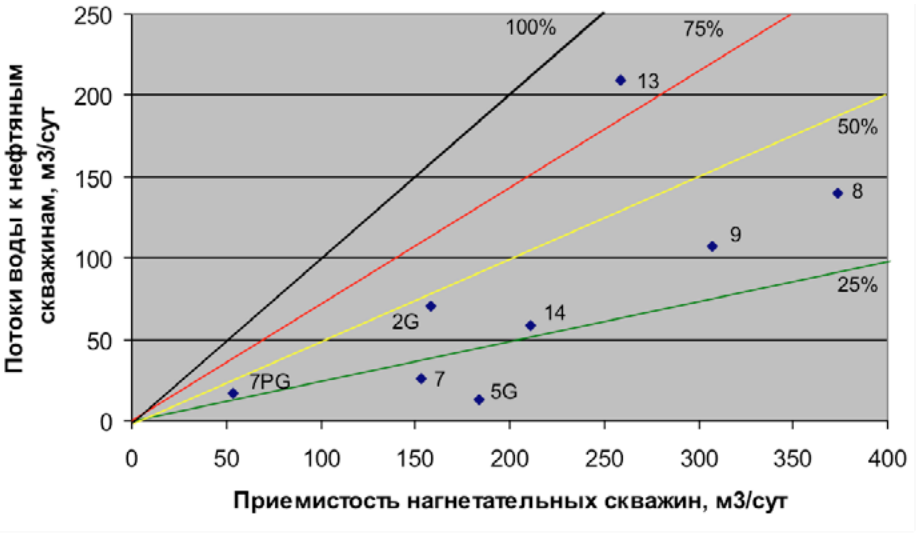


Рис. 7. Классификация нагнетательных скважин по влиянию на обводнение продукции добывающих скважин [2]

3. Выбор малоэффективных нагнетательных скважин – кандидатов для остановки при ограниченных мощностях СППД по закачке воды.

Влияние *i*-той нагнетательной скважины на обводнение продукции конкретной *p*-той добывающей скважины может быть оценено $WAF^{(j)}_{pi}$ – коэффициентом (2) для *j*=1.

Суммарное влияние *i*-той нагнетательной скважины на обводнение продукции *n* гидродинамически связанных добывающих скважин может быть рассчитано по формуле:

$$Iwcut_i = \sum_{p=1}^{n} WAF^{(1)}{}_{pi} \qquad (10)$$

Параметр *Iwcut* показывает: сколько воды отобрано *n* скважинами в результате закачки единицы объема воды (газа) в *i*-тую нагнетательную скважину.

Параметры $WAF^{(1)}_{pi}$ и $Iwcut_i$ дают количественные оценки для организации управления нагнетанием с целью уменьшения обводнения продукции добывающих скважин.

На основе классификации нагнетательных скважин по величине параметра *Iwcut* и пар скважин «нагнетательная – нефтяная» по параметру $WAF^{(1)}_{pi}$ могут быть решены следующие задачи:

1. Выявление нагнетательных скважин – источников нерациональных циклов воды «от нагнетательных к добывающим скважинам».

2. Выбор технологических режимов работы нагнетательных скважин для уменьшения обводнения продукции добывающих скважин.

3. Обоснование выбора нагнетательных скважин для дополнительных исследований и ремонтных работ.

Опыт разработки отечественных и зарубежных месторождений показывает, что заводнение оказывается довольно эффективным методом воздействия на низкопроницаемые коллектора при строгом соблюдении необходимых требований к технологии его осуществления.

В числе основных причин, вызывающих снижение эффективности заводнения низкопроницаемых пластов оказываются:

• ухудшение фильтрационных свойств породы за счет:

• набухания глинистых составляющих породы при контакте с закачиваемой водой,

• засорения коллектора мелкодисперсными механическими примесями, находящимися в закачиваемой воде,

• выпадением в пористой среде коллектора осадков солей при химическом взаимодействии нагнетаемой и пластовой воды,

• уменьшение охвата пласта заводнением вследствие образования вокруг нагнетательных скважин трещин-разрыва и распространения их в глубь пласт (для прерывистых пластов возможно также некоторое увеличения охвата пласта по разрезу),

• значительная чувствительность к характеру смачиваемости пород нагнетаемым агентом значительное снижение проницаемости коллектора за счет выпадения парафинов.

Для устранения влияния на процесс заводнения указанных факторов используются соответствующие технологические решения: оптимальные сетки скважин и технологические режимы эксплуатации скважин, нагнетание в пласты воды необходимого типа и состава, соответствующая ее механическая, химическая и биологическая очистка, а также добавка в воду специальных компонентов.

Для Приобского месторождения заводнение следует рассматривать в качестве основного метода воздействия [5].

В процессе разработки Самотлорского месторождения широкое применение нашел гидродинамический метод – заводнение нефтяных залежей, которое обеспечило сравнительно высокую нефтеотдачу пластов. Однако показатели по добыче постепенно снижались, особенно в сильно неоднородных объектах, что было связано с прорывом нагнетаемой воды к добывающим скважинам по высокопроницаемым пропласткам, оставляя неохваченными низкопроницаемые зоны. Для повышения эффективности добычи нефти было предложено циклическое заводнение, сущность которого заключается в создании в пластовых условиях периодических волн увеличения и снижения давления. Под воздействием знакопеременных перепадов давления происходит перераспределение потоков жидкости в неравномерно насыщенном пласте. Нестационарное заводнение, как метод регулирования разработки, начало применяться на Самотлорском месторождении с 1975 года. Нестационарное заводнение осуществляется путем попеременной работы двух, трех и более групп нагнетательных скважин, группируемых в ряду по принципу «подряд». Продолжительность полуциклов изменяется в диапазоне 30–120 суток [6].

Выводы

Модели линий тока в отличие от традиционных сеточных методов позволяют получить принципиально новую информацию о взаимодействиях каждой пары добывающая-нагнетательная скважина, дифференцировать дренируемые поровые объемы и рассчитывать компенсацию отборов закачкой по скважинам.

Результаты моделирования методом линий тока – основа для быстрого и эффективного количественного анализа реальных систем заводнения любой сложности и классификации скважин по эффективности. Такой анализ – основа оптимизации систем заводнения.

Количественный анализ необходим для определения эффективности систем воздействия на залежи и формирования гидродинамических регулирующих воздействий как при проектировании систем разработки месторождения, так и при оперативном геолого-промысловом анализе.

Автор выражает глубокую признательность научному консультанту д.т.н. Костюченко С.В. за помощь в подготовке данного доклада

Список литературы

1. Горная энциклопедия. Date Views 10.01.2019 http://www.mining-enc.ru/z/zavodnenie/

2. Костюченко С.В., Зимин С.В. Количественный анализ эффективности систем заводнения на основе моделей линий тока. «Нефтяное хозяйство», 1, 2005. – С. 56–60.

3. Сидельников К.А., Васильев В.В. Анализ современных способов увеличения эффективности моделирования нефтяных месторождений // Надежность и качество. Труды международного симпозиума / Под ред. Н.К. Юркова. – Пенза: Изд-во Пенз. гос. ун-та, 2005. – С. 227–230.

4. Streamsim Connect. Date Views 10.01.2019 https://www.streamsim.com/technology/reservoir-surveillance-streamlines?field_product1_value=2

5. Allbest
Date Views 28.01.2019 https://allbest.ru/

6. Российская академия естествознания
Date Views 28.01.2019 https://www.rae.ru/

УДК 621.391

Р.Ч. Гаянов, М.М. Гаянова, А.М. Вульфин

ИССЛЕДОВАНИЕ И РАЗРАБОТКА ЭНЕРГОЭФФЕКТИВНОГО МЕТОДА ВЗАИМОДЕЙСТВИЯ В ГЕТЕРОГЕННОЙ СЕТИ ИНТЕРНЕТА ВЕЩЕЙ С ИСПОЛЬЗОВАНИЕМ СПУТНИКОВОЙ СВЯЗИ

Аннотация

Областью исследования является передача данных в гетерогенной среде Интернета вещей с использованием спутниковой связи. Технологии Интернета вещей широко распространены в различных отраслях, что приводит к возникновению задачи построения гетерогенных сетей в условиях недоступности традиционных каналов связи.

На данный момент количество физических объектов Интернета вещей растет экспоненциально и по прогнозам компании Cisco достигнет 50 млрд. устройств до 2020 года [1]. При этом стандартом для связи физических объектов Интернета вещей являются три основных технологии передачи данных: LoRa, LTE и спутниковые каналы связи.

Для использования в труднодоступных районах в таких областях, как: добыча полезных ископаемых, использование в сельском хозяйстве, в грузоперевозках, рассматривается гетерогенная сеть, состоящая из сети LoRaWAN и спутниковых каналов связи.

Разработка протоколов взаимодействия в гетерогенной сети является одним из основных методов повышения эффективности передачи данных. Таким образом, исследование и разработка методов повышения эффективности взаимодействия в гетерогенной сети Интернета вещей является актуальным и позволяет расширить сферу использования Интернета вещей.

Ключевые слова: гетерогенная сеть, Интернет вещей, спутниковый канал.

Введение

В результате того, что технологии Интернета вещей широко распространились в различных отраслях производства, появилась задача построения гетерогенных сетей в труднодоступных районах с отсутствием традиционных каналов связи. Основными требованиями, предъявляемой разрабатываемой системе являются требования по обеспечению автономности устройств и энергоэффективности передачи данных. Целью работы является повышения эффективности взаимодействия цифровых объектов в гетерогенной сети промышленного Интернета вещей.

Для связи физических объектов Интернета вещей используется три технологии – LTE, LoRa и спутниковые каналы связи [2]. Комбинируя эти каналы связи для передачи данных, можно расширить сферу применения технологий Интернета вещей.

Для дальних и труднодоступных районов, в областях, где ведется добыча полезных ископаемых, в сельском хозяйстве и в грузоперевозках [3] в качестве беспроводной транспортной сети рассматривается гетерогенная сеть, состоящая из сети LoRaWAN и низкоскоростных спутниковых каналов.

Разработка протоколов взаимодействия устройств в гетерогенной сети Интернета вещей является одним из методов повышения эффективности передачи данных.

Обзор протоколов передачи данных

Предлагаются следующие критерии при выборе протокола передачи данных:

- Максимальная энергоэффективность;
- Возможность использования сетевого оборудования в труднодоступных районах;
- Максимальная зона покрытия одной базовой станции;
- Возможность использования нелицензируемого радиочастотного диапазона.
- LPWAN (Low Power Wide Area Networks) является базовой технологией передачи данных. Основными ее отличиями являются:
- Дальность передачи данных (до 10 км) [4];
- Более высокая энергоэффективность по сравнению с другими протоколами передачи данных. Расчётное время эксплуатации без замены источника питания в зависимости от типа сенсора – 2–10 лет;
- Меньшая пропускная способность (сенсоры Интернета вещей предают минимальный объем информации);
- Увеличенная задержка в процессе передачи данных;
- Поддержка большого количества устройств на одну базовую станцию;
- Низкая мобильность подключаемых устройств, так как в основном устройства являются стационарными датчиками;
- Относительно высокая надежность передачи данных.

Рассмотрим несколько технологий.

Технология СТРИЖ является примером технологии передачи данных на базе LPWAN, разработанной в России. Протокол использует нелицензируемый диапазон 868 МГц, но, в отличие от большинства существующих LPWAN технологий, является узкополосной – полоса для передачи данных составляет 100 Гц, что позволяет в диапазоне 868 МГц использовать 1250 каналов [5]. Скорость передачи данных состав-

ляет 100 бит/с. Использование узкополосного сигнала позволяет увеличить радиус передачи и плотность устройств за счет снижения скорости передачи. Но, несмотря на соответствие техническим требованиям, эта технология не подходит для использования в сети удаленного Интернета вещей, так как является полностью закрытой.

Технология Sigfox была разработана во Франции. Она похожа на российскую технологию СТРИЖ и используется нелицензированный частотный диапазон, зависящий от страны (в Европе используется частота 868 МГц. Как и СТРИЖ, Sigfox использует узкополосный сигнал с шириной канала в 100 Гц [6]. Однако, в отличие от технологии СТРИЖ, Sigfox поддерживает режим прямой связи между узлами – режим P2P. Sigfox поддерживает скорость передачи 600 бит/с [6]. Если сеть СТРИЖ строится только на территории России, то сеть Sigfox в настоящий момент развернута в 60 странах мира. Однако технология Sigfox подразумевает построение транспортной сети, а также является закрытым стандартом, что приводит к невозможности использования этой технологии.

LoRaWAN является одной из спецификаций LPWAN, которая для передачи данных использует метод модуляции LoRa. Метод модуляции LoRa в настоящий момент наиболее распространен и использует нелицензированный частотный диапазон до 1 ГГц (зависит от страны использования). На территории России, например, стандартом для сетей LoRa выделен частотный диапазон 864-869 МГц [7].

Технология использует линейную частотную модуляцию, которая обеспечивает высокую помехоустойчивость и дальность передачи. Метод модуляции LoRa использует технологию адаптивного расширения спектра, которая позволяет динамически изменять полосу пропускания в зависимости от диапазона используемых частот.

Также спецификация LoRaWAN является условно открытой технологией и не требует построения базовой инфраструктуры.

Другая технология, которая используется при передаче данных Интернета вещей, является технология NB-IoT [8]. Разработкой технологии занимался консорциум 3GPP и встроена в LTE стандарт. Она сильно упрощена по сравнению с LTE для того, чтобы для того, чтобы ее можно было использовать в энергоэффективных автономных устройствах. По сравнению с LTE, в NB-IoT отсутствует агрегация каналов для повышения производительности и роуминг между базовыми станциями, что сужает область применения технологии. К тому же она требует наличия сотовой связи, а значит, не может быть использована в труднодоступных районах.

Сравнение характеристик приведено в таблице 1.

Таблица 1

Сравнение технологий LPWAN

	СТРИЖ	**Sigfox**	**LoRa**	**NB-IoT**
Частотный диапазон	868 МГц	868 МГц (Европа) 902 МГц (США)	863-870 МГц (Европа) 902-928 МГц (США) 779-787 МГц (КНР)	450 МГц-3,5 ГГц
Ширина канала	50/100/1000 Гц	100 Гц	125/250/500 КГц	180/200 КГц
Радиус действия	До 50 км	До 50 км	1,5 км в городских условиях, 10 км в сельской местности	До 20 км
Скорость передачи	50/100/1000 бит/с	100-600 бит/с	0,5–5 Кбит/с	До 200 Кбит/с
Тип модуляции	Сверхузкополосный метод с дифференциальной двоичной фазовой манипуляцией DBPSK	DBPSK/GFSK	LORA, GFSK, MSK	QPSK
Режим доступа	FDMA и TDMA	TDMA	ALOHA	CSMA/CD
Шифрование	ГОСТ Р 34.12-2005 128 бит	AES	AES-128	AES-128
Прочие особенности	Используется узкополосный сигнал, что позволяет повысить дальность передачи и плотность устройств за счет снижения скорости передачи. Закрытая технология	Поддерживает режим прямой связи между узлами. Используется узкополосный сигнал. Закрытая технология	Используется линейная частотная модуляция, обеспечивающая высокую дальность передачи и помехоустойчивость. Условно открытая технология	Использует сотовую связь для передачи данных

На основании проведенного анализа, наиболее подходящей технологией связи в гетерогенной сети Интернета вещей является LoRaWAN.

Обзор спутниковых каналов связи

В труднодоступных регионах связь реализуется в основном с помощью спутниковой связи.

Существует следующая классификация систем подвижной спутниковой связи [9]:

- Глобальные геостационарные;
- Региональные геостационарные;
- Глобальные негеостационарные;
- Региональные негеостационарные.

Системы спутниковой связи (ССС) на геостационарных и негеостационарных орбитах имеют следующие отличия.

Таблица 2

Сравнение ССС на геостационарных и негеостационарных орбитах

ССС на геостационарных орбитах	ССС на негеостационарных орбитах
Невозможность обслуживания приполярного региона	Высокая пропускная способность канала
Большое время задержки	Полное покрытие земного шара
Постоянная подписка на услугу	Короткие сеансы связи
–	Прием на ненаправленную антенну

Рассмотрим спутниковые системы на геостационарной орбите (Inmarsat) и на негеостационарной орбите (Iridium). Эти спутниковые системы предоставляют услуги для Интернета вещей.

Спутниковая система Inmarsat

Спутники Inmarsat находятся на геостационарной орбите и обеспечивают покрытие до 70° северной и южной широты [10].

Эта спутниковая система предлагает Интернет-соединение на скорости 384 Кбит/сек. Ее ключевой особенностью является наличие контроллера, которое позволяет реализовать решение в короткие сроки. Существуют готовые M2M решения, выбор которых зависит от скорости, объема и сложности принимаемой и передаваемой информации. Среди таких решений IsatM2M [11], IsatData Pro [12], BGAN M2M [13].

Спутниковая система Iridium

Спутниковая система Iridium предлагает Интернет-соединение на скорости 128 Кбит/сек. Компания предлагает оборудование M2M на основе спутниковых модемов SBD типа 9602 [14], 9603N [15] и 9523 [16], позволяющих передавать короткие сообщения. Ключевыми особенностями этих модемов является возможность обмена текстовыми сообщениями и низкое энергопотребление.

Таблица 3

Сравнение ССС Iridium и Inmarsat

	Iridium	Inmarsat
Скорость передачи	128 Кбит/с	384 Кбит
Покрытие Арктики	Полное	Только часть
Наличие контроллера	Нет	Есть

Необходимость предоставления услуг связи в труднодоступных районах приводит к тому, что требуется зона покрытия в районах крайнего севера и Арктики. Единственной ССС, удовлетворяющая этому требованию, является спутниковая система Iridium.

Спутниковый SBD модем Iridium подключается по RS-232 интерфейсу к передающему устройству и не может быть подключен к каждому сенсору сети Интернета вещей. Таким образом для организации подключения сенсоров Интернета вещей по спутниковым каналам связи требуется шлюз между LPWAN сетью и спутниковым каналом связи.

Таким образом, в дальнейшем будет проводиться исследования гетерогенной сети Интернета вещей, состоящей из сети LoRaWAN и низкоскоростных спутниковых каналов связи Iridium.

Обзор форматов представления данных

При передаче информации через сеть Интернета вещей и спутниковую сеть, большое значение представляет метод представления информации, так как именно метод представления информации влияет на соотношение полезных данных и служебных в потоке информации.

Между сенсором и шлюзом LoRa информация передается в бинарном виде. В таком случае размер полезных данных равен общему размеру передаваемого пакета.

Сенсор может быть использован для передачи данных с более чем одного датчика. Тогда помимо передачи значения датчика требуется передать информацию об идентификаторе датчика и характере измеряемой величины. При дальнейшей передаче информации от шлюза LoRa в общую систему сбора и анализа данных, к данным добавляется идентификатор сенсора и параметры сети Интернета вещей.

Вследствие всего этого предлагается использовать формат ключ-значение при передаче информации в общую систему сбора и анализа данных

Рассмотрим некоторые из форматов представления информации.

Формат JSON [17] является подвидом формата ASN.1 и в настоящий момент является стандартом для передачи информации в системах Интернета вещей. В частности, это наиболее часто используемый формат в сообщениях, использующих MQTT: специально упрощенный сетевой протокол для об-

мена данных в сети Интернета вещей [18]. Текстовое представление информации в формате JSON дает возможность анализировать получаемые сообщения без использования декодирования, а передача информации в формате «ключ-значение» без определения значений передаваемых полей дает возможность гибкого использования формата. Наличие библиотек для обработки формата JSON в таких языках программирования, как C++, Java, Python, PHP, значительно облегчает использование этого формата при разработке приложений. Однако у этого формата есть и недостатки. Текстовый формат приводит к большому размеру сообщений, а отсутствие предварительного определения значений передаваемых полей приводит к тому, что в каждом сообщении все определения полей должны передаваться заново, что в свою очередь приводит к увеличению числа сообщений.

Формат BSON [19] является бинарной разновидностью JSON, который позволяет передавать бинарные данные в сообщениях. В зависимости от кодируемой информации формат может быть как более, так и менее эффективным по сравнению с JSON, но в целом BSON соответствует ему по размеру сообщений.

Форматы CBOR [20] и MsgPack [21] используют бинарное представление данных и уменьшают объем передаваемых сообщений за счет изменения формата передаваемых значений. Эти форматы не требуют использования текстовых разделителей полей, выполняют преобразования текстовых представлений целых чисел, чисел с плавающей запятой и дат в бинарное представление, что позволяет уменьшить объем передаваемых сообщений. CBOR и MsgPack не требуют предварительного определения передаваемых полей данных и сопоставимы по соотношению полезных данных к общему объему передаваемых данных.

Формат JSONC [22] использует компрессию данных при помощи алгоритма zlib без анализа передаваемых данных и степень сжатия зависит от характера и объема передаваемых данных.

При передаче информации сенсоров Интернета вещей, структуру данных можно определить заранее и не передавать информацию о полях передаваемых данных в каждом сообщении. Это позволяет уменьшить объем информации в сообщении.

Для передачи структурируемой информации используется формат Protobuf, разработанный компанией Google [23]. Protobuf является бинарным форматом, в котором уменьшение размера данных достигается за счет оптимизированного представления типов данных в бинарном виде и использования предопределения типов передаваемых значений. Таким образом, в парах ключ-значение передается только бинарный идентификатор ключа, а не его имя, что сильно снижает объем передаваемых данных.

С учетом возможности предварительного определения структуры передаваемых данных, форматом с минимальным объемом передаваемой информации является Protobuf.

Таким образом, для разработки энергоэффективного протокола взаимодействия в гетерогенной сети Интернета вещей с использованием спутниковой связи необходимо разработать метод представления информации с использованием формата Protobuf.

Выводы

Проведен анализ протоколов передачи данных, спутниковых каналов связи и методов представления информации, которые могут быть использованы для сбора информации с удаленных сенсоров в условиях отсутствия доступных каналов связи. Показаны особенности использования спутниковых каналов связи с целью увеличения скорости передачи данных из труднодоступных районов с помощью сети Интернета вещей.

Дальнейшими шагами проведения исследования являются:

• Разработка структурной модели гетерогенной сети Интернета вещей с использованием спутниковой связи для передачи информации;

• Разработка нового метода представления информации в сети Интернета вещей.

Список литературы

1. Если взглянуть на Интернет вещей шире / Cisco. Date Views 28.01.19 https://www.cisco.com/c/ru_ua/about/press/2015/02-021615c.html

2. Эволюция интернета вещей / Леонид Восков. Date Views 28.01.19 https://postnauka.ru/talks/80081

3. Satellite technologies for IoT applications / IoTUK. Date Views 28.01.19 https://iotuk.org.uk/wp-content/uploads/2017/04/Satellite-Applications.pdf

4. Особенности радиопокрытия LoRa / iTech Технологии связи. Date Views 28.01.19 https://itechinfo.ru/content/особенности-радиопокрытия-lora

5. LoRa против «Стрижа»: сравнение технологий/ iot.ru. Date Views 28.01.19 iot.ru – Режим доступа: https://iot.ru/gorodskaya-sreda/lora-protivstrizha-sravnenie-tekhnologiy, свободный

6. Д.Л. Кумаритова, Р.В. Киричек, 2016. Обзор и сравнительный анализ технологий LPWAN сетей, СПбГУТ, Date Views 28.01.19 http://www.sut.ru/doci/nauka/review/20164/33-48.pdf

7. Строительство сетей LoRa в РФ / iTech Технологии связи. Date Views 28.01.19 https://itechinfo.ru/content/строительство-сетей-lora-в-рф

8. An Overview of Narrowband IoT (NB-IoT). Date Views 28.01.19 https://www.link-labs.com/blog/overview-of-narrowband-iot

9. Спутниковые системы Часть 1. Орбиты и параметры / Сети Network world. Date Views 28.01.19 https://www.osp.ru/nets/1999/01-02/143941/

10. Sattelite Coverage | Our Satellite | Inmarsat. Date Views 28.01.19 https://www.inmarsat.com/about-us/our-satellites/our-coverage/

11. IsatM2M. Date Views 28.01.19 https://www.marsat.ru/en/satelliteequipment-isatm2m.

12. IsatData Pro. Date Views 28.01.19 https://www.marsat.ru/en/satelliteequipment-isatdatapro.

13. BGAN M2M. Date Views 28.01.19 https://www.marsat.ru/en/satelliteequipment-bgan-m2m.

14. Iridium 9602 | Iridium Satellite Communications. Date Views 28.01.19 https://www.iridium.com/products/iridium-9602/.

15. Iridium 9603 | Iridium Satellite Communications. Date Views 28.01.19 https://www.iridium.com/products/iridium-9603-3/.

16. Iridium Core 9523 | Iridium Satellite Communications. Date Views 28.01.19 https://www.iridium.com/products/iridium-core-9523/.

17. Introducing JSON / JSON. Date Views 28.01.19 http://www.json.org.

18. MQTT Version 5.0. Date Views 28.01.19 http://docs.oasis-open.org/mqtt/mqtt/v5.0/cs01/mqtt-v5.0-cs01.html.

19. BSON – Binary JSON / BSON. Date Views 28.01.19 http://bsonspec.org.

20. C. Bormann and P. Hoffman, 2013. Concise Binary Object Representation (CBOR). Internet Engineering Task Force (IETF). Date Views 28.01.19 https://tools.ietf.org/html/rfc7049.

21. MessagePack. It's like JSON, but fast and small. Date Views 28.01.19 http://msgpack.org.

22. JSON compressor and decompressor / tcorral. Date Views 28.01.19 https://github.com/tcorral/JSONC.

23. Protocol Buffers / Google Developers. Date Views 28.01.19 https://developers.google.com/protocolbuffers/.

УДК 519.25 004.852

К.И. Завалич, В.А. Мельников

ПРОГНОЗИРОВАНИЕ СПРОСА НА ПРОДУКЦИЮ ПРОИЗВОДСТВЕННОГО ПРЕДПРИЯТИЯ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Аннотация

В статье поставлена задача прогнозирования объема продаж для промышленного предприятия. Актуальность исследования обусловлена тем, что точность прогноза для подобного предприятия имеет принципиально важное значение. Прогноз спроса, находясь в самом начале цепочки поставок, оказывает большое влияние на расходы компании. Цель исследования состоит в повышении точности прогнозирования спроса на предприятии. Приведена классификация методов и моделей прогнозирования временных рядов. Подробнее рассмотрены методы, применяемые для исследования, а именно авторегрессионные модели, модели экспоненциального сглаживания и нейронные сети. Описаны достоинства и недостатки их применения. Обосновывается выбор метрик качества для оценки адекватности получившегося прогноза и его сравнения с показателями, достигнутыми на предприятии. В качестве исходных данных для построения прогнозов используются месячные данные фактических продаж за последние шесть лет. Исходная выборка делится на обучающую и тестовую. Метрики качества считаются на последних 12 месяцах. Для проведения эксперимента отобраны шесть изделий, по данным которых были построены модели прогнозирования. На основании проведенного эксперимента делаются выводы о необходимости дальнейшего ручного подбора параметров моделей для достижения цели работы. Направления будущего исследования относятся к исследованию возможности применения других методов прогнозирования временных рядов и попыткам распространить полученные на шести изделиях результаты на весь товарный ряд предприятия.

Ключевые слова: прогнозирование спроса на товар, анализ временных рядов, метрики качества прогноза, B2B, нейронные сети, *impact mapping, IBM SPSS, Deductor*.

Введение

Объектом нашего исследования является прогноз продаж производственного предприятия – поставщика решений и производителя оборудования в области промышленной автоматизации техпроцессов.

Высокая точность прогноза спроса для такого предприятия является залогом отсутствия потерь во всей цепочке поставок и одной из важнейших задач для оптимальной работы компании в целом. Ведь превышение прогноза над фактом может привести к излишне замороженные средствам, закупке ненужных материалов и производству ненужных полуфабрикатов, затовариванию складов. Наоборот, превышение факта над прогнозом приводит к отсутствию необходимых запасов продукции и к упущенной прибыли в следствие нереализованных возможностей. Подробная схема с причинами потерь указана на рисунке 1.

Рис. 1. Причины потерь [1]

Целью исследования является повышение точности прогнозирования спроса по каждому продукту предприятия на более чем 5 %. Точностью прогнозирования для целей нашего исследования будем считать среднемесячное относительное отклонение прогноза от факта (MAPE):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Yt - \widehat{Yt}|}{Yt} \quad (1)$$

Y_t – фактический объем продаж за анализируемый период;

$\hat{Y}_t$ — значение прогнозной модели за анализируемый период;

n — количество периодов.

Для достижения цели исследования будет использован инструментарий анализа временных рядов.

Методы прогнозирования временных рядов

Несмотря на то, что существует множество методов прогнозирования и множество литературы, посвященной описанию этих методов, например [2, 3, 4], на исследуемом предприятии используется экспертный метод и коэффициент сезонности, исчисляемый по историческим данным. Использование такого сочетания при прогнозировании спроса – это обычная практика согласно исследованию [5]. И это значит, что есть потенциал для улучшения точности прогнозирования спроса на исследуемом предприятии за счет использования более подходящих методов.

Процесс прогнозирования спроса в B2B – это сложный процесс, имеющий ряд особенностей и отличий от прогнозирования в B2C, обусловленных особенностями процесса покупки в B2B. [6, 7] Многие авторы занимаются исследованием прогнозирования спроса для B2C. В статье [8] авторы используют Гауссовский процесс в отрасли FMCG (товары повседневного спроса). В работе [9] применяется свёрточная нейронная сеть для повышения точности прогнозирования спроса в области электронной коммерции. Статья [10] описывает линейную регрессию, ARIMA алгоритм и XGBoost алгоритм машинного обучения, а также комбинированные модели для прогнозирования продаж на примере магазина. Но по причине того, что между процессом покупки в B2B и B2C существует множество отличий, то для целей нашего исследования рассмотрим только те научные работы [11–18], которые проведены в целях повышения точности прогнозирования спроса в области исключительно B2B.

Выбор методов зависит от специфики конкретной задачи, в нашем исследовании мы будем предсказывать значение временного ряда и использовать методы, которые показали лучшие результаты в исследованиях [11–18], а именно:

1. Нейросетевые модели и методы. Достоинства моделей – это нелинейность, адаптивность, масштабируемость и единообразие их анализа и проектирования [19]. При этом недостатками являются отсутствие прозрачности моделирования, сложность выбора архитектуры, высокие требования к непротиворечивости обучающей выборки, сложность выбора алгоритма обучения и ресурсоемкость процесса их обучения [20].

2. Авторегрессионные модели и методы. Важными достоинствами данного класса моделей являются их простота, прозрачность моделирования и единообразие анализа и проектирования. Недостатками данного класса моделей являются большое число параметров модели, идентификация которых неоднозначна и ресурсоемка [21], низкая адаптивность моделей, а также линейность и, как следствие, отсутствие способности моделирования нелинейных процессов, часто встречающихся на практике [22].

3. Модели и методы экспоненциального сглаживания. Достоинствами данного класса моделей являются простота и единообразие их анализа и проектирования. Недостатком данного класса моделей является отсутствие гибкости [23].

Для оценки качества применяемых методов исходный временной ряд разделяется на обучающий и тестовый: на обучающем отрезке рассчитывается модель, на тестовом – вычисляется заданный показатель качества, который сравнивается с показателем качества прогноза, существующем на предприятии за аналогичный период.

В работе [24] рассмотрены многие метрики качества прогнозирования временных рядов. Однако большинство из этих мер страдают от одной или нескольких проблем. Для целей нашего исследования предлагается использовать три меры и сравнивать их результаты между собой:

1. MAPE не обладает устойчивостью к выбросам, так как основана на арифметическом среднем, но зато эта метрика проста и информативна. На практике слабая устойчивость к прогнозированию выбросов может приводить к ошибочным результатам.

2. GMRAE показал замечательную стойкость к прогнозированию выбросов. Однако одним из недостатков мер, основанных на среднем геометрическом значении, является то, что прогнозы с нулевой ошибкой должны быть исключены. Таким образом, эти меры могут быть недостаточно информативными.

$$GMRAE = \left(\prod_{t=1}^{n}\left|\frac{et}{et*}\right|\right)^{\frac{1}{n}} = \exp\left(\frac{1}{n}\sum_{t=1}^{n} ln\left(\frac{et}{et*}\right)\right) \quad (2)$$

Пусть et* обозначает ошибку прогнозирования во время t, полученную некоторым эталонным методом. Это означает et*=(Yt−Ft*), где Ft*-прогноз в момент времени t эталонным методом. В качестве эталонного метода может быть использован «Наивный метод», то есть прогноз во время t равен факту во время t-1

3. Третьей мерой попробуем использовать UMBRAE, так как она с одной устойчива к выбросам, с другой стороны, нет необходимости удалять нулевые ошибки. Для случаев, когда такие меры, как GMRAE являются предпочтительными, UMBRAE может быть альтернативной мерой, так как она гораздо проще в использовании без необходимости обрезать нулевые ошибки.

$$MBRAE = mean\ BRAE = \frac{1}{n}\sum_{t=1}^{n}\frac{|et|}{|et| + |et*|} \quad (3)$$

$$UMBRAE = \frac{MBRAE}{1 - MBRAE} \quad (4)$$

Результаты исследования

Входные данные предприятия состоят из фактических продаж исследуемого предприятия за 2012–2018 годы. Эксперты с предприятия выделили 6 (шесть) наиболее важных изделий, на которых был проведен эксперимент по выбранным методам прогнозирования.

Авторегрессионные модели и модели экспоненциального сглаживания. В программе IBM SPSS был использован модуль «эксперт построения моделей временных рядов», который подобрал оптимальные из моделей АРПСС (модели авторегрессии и проинтегрированного скользящего среднего) и экспоненциального сглаживания. Для всех изделий, кроме 2-го была выбрана модель сезонная АРПСС, для 2-го – простая сезонная экспоненциальное сглаживание.

Чтобы проанализировать адекватность получившихся моделей, можно сравнить имеющиеся наблюдения с предсказанными значениями, полученными с помощью подобранной модели и посчитать метрики точности, что будет сделано дальше. А также можно проанализировать остатки, то есть разницу между наблюдениями и предсказанными значениями. Модель адекватна, если выполняется три требования:

Остатки несмещенные, то есть в среднем равны нулю. Остатки по всем 6 (шести) изделиям несмещенные.

Остатки стационарны, то есть отсутствует зависимость от времени, не меняется дисперсия и остатки не меняются систематически. Не по всем изделиям остатки стационарны, по изделию 2 систематически меняется закон распределения.

Остатки независимы, неавтокоррелированы. Для проверки этого критерия строим коррелограммы остатков. По изделиям 1 и 6 наблюдаем автокоррелированность остатков, есть выбросы за пределы доверительного коридора.

В результате эксперимента рассмотрим значения трех метрик, посчитанным по 6 изделиям. Расчёт проведен на данных за последние 12 месяцев и представлен в таблице 1.

Таблица 1

Метрики качества прогнозирования спроса

MAPE	18 %	72 %	44 %	42 %	419 %	112 %
GMRAE	1,31	0,71	0,60	0,51	0,66	0,89
UMBRAE	1,12	0,80	0,83	0,55	0,72	0,92
MAPE орг	14 %	63 %	40 %	52 %	355 %	166 %
GMRAE орг	0,76	0,72	0,76	1,00	0,62	1,15
UMBRAE орг	1,00	0,87	0,79	1,00	0,77	1,14

Сравнивая с показателями, достигнутыми на предприятии за этот же период соответственно (MAPE **орг,** GMRAE **орг,** UMBRAE **орг**), делаем вывод о том, что показатели предприятия лучше по 4-ем из 6-ти изделий.

Нейронные сети. Был использован сценарий в Deductor, который представляет собой последовательность следующих действий:

Редактирование аномалий и выбросов (стандартное отклонение от среднего для выбросов = 3, для экстремумов = 5);

Проведение спектрального анализа для освобождения данных от шумов методом Фурье-преобразования, степень вычитания шума – малая;

Построение скользящего окна глубиной 12 месяцев согласно выявленной сезонности и экспертному мнению сотрудников предприятия.

Проведение обучения нейронной сети на 80 % данных и её тестирование на 20 % данных со следующими параметрами:

Многослойный персептрон;

Нейроны во входном слое – 12, выходном – 1, скрытых слоев – 1 (2 нейрона в скрытом слое);

Тип функции – Сигмоида, крутизна 1;

Алгоритм – Resilient Propagation, шаг спуска 0,5, шаг подъёма 1,2.

Выгрузка данных в MS Excel, проведение расчета выбранных метрик качества прогнозирования (табл. 7) и построение графиков (рис. 2), где синей линией отражаются фактические данные, а красной – прогноз нейронной сети.

Результаты эксперимента представлены в таблице 2. Расчёт проведен по данным за последние 12 месяцев.

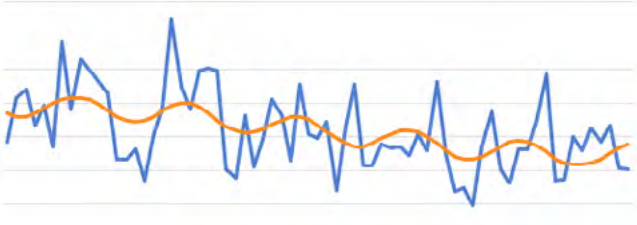


Рис. 2. Фактические и прогнозные данные по изделию 1

Таблица 2

Метрики качества прогнозирования спроса

Метрика	Изделие 1	Изделие 2	Изделие 3	Изделие 4	Изделие 5	Изделие 6
MAPE	20 %	56 %	45 %	54 %	245 %	127 %
GMRAE	1,73	0,62	0,93	0,58	0,65	0,50
UMBRAE	1,43	0,72	0,95	0,68	0,68	0,76
MAPE орг	14 %	63 %	40 %	52 %	355 %	166 %
GMRAE орг	0,76	0,72	0,76	1,00	0,62	1,15
UMBRAE орг	1,00	0,87	0,79	1,00	0,77	1,14

Сравнивая с показателями, достигнутыми на предприятии за этот же период соответственно (MAPE **орг,** GMRAE **орг,** UMBRAE **орг**), делаем вывод о том, что показатели предприятия хуже по 4-ем из 6-ти изделий.

Выводы

В результате проведения эксперимента методами АРПСС, экспоненциального сглаживания и нейронных сетей лучший результат показали нейронные сети, по изделиям 2, 4, 5 и 6. По изделиям 1 и 3 – на исследуемом предприятии прогнозируют точнее.

Таким образом, на данный момент цель исследования – повысить точность прогнозирования на 5 % по всем изделиям – не достигнута. Но в рамках этой итерации были использованы стандартные параметры методов, не был проведен тщательный ручной подбор параметров. На очереди следующая итерация и следующий эксперимент, который будет реализован на языке программирования Python и в рамках которого:

Будет проведен ручной подбор параметров по тестируемым методам;

Будут протестированы дополнительные методы для анализируемых изделий (XGBoost, Facebook Prophet).

Список литературы

1. Deductor Demand Planning [Электронный ресурс] // URL: https://basegroup.ru/system/files/slideshare/deductor-demand-planing.pdf (дата обращения: 20.08.2018).
2. Forecasting Methods and Principles: Evidence-Based Checklists [Электронный ресурс] // URL: https://www.researchgate.net (дата обращения: 20.08.2018).
3. The design features of forecasting support systems and their effectiveness [Электронный ресурс] // URL: https://www.sciencedirect.com (дата обращения: 20.08.2018).
4. Time Series Analysis: Forecasting and Control [Электронный ресурс] // URL: https://books.google.ru (дата обращения: 20.08.2018).
5. Improving Forecast Quality in practice [Электронный ресурс] // URL: http://orca.cf.ac.uk (дата обращения: 20.08.2018).
6. What are the key differences between B2C and B2B ecommerce? [Электронный ресурс] // URL: https://www.ecommercewiki.org (дата обращения: 20.08.2018)
7. The Differences & Similarities Between B2B and B2C Markets and Selling [Электронный ресурс] // URL: http://www.dwsassociates.com (дата обращения: 20.08.2018).
8. Forecasting of commercial sales with large scale Gaussian Processes [Электронный ресурс] // URL: https://arxiv.org/pdf/1709.05548.pdf (дата обращения: 20.08.2018)
9. Sales Forecast in E-commerce using Convolutional Neural Network [Электронный ресурс] // URL: https://arxiv.org/pdf/1708.07946.pdf (дата обращения: 20.08.2018).
10. Linear, Machine Learning and Probabilistic Approaches for Time Series Analysis [Электронный ресурс] // URL: https://arxiv.org/pdf/1703.01977.pdf (дата обращения: 20.08.2018).
11. Organizational Learning Supported by Machine Learning Models Coupled with General Explanation Methods: A Case of B2B Sales Forecasting [Электронный ресурс] // URL: https://www.degruyter.com (дата обращения: 20.08.2018).
12. Sales pipeline win propensity prediction: a regression approach [Электронный ресурс] // URL: https://arxiv.org/pdf/1502.06229.pdf (дата обращения: 20.08.2018).
13. Analysis & prediction of sales data in sap-erp system using clustering algorithms [Электронный ресурс] // URL: https://arxiv.org/pdf/1312.2678.pdf (дата обращения: 20.08.2018).
14. Inter Time Series Sales Forecasting [Электронный ресурс] // URL: https://arxiv.org/pdf/1303.0117.pdf (дата обращения: 20.08.2018).
15. Sales forecasting using WaveNet within the framework of the Kaggle competition [Электронный ресурс] // URL: https://arxiv.org/pdf/1803.04037.pdf (дата обращения: 20.08.2018).

16. Метод машин опорных векторов для прогнозирования объема инвестиций [Электронный ресурс] // URL: https://cyberleninka.ru (дата обращения: 20.08.2018)

17. PSO based Neural Networks vs. Traditional Statistical Models for Seasonal Time Series Forecasting [Электронный ресурс] // URL: https://arxiv.org/pdf/1302.6615.pdf (дата обращения: 20.08.2018)

18. An Introductory Study on Time Series Modeling and Forecasting [Электронный ресурс] // URL: https://arxiv.org/pdf/1302.6613.pdf (дата обращения: 20.08.2018)

19. Хайкин С. Нейронные сети: полный курс. М.: ООО «И.Д. Вильямс», 2006. 1104 с

20. Mazengia D.H. Forecasting Spot Electricity Market Prices Using Time Series Models: Thesis for the degree of Master of Science in Electric Power Engineering. Gothenburg, Chalmers University of Technology, 2008. 89 p.

21. Morariu N., Iancu E., Vlad S. A neural network model for time series forecasting // Romanian Journal of Economic Forecasting. 2009, No. 4. P. 213–223.

22. Pradhan R.P., Kumar R. Forecasting Exchange Rate in India: An Application of Artificial Neural Network Model // Journal of Mathematics Research. 2010, Vol. 137 2, No. 4. P. 111–117.

23. Alfares H.K., Nazeeruddin M. Electric load forecasting: literature survey and classification of methods // International Journal of Systems Science. 2002, Vol 33. P. 23–34.

24. A new accuracy measure based on bounded relative error for time series forecasting [Электронный ресурс] // URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5365136/ (дата обращения: 20.08.2018)

УДК 004.65.001.56

К.М. Карадобри, Ю.В. Петриченко, С.Н. Полторак

АВТОМАТИЗАЦИЯ БИЗНЕС-ПРОЦЕССОВ ВЕДЕНИЯ РЕКЛАМНЫХ КАМПАНИЙ КАК ИННОВАЦИЯ В СФЕРЕ ИНТЕРНЕТ-МАРКЕТИНГА

Аннотация

Данная статья рассматривает проблему автоматизации бизнес-процессов отделов рекламы в интернет-агентствах. Был проведен анализ основных направлений рекламного рынка в России за 2018 год, определены наиболее популярные способы продвижения бизнеса, такие как SEO и контекстная реклама. Были рассмотрены и изучены действующие способы автоматизации определенных процессов и выявлены недочеты в данных системах. Также был проведен опрос среди специалистов по интернет-рекламе для определения основных задач, которые интернет-маркетологи хотели бы решить с помощью автоматизации процессов в своей работе. Такой опрос позволил избежать ошибок и упущений, которые были выявлены при анализе конкурентных автоматизированных систем, а также ответил на определенный запрос специалистов о создании системы, максимально удобной и адаптированной под конкретные цели и задачи. В ходе анализа полученных результатов был составлен список требуемого функционала для создания универсальной системы, которая подошла бы многим digital-агентствам и могла бы быть успешно внедрена в любое из них. В итоге всех проведенных мероприятий был предложен и внедрен свой способ автоматизации работы интернет-маркетологов, основанный на полученных данных. Данная система была реализована в одном из региональных интернет-агентств и продемонстрировала значимость оптимизации бизнес-процессов для компаний, которые работают в сфере интернет-маркетинга.

Ключевые слова: Продвижение, контекстная реклама, SEO, Интернет, интернет-технологии, digital.

Введение

Интернет сейчас – это огромный пласт для развития любого бизнеса в 21 веке. Большинство компаний обращаются в digital-агентства для продвижения своего сайта или групп в социальных сетях, чтобы привлекать новых клиентов и повысить лояльность действующих. В связи с этим агентства сталкиваются с обработкой большого объёма данных, так как с каждым годом интернет-реклама только набирает обороты и всё больше компаний хотят продвигать свой бизнес через различные интернет-площадки. В результате этого остро встаёт вопрос об автоматизации рутинных процессов, на которые специалисты тратят значительную часть времени (например: составление отчётности в ручную. Из-за данной проблемы digital-агентства теряют прибыль, так как вместо выполнения задач, которые направлены на увеличение числа клиентов или расширения списка услуг (например: увеличение рекламируемых площадок, сквозная аналитика, подключение рассылок и многие другие процессы, интернет-маркетологам приходится нецелесообразно тратить время на процессы, которые требуют автоматизации. Для определения основных проблем и для постановки конкретных задач для автоматизации было решено проанализировать рекламный рынок России, а также провести анкетирование среди специалистов в сфере интернет-маркетинга. Объектом исследования было выбрано 5 региональных агентств и 6 специалистов.

Основная часть

В 2018 году по итогам 3-х кварталов, digital занимает лидирующее место среди остальных видов рекламы. Также, рынок интернет-рекламы впервые обогнал телевидение по объёму продаж.

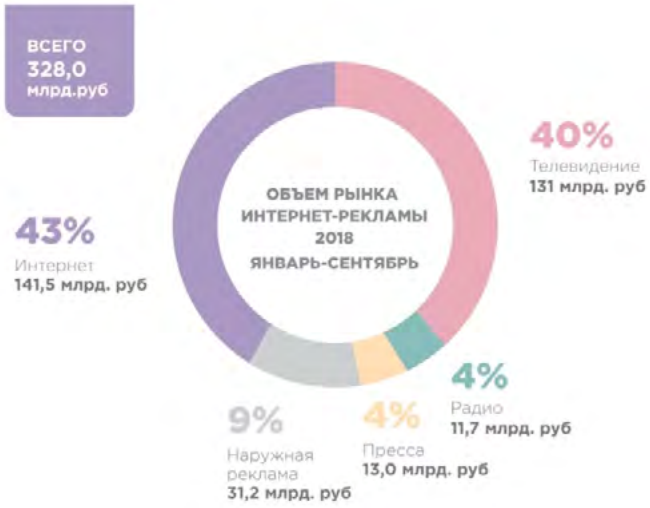


Рис. 1. Объём рынка рекламы в 2018 году по итогам 3-х кварталов, источник: АКАР (Ассоциация Коммуникационных Агентств России)

Как мы видим из рисунка 1, на данный момент интернет-реклама является наиболее востребованным инструментом для продвижения бизнеса. Большим плюсом, в отличие от других видов рекламы, является то что компания может не только нацелиться на потенциальных или действующих покупателей, но и получить детальную информацию о продвижении, провести анализ и скорректировать стратегию рекламной кампании.

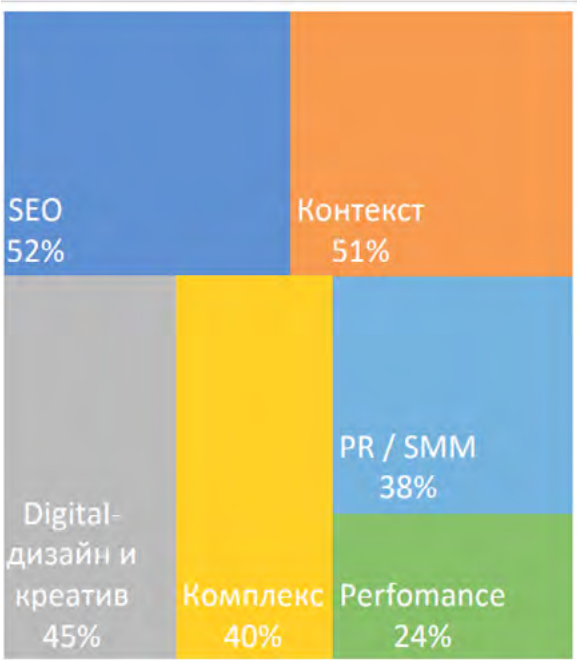


Рис. 2. Структура рынка интернет-маркетинга в 2017 году и прогноз на 2018 год, источник: RUWARD

Определим наиболее популярные услуги в рекламных агентствах, для этого рассмотрим структуру рынка интернет-маркетинга за 2018 год.

Как мы видим из рисунка 2, наиболее популярными способами продвижения в Интернете являются:

• SEO-продвижение (Search Engine Optimization – поисковая оптимизация, которая включает в себя комплекс мер для повышения позиций выдачи сайта в е по заранее подобранным запросам и включает в себя адаптацию страниц сайта для наиболее эффективной выдачи.

• Контекстная реклама – это вид интернет-рекламы, который включает в себя поисковую и баннерную рекламу. Объявления на поиске отображаются на страницах поисковой выдачи Яндекс и Google под строкой поиска или чуть выше футера страницы. Баннерные объявления отображаются на сайтах-партнёрах поисковых систем в специальных блоках.

Несмотря на то что контекстная реклама является одним из первых видов рекламы, по данным РАЭК (Ассоциация электронных коммуникаций и годовых отчётов Яндекса, с каждым годом она растёт и на данный момент этот рост составляет 17–25 % (рис. 3).

Рассмотрим интернет-рекламу с точки зрения продажи её как услугу.

На данный момент большинству digital-агентств приходится сталкиваться с обработкой большого объёма данных, так как всё больше компаний обращаются за профессиональными услугами интернет-маркетологов.

Для выявления основных проблем, с которыми сталкиваются специалисты при обработке большого объёма данных, было проведено анкетирование, в котором участвовало 6 специалистов 5 региональных агентств (табл. 1).

По результатам опроса было выявлено, что в большинстве digital-агентств сейчас составление отчетности по рекламным кампаниям подготавливается вручную. Для того, чтобы подготовить отчёт, специалисту необходимо зайти в каждую систему и собрать все данные в один отчёт. По подсчетам, специалист, который, в среднем, ведёт 25 клиентов, тратит около 16 часов в месяц на составление всех отчётов. Для того, чтобы минимизировать время на составление отчётов и повысить лояльность клиентов агентства, было решено автоматизировать такие процессы как составление отчётов, отслеживание эффективности рекламных кампаний и напоминание об окончании акций клиентов.

Перед началом разработки системы автоматизации, был проанализирован рынок альтернативных систем и удалось найти только системы, через которые нужно полностью вести рекламу, такие системы как Elama, Begun, Aori. В этих системах можно вести рекламу сразу в нескольких системах и анализировать статистику сразу по всем системам. Но данные

Рис. 3. Объём рынка контекстной рекламы в России 2015–2018 гг., источник: RUWARD.

Таблица 1

Анкетирование интернет-маркетологов региональных агентств

Вопрос	Агентство 1	Агентство 2	Агентство 3	Агентство 4	Агентство 5
Вы отслеживаете РК в каждой системе отдельно?	+	+	+	+	+
Назовите число рекламных систем, в которых Вы ведете клиентов	8	4	2	2	3
У Вас автоматизировано составление отчетов?	–	+	–	–	–
Вам приходят напоминания об окончании акции	–	–	+	–	–
Вы отслеживаете сколько часов тратится на каждого клиента в месяц?	–	–	+	+	+
У Вас автоматизировано отслеживание показателей по сертификации и срок окончания сертификатов?	–	–	–	–	–
Хотели бы Вы автоматизировать отслеживание ведения РК клиентов?	+	+	+	+	+

сервисы не подходят профильным агентствам, так как они не заменяют такие системы как Яндекс.Директ, Google Ads. Они ограничены функционалом по сравнению с официальными площадками. Кроме того, в данных системах нельзя полностью контролировать активность рекламных кампаний, система сама откручивает объявления и не всегда это делает эффективно, из-за чего наблюдается снижение конверсии. Поэтому, ни одна из данных систем автоматизации не подходит для реализации наших целей и задач. Перед началом постановки технического задания, кроме анкетирования интернет-маркетологов были составлены бизнес-требования генеральным директором и интернет маркетологом агентства, в котором планировалась и в дальнейшем была внедрена наша система (рис. 4).

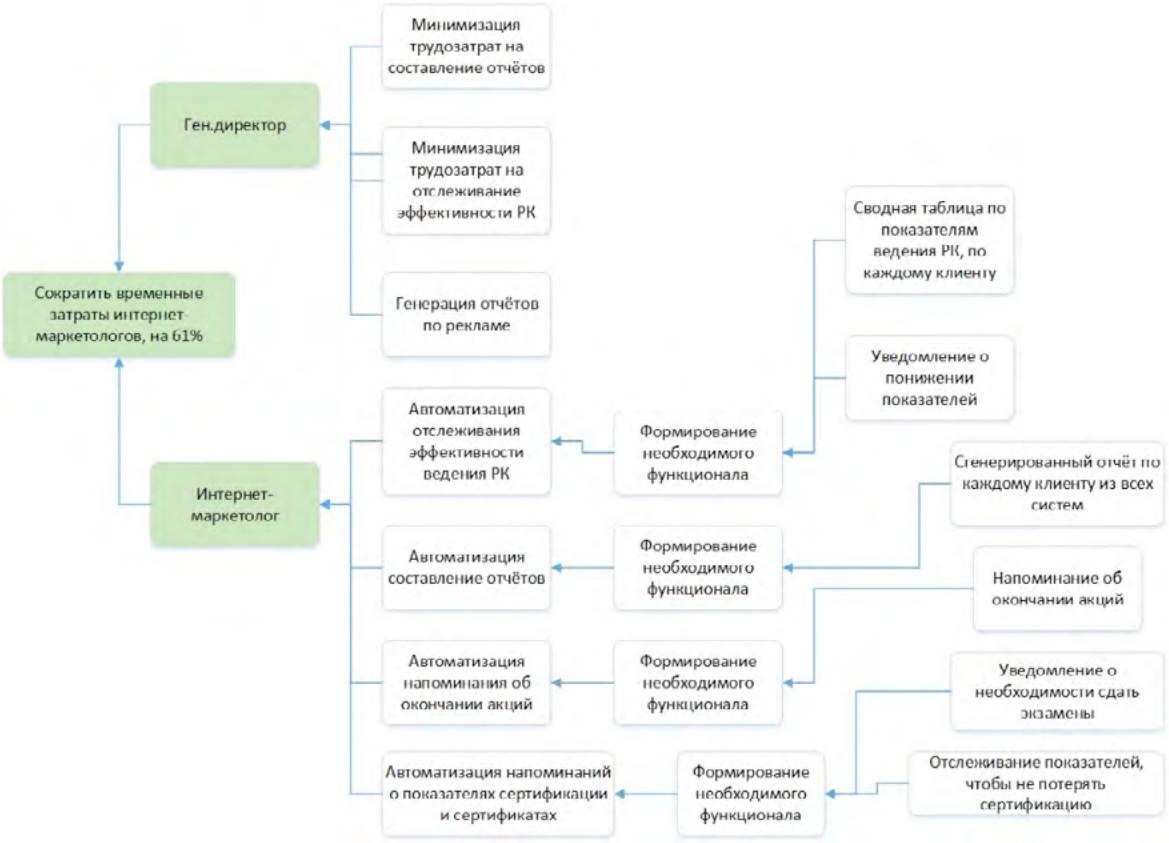


Рис. 4. Бизнес-требования (Impact Mapping)

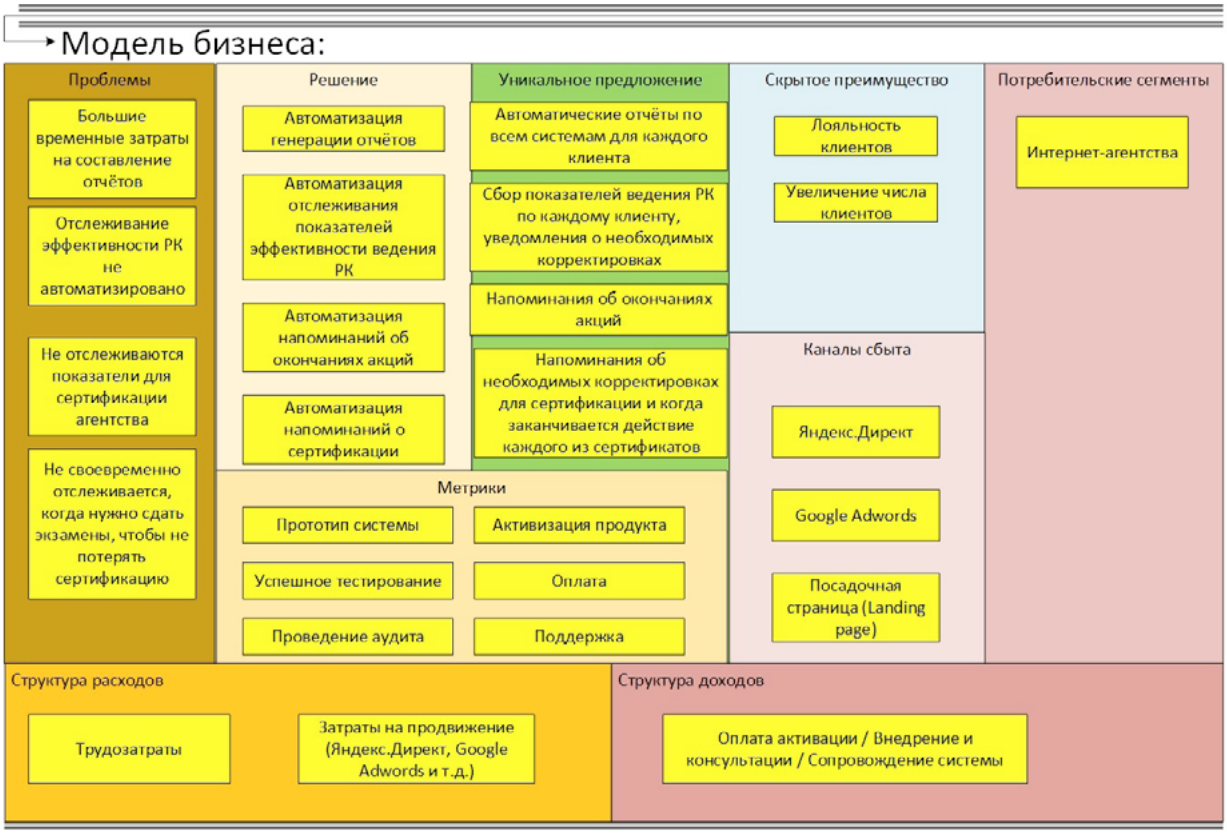


Рис. 5. Решение (Lean Canvas)

После определения бизнес-требований было предложено следующее решение (рис. 5).

В результате проведенного анализа было составлено техническое задание на создание системы, которая отвечала бы требованиям большинства агентств, была гибкой и встраиваемой в любое из них.

В апреле 2018 года система автоматизации была внедрена в выбранное агентство. После внедрения данной системы мы получили сокращение временных затрат на отслеживание эффективности ведения рекламных кампаний, составление отчётов, а также было добавлено информирование специалистов об окончании акций для конкретного клиента (табл. 2).

Все данные, которые представлены в таблице 2 были зафиксированы специалистами в системе управления «Битрикс 24» до начала внедрения и после. Как мы видим из таблицы, время на отслеживание сократилось на 95,23 %, время на составление отчётов сократилось на 98,44 %.

Таблица 2

Показатели эффективности внедрения автоматизации процессов

Наименование	**Тратилось времени до внедрения, часы**	**Тратится время после внедрения, часы**	**%**	**Сократилось / Результат**
Отслеживание эффективности ведения РК	87,5	4,17	95,23	18 228,44
Составление отчетов	16	0,25	98,44	3 445,31
Окончание акции	Не отслеживалось	0,1		Лояльность клиента
Отслеживание затраченного времени на каждого клиента	Не отслеживалось	1		• Контроль работы интернет-маркетологов • Увеличение сопровождения (2 клиента в месяц) = 2 000 руб.
Отслеживание сертификации агентства	Не отслеживалось	0,25		Сертифицированное агентство
Прибыль/экономия, руб.				

Выводы

Научная новизна заключается в создании автоматизированной системы, которой на данный момент нет на российском рынке и её внедрении в одном из рекламных агентств. Благодаря автоматизации таких процессов как составление отчётов и отслеживание данных по рекламным кампаниям, нам удалось сократить не только временные затраты, но и увеличить доход компании. А это значит, что все поставленные цели и задачи были реализованы и успешно внедрены на конкретном примере.

Список литературы

1. Анар Бабаев, Николай Евдокимов, Андрей Иванов, 2011 год. Контекстная реклама. Издательство «Питер», 304 с.
2. Катернюк А.В., Марченко О.Г. Современные рекламные технологии: коммерческая реклама. Дата обращения 30.01.2019 https://abc.vvsu.ru/books/u_reklama/page0026.asp#xex33.
3. Джесси Стрикчиола, Рэнд Фишкин, 2014 год. SEO – искусство раскрутки сайтов. Издательство «Питер», 810 с.
4. Ассоциация Коммуникационных Агентств России, объём рынка цифрового контента в России составил 75 млрд. рублей в 2018 году. Дата обращения 17.12.2018 https://adindex.ru/news/digital/2018/12/13/228137.phtml.
5. Ассоциация Коммуникационных Агентств России, рунет обошёл ТВ по рекламным доходам в первом полугодии. Дата обращения 17.12.2018 https://adindex.ru/news/researches/2018/08/22/173671.phtml.
6. Что такое SEO-оптимизация. Дата обращения 26.01.2019 https://www.i-media.ru/seo/searchengineoptimization.
7. Маркетинг в социальных сетях. Дата обращения 26.01.2019 https://ru.wikipedia.org/wiki/Маркетинг_в_социальных_сетях.
8. Интернет-рекламы, рынок России. Дата обращения 26.01.2019 http://www.tadviser.ru/index.php/Статья:Интернет-реклама_%28рынок_России.
9. Тренды интернет-рекламы-2019: технологии есть, знаний не хватает. Дата обращения 30.01.2019 https://vc.ru/marketing/52183-trendy-internet-reklamy-2019-tehnologi-est-znaniy-ne-hvataet.
10. Cybermarketing 2018: Изнанка бизнеса агентств контекстной рекламы. Как выжить в новых реалиях. Дата обращения 30.01.2019 https://www.seonews.ru/analytics/cybermarketing-2018-iznanka-biznesa-agentstv-kontekstnoy-reklamy-kak-vyzhit-v-novykh-realiyakh/
11. Russian Digital Profile 2018. Дата обращения 31.01.2019 http://www.ruward.ru/russian-digital-profile-2018/
12. Экономика рунета. Дата обращения 31.01.2019 https://raec.ru/activity/analytics/9884/

УДК 372.862

З.Д. Леонтьев, И.И. Башаров

СОВРЕМЕННЫЕ МЕТОДЫ И ТЕХНОЛОГИИ СОЗДАНИЯ ВИРТУАЛЬНЫХ ЛАБОРАТОРНЫХ РАБОТ ДЛЯ ТЕХНИЧЕСКИХ СПЕЦИАЛЬНОСТЕЙ

Аннотация

В статье большое внимание уделено описанию использования современных компьютерных программ для разработки виртуальных лабораторных работ. Проведённые исследования подчёркивают актуальность настоящей работы. Авторами сделан акцент на применение в учебном процессе современных информационных и мультимедийных образовательных ресурсов. Виртуальные лабораторные работы, как важная составляющая образовательного процесса, позволяют студентам приобрести навыки самостоятельного решения практических задач, осваивать материал дистанционно при заданных конкретных условиях, а также минимизируют погрешность измерений.

Ключевые слова: виртуальная лабораторная работа, лабораторный практикум, Adobe Flash, LABView, Algodoo, дистанционное обучение.

Введение

Выполнение лабораторных является важной частью учебного процесса, где студенты приобретают навыки самостоятельного решения практических задач. В настоящее время разрабатывается огромное количество виртуальных лабораторных работ (ВЛР), особенно для образовательно-практических целей. Данную тему затрагивают многие авторы и разработчики. Например, Мурзин С.В. в статье «Виртуальные приборы в лабораторном практикуме» [1] пишет, что лабораторный практикум с использованием виртуальных приборов представляет собой один из прогрессивно развивающихся видов проведения лабораторных занятий. «Виртуальные лабораторные работы – это насущная необходимость, позволяющая в значительной мере оптимизировать и поднять на должный уровень учебный процесс по техническим дисциплинам». Также, данную тему поднимали авторы Савкина Т.А. и Вайда О.В. в работе «Виртуальный лабораторный практикум» [2]. В ней рассматривается понятие виртуального лабораторного практикума, его применение в инженерном образовании, преимущества и недостатки, а также производится анализ существующих систем дистанционного обучения. Практическую значимость подтвердила Бочкарева Ю.Г. в работе «Формирование информационной компетенции бакалавров в виртуальном лабораторном практикуме по физике» [3]. В статье обсуждаются проблемы формирования компетенций при изучении курса физики с использованием современных информационных технологий. Учёные и изобретатели [4] предлагают использовать в учебном процессе живые лаборатории («living labs»). Авторы отмечают преимущества живых лабораторий: эффективность, управляемость, время, устойчивость, масштабируемость и непредсказуемость результатов. В результате студенты получают выгоды в виде огромного разнообразия материальных и нематериальных инноваций. В работе [5] описана визуализация удаленных лабораторий при обучении дисциплинам, связанных с электроникой и смежных дисциплин. Литературный анализ [6–15] показал, что ВЛР являются весьма эффективными на всех уровнях и направлениях обучения, что доказывается приобретаемыми компетенциями и практическими навыками. Авторами [16] научно обоснованы результаты более успешного освоения профессиональных компетенций студентами университета в первый год обучения.

Актуальность данного исследования заключается в потребности проведения лабораторной работы дистанционно при конкретных обстоятельствах, а также необходимости минимизировать погрешность, проводимых на занятии, измерений. Удовлетворению этих требований соответствует внедрение в учебный процесс информационных и мультимедийных образовательных ресурсов, таких как виртуальные лабораторные работы.

Методика

Целью работы является анализ современных компьютерных программ для разработки виртуальных лабораторных работ. Достижение поставленной цели возможно при решении следующих задач:

1. провести анализ средств разработки виртуальных лабораторных работ;
2. создать виртуальную лабораторную работу с помощью выбранной программной среды в соответствии с техническим заданием.

Основная часть

Виртуальная лабораторная работа – информационная система, интерактивно моделирующая реальный технический объект и его существенные для изучения свойства с применением компьютерной виртуализации [17].

Для проектирования и дальнейшего использования виртуальной лабораторной работы определены ее структурные элементы: последовательность выполнения задания, ключевые компоненты системы,

функционал элементов пользовательского интерфейса. ВЛР разрабатываются с использованием самых современных, технически развитых программных средств [18, 19]. В соответствии с требованиями к виртуальной работе изучено современное программное обеспечение [20, 21]. Для сравнительного анализа выбраны компьютерные программы, в которых возможно создание виртуальной лабораторной работы:

LabVIEW – программное обеспечение для системного проектирования в отраслях, где требуется проведение испытаний, измерений и осуществление управления, с его помощью разрабатываются и выполняются программы, созданные на основе графического языка программирования «G». LabVIEW предоставляет графический метод программирования, позволяющий отобразить любую часть выполняемой задачи. Визуализация в большой степени упрощает разработку интерфейсов пользователя. В составе программы имеется большой набор библиотек, ускоряющий разработку. LabVIEW можно использовать на различных операционных системах. Программа незначительно уступает в производительности другим языкам программирования. Для использования объекта, созданного с помощью LabVIEW на компьютере, в котором сама среда не установлена, необходимо установить LabVIEW Run-Time. Стоимость данного программного обеспечения – 400 долларов в год.

Наиболее эффективная область применения LabVIEW – это построение графиков, моделирование измерительных приборов, переключателей, осциллографов.

Adobe Flash, или просто Flash – это мультимедийная платформа компании Adobe для создания web-приложений или мультимедийных презентаций платформу Flash можно использовать при создании виртуальных лабораторных работ. Программное средство позволяет реализовывать анимированные проекты с использованием векторной графики и использованием встроенного языка программирования ActionScript 3.0. ActionScript поддерживает события, условия, циклы, массивы, функции и наследуемые классы. Flash использует векторный формат изображений, что уменьшает размер выходных файлов. Для внесения любого изменений необходимо изменить Flash-файл целиком. Для просмотра Flash-анимации в браузерах необходимо установить на Adobe Flash Player. Adobe Flash Player сейчас используется большинством компьютеров, что в совокупности с хорошей сочетаемостью с HTML позволяет открывать виртуальные лабораторные работы с помощью web-браузеров. Стоимость Adobe Flash составляет 15 456 рублей в год.

С помощью Adobe Flash возможно создать множество разновидностей виртуальных лабораторных работ, обладающих линейностью выполнения. Программная среда удобна для моделирования измерительных приборов, физических опытов.

Algodoo – анимационный графический редактор (симулятор), использующий встроенный язык программирования Thyme. Приложение дает возможность наглядно увидеть работу механической системы, а также, меняя параметры сцены, оптимизировать и анализировать полученные решения. Имеется богатый инструментарий для создания различных систем и механизмов. Поддерживаются слои, каждый объект на двумерном поле имеет свой слой и взаимодействует лишь с ним. Имеется возможность загружать в симулятор любые сторонние рисунки. Для всех объектов сцены устанавливаются различные физические параметры – объем, вес, характер материала, упругость, скорость, длина и сила натяжения пружины, высота и шаг зубцов шестеренок и многое другое. Симулятор интерактивен, в любой момент во время демонстрации можно вносить коррективы. Программа Algodoo имеет встроенный скриптовый язык под названием Thyme, привносящий в симулятор еще большую свободу действий и позволяющий создавать: объекты с уникальными физическими свойствами, явления и эффекты В Thyme доступны массивы, условные операторы, переменные, обработка событий, происходящих в среде моделирования. Минусом является отсутствие поддержки русского языка. Algodoo распространяется на бесплатной основе.

Таблица 1

Критерии выбора технический средств для создания модели виртуальной лабораторной работы

Критерии	LABView	Adobe Flash	Algodoo
Область применения	Измерительные приборы, переключатели, графики	Линейные физические опыты, нетиповые приборы	Физические процессы, ускоренные процессы
Тип языка программирования	Графический	Текстовый	Текстовый
Возможность использования в HTML	–	+	–
Русскоязычная поддержка	+	+	–
Цена	400 долларов (26 320 рублей) в год	15 456 рублей в год	Бесплатно

Algodoo прекрасно подходит для моделирования физических и химических процессов, движения, моделирование процессов с измененной скоростью времени.

Практическая значимость

Конечной целью анализа программных сред программирования виртуальных лабораторных работ является разработанная виртуальная лабораторная работа. Основываясь на анализе из таблицы 1, для программной реализации была выбрана среда Adobe Flash Professional по причине сочетаемости его выходных документов в html-файле. Выходные файлы могут быть использованы на большинстве современных компьютеров, поддерживающих последнюю версию Adobe Flash Player.

С помощью объектов, реализуемых в языке программирования ActionScript 3.0 создана модель лабораторной работы, в которой пользователь имеет возможность соединять два конденсатора необходимым ему способом (рис. 1).

Для проведения измерений в виртуальной лабораторной реализована модель микроамперметра (рис. 2).

В ходе проведения тестирования ВЛР (совместно с обучающимися гр. УТСбп-17-1 ИГиН ТИУ [22]) сформулированы следующие преимущества по сравнению с традиционными лабораторными работами:

1. Отсутствие поломок лабораторного оборудования;

2. Безопасность (т.е. отсутствие высокого напряжения, которое сопровождает лабораторное оборудование);

3. Возможность проводить эксперименты большое количество раз и с более сложными условиями;

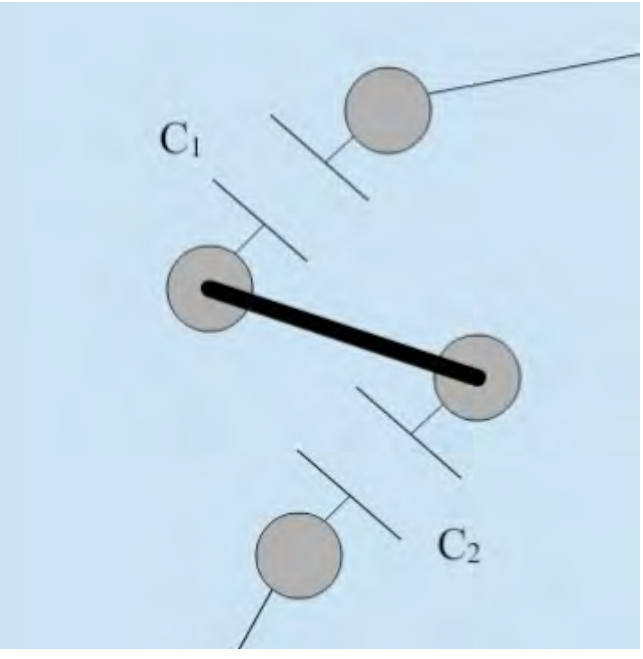


Рис. 1. Последовательно соединенные конденсаторы

Рис. 2. Микроамперметр

4. Освобождение обучающихся от монотонной, трудоёмкой работы, освобождая больше времени мыслительной и творческой деятельности;

5. Возможность выполнения лабораторной работы дистанционно.

При всех преимуществах ВЛР, безусловно, необходимо отметить, что компьютерные технологии не заменяют интеллект человека и изучение базовых основ и навыков при изучении дисциплины. Не зная формул и алгоритма решения задач «вручную», ВЛР становится бесполезным инструментом для обучающегося. В таком случае, ни одна ВЛР не является заменой реальных экспериментов, а служит интерактивным и удобным дополнением.

В ходе выполнения научно-исследовательской работы получены следующие выводы:

- был проведен анализ современных средств разработки виртуальных лабораторных работ;
- разработана работоспособная модель виртуальной лабораторной работы;

Поставленные задачи решены и цель работы достигнута.

Выводы

В результате работы можно сделать вывод о том, что изучены современные методы и технологии создания виртуальных лабораторных работ. Был проведен анализ средств разработки виртуальных лабораторных работ, учтены преимущества и недостатки каждой из рассматриваемых программ, выбрана программа, с помощью которой создана виртуальная лабораторная работа в соответствии с техническим заданием.

Список литературы

1. Мурзин С.В. Виртуальные приборы в лабораторном практикуме [Текст] / С. В. Мурзин // Информатика и прикладная математика: межвузовский сборник научных трудов. – 2010. № 16. – С. 61–63.

2. Савкина Т.А. Виртуальный лабораторный практикум [Текст] / Т.А. Савкина, О.В. Вайда, Е.Г. Зальцман, Е.С. Штаговец // Решетневские чтения. – 2013. № 17. – С. 506–507.

3. Бочкарева Ю.Г. Формирование информационной компетенции бакалавров в виртуальном лабораторном практикуме по физике [Текст] / Ю.Г Бочкарева, Н.И. Чижухина // Современные проблемы науки и образования. – 2012. № 2. – С. 176.

4. Hossain M.A systematic review of living lab literature / M. Hossain, S. Leminen, M. Westerlund // Journal of Cleaner Production. – Volume 2013. – 2019. – PP. 976-988. DOI: 10.1016/j.jclepro.2018.12.257.

5. Baez H. Use of remote laboratories in the teaching of technological degree / H. Baez, O. E. Mojica // IEEE Global Engineering Education Conference, EDUCON. – Volume 2018. – PP. 1957–1961.

6. Hidayati H. The Development of Virtual Laboratory Using ICT for Physics in Senior High School / H. Hidayati, Y. Darvina // IOP Conference Series: Materials Science and Engineering. 2nd International Conference on Mathematics, Science, Education and Engineering, ICOMSET 2017. – Volume 335. – 2017. – 012069.

7. Li C. Innovate engineering education by using virtual laboratory platform based industrial robot / C. Li, L. Fu, L. Wang // Proceedings of the 30th Chinese Control and Decision Conference, CCDC 2018. –2018. – PP 3467–34723.

8. Zhao H. Research on application of novel virtual reality technology in medical teaching / H. Zhao, G. Li, W. Feng // Proceedings - 2018 International Conference on Robots and Intelligent System, ICRIS 2018. – PP. 443–445.

9. Sharma S. Can virtual labs become a new normal? A case study of Millikan's oil drop experiment / S. Sharma, P.K. Ahluwalia // European Journal of Physics. – Volume 39. – 2018. – 065804.

10. Baranov A.V. Students' Project Developments of Wave Optics Virtual Labs / Baranov, A.V. // 14th International Scientific-Technical Conference on Actual Problems of Electronic Instrument Engineering, APEIE 2018 – Proceedings 26. – 2018. – PP. 240-242.

11. Virtual Laboratories in Biotechnology are Significant Educational Informatics Tools / Radhamani, R. [et] // 2018 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2018. – PP. 1547–1551.

12. Donose C. Using the Virtual Online Laboratories for the pre-University / C. Donose, C.-M. Schreiner, S. C. MacOvei // EPE 2018 – Proceedings of the 2018 10th International Conference and Expositions on Electrical And Power Engineering. – 2018. – PP. 609-613;

13. Bistak P. Disturbance Analysis Virtual Laboratory for PID Controllers with Higher Derivative Degrees / P. Bistak // ICETA 2018 - 16th IEEE International Conference on Emerging eLearning Technologies and Applications, Proceedings 10. – 2018. – PP. 69-74. DOI: 10.1109/ICETA.2018.8572065.

14. Estriegana R. Analysis of competence acquisition in a flipped classroom approach / R. Estriegana, J.-A. Medina-Merodio, R. Barchino // Computer Applications in Engineering Education. – Volume 27. – 2019. – PP. 49-64. DOI: 10.1002/cae.22056.

15. Modelling and simulation methodology for unidirectional composite laminates in a Virtual Test Lab framework / O. Falcó [et] // Composite Structures. – Volume 190. – 2018. – Pages 137-159.

16. Characterization of biodiesel using virtual laboratories integrating social networks and web app following a ubiquitous- and blended-learning / P.-M. Manuel // Journal of Cleaner Production. – Volume 215. – 2019. – PP. 399-409. DOI: 10.1016/j.jclepro.2019.01.098.

17. Баймульдин М.К. Актуальность использования виртуальных лабораторных работ для дисциплины «Основы компьютерного моделирования» / М.К. Баймульдин, Э.К. Сейпишева, Б. О. Мухаметжанова, Д. Е. Сагатбекова // Молодой ученый. – 2016. – № 26. – С. 3–7.

18. Lincoln J. Using Python as a toolbox for student-directed virtual labs / J. Lincoln // Physics Teacher. – Volume 57. – 2019. – PP. 60-61. DOI: 10.1119/1.5084937;

19. Foust E.C. Optimizing students' learning experiences in instrumentation and measurement laboratory / E. C. Foust // ASEE Annual Conference and Exposition, Conference Proceedings. – Volume 2018.

20. ГОСТ Р 55751-2013. Информационно-коммуникационные технологии в образовании. Электронные учебно-методические комплексы. Требования и характеристики. – Введ. 2015-01-01. – Москва: Стандартинформ, 2014;

21. ГОСТ Р МЭК 60073-2000 Интерфейс человекомашинный. Маркировка и обозначения органов управления и контрольных устройств. Правила кодирования информации. – Введ. 2002-01-01. – Москва: ИПК Издательство стандартов, 2001.

22. Леонтьев З.Д. Разработка виртуальной лабораторной работы / З.Д. Леонтьев, И.И. Башаров // Кристаллы творчества. – 2018. – Т. 1. – С. 209–211.

УДК 33:001

Д.А. Сичинский, А.В. Мельников

МОДЕЛЬ ОЦЕНКИ IT-СЕРВИСОВ УМНОГО ГОРОДА. ТОЧКИ ВЛИЯНИЯ НА ОБЩЕСТВЕННО-ПОЛИТИЧЕСКУЮ ОЦЕНКУ ДЕЯТЕЛЬНОСТИ ОРГАНОВ ГОСУДАРСТВЕННОЙ ВЛАСТИ

Аннотация

В настоящее время нет количественной модели для оценки удовлетворенности IT-сервисами «Умного города», основанной на объективных данных. Реализация концепции «максимально возможная эффективность для граждан, при оптимальных затратах» невозможна без таких оценок. концепции «умный город». В связи с этой проблемы, в статье описывается подход к построению модели оценки IT-сервисов. Для получения объективной оценки необходимо использовать модели методы, основанные на статистической информации, собираемых данных из социальных сетей, форумов, комментариев к новостным ресурсам. Для обработки таких данных предлагается использовать методы мягкого рандомизированного оценивания характеристик моделей удовлетворенности IT-сервисами. В работе была собрана статистическая информация за последний год по государственным услугам в Ханты-Мансийском автономном округе Югра, выделены ключевые слова для поиска постов и комментариев в социальной сети, предложена модель оценки IT-сервисов.

Ключевые слова: рандомизированное машинное обучение, «Умный город», социально-экономический эффект, анализ эффективности.

Введение

Исследования по «умным городам» вызывают пристальное внимание граждан, частных инвесторов и органов государственной власти. Накопленные знания по машинному обучению и концепции «умный город» на сегодняшний день дают основание для переосмысления разрабатываемой модели оказания услуг и актуальности темы в целом. Предметом исследования по умным городам является с одной стороны, осведомленность граждан о приложениях и решениях, которые считаются «умными», а с другой востребованность разрабатываемых приложений и решений для частичного удовлетворения ежедневных, рутинных или типовых потребностей граждан. Стоит отметить, что разрабатываемые сервисы для «умного города» создаются и развиваются за счет средств налогоплательщиков, которые вправе получать качественные сервисы и услуги, за которые они опосредованно заплатили. На основе анализа результатов пилотного международного исследования авторов Miltiadis D. Lytras и Anna Visvizi можно сделать вывод, что даже, наиболее осведомленные и имеющие навыки эффективного использования услуг «умного города» граждане, испытывают очень серьезную обеспокоенность по поводу полезности, безопасности, доступности и эффективности получаемых услуг [2]. Для того, чтобы исследования «умных городов» оставались полезными и актуальными для всех заинтересованных сторон, необходимо применять подход, основанный на концепции «максимально возможная эффективность для граждан, при оптимальных затратах», которая должна быть реализована при создании IT-сервисов. Цель работы состоит в создании модели оценки удовлетворенности граждан существующими и разрабатываемыми IT-сервисами.

Для решения задач по прогнозированию и оценке применяют различные алгоритмы машинного обучения, которые определяются, как обученные детерминированные модели, с эмпирическими вероятностными оценками их достоверности. [3].

Предварительным этапом, необходимым при выборе модели обучения для оценки IT-сервисов является нахождение данных, называемых data set. В настоящей работе используются цифровые источники данных, получаемые из: социальных сетей, тематических блогов (в том числе комментариев, оставляемых под различными постами в блогах и социальных сетях), сайтов органов государственной власти, архивов федеральной службы государственной статистики, портала государственных услуг Российской Федерации, данных социологических опросов, порталов открытых данных, публикаций в периодической интернет-прессе.

Из полученных data set выделяются параметры, которые необходимы для оценки IT-сервисов.

Учитывая, что в результате поиска косвенных оценок качества и полезности IT-сервисов, как правило, сформированные data sets, погружены в некоторую неопределенную среду. Это значит, что данные получены с ошибками, пропусками, сомнительной достоверностью. Необходимо подобрать соответствующий способ имитации неопределенности [3]. В данной работе выбрана концепция, которая основывается на рандромизированном представлении параметров модели и ошибок измерений данных. Модель минимизирующая эмпирический риск – модель рандомизированного машинного обучения, предложенная российским ученым Попковым Ю.С.

В рассматриваемой модели оценки являются условными. Отличие от существующих процедур ма-

шинного обучения в том, что в рамках процедур рандомизированного машинного обучения строятся оптимальные оценки не параметров, а функций плотности распределения вероятностей (ПРВ) случайных параметров и «наихудшие» плотности распределения вероятностей случайных ошибок в данных. [3]

Описание данных и модели оценки IT-сервисов

Для оценки IT-сервисов будут использоваться два типа данных получаемые из data set:

1) Текстовые данные

2) Данные государственной статистики.

При использовании текстовых данных, необходимо учитывать частотные характеристики текста, например, вектор количеств вхождений ключевых слов в документ. [3]

При обращении к данным государственной статистики, следует принимать к вниманию то, что данные получаемые при ведении статистики имеют уже заложенные ранее ошибки, так как используют единые методы расчетов для разных регионов и типов анализируемых данных.

Для получения данных из социальных сетей, блогов и иных социальных медиа, после анализа, были выбраны 2 способа:

1) Brand24.com – англоязычный, коммерческий сервис для мониторинга англоязычных социальных медиа, позволяющий составлять график масштабности обсуждений для анализа позитивного отношения к услугам, о негативных отзывах об услугах с популярных блогов.

2) API социальных сетей

Для сбора первичного data set о государственных услугах из социальных сетей, была написана программа по выгрузке записей и комментариев из группы «Госуслуги Югры» в социальной сети вконтакте. По полученным данным была отфильтрована информация по следующим критерием:

- Количество просмотров >1000
- Количество лайков >10

Комментарии относящиеся к записям удовлетворяющими данные условия, также вошли в сформированную выборку.

На основе полученных данных было сформировано семантическое ядро ключевых слов и словосочетаний, данный список был отфильтрован по значению более 5 вхождений в наборе данных.

Данный набор ключевых слов и словосочетаний мы сравнили методом морфологическое вхождения, одно или несколько слов из ключевой фразы изменены по форме с помощью склонения или спряжения, регистр не важен, последовательность слов важна, с услугами и подуслугами портала «Госуслуги». Отсортировав по количество встречаемых значений от 5 вхождений, сформировалась выборка ключевых слов и фраз, состоящая из 30 слов.

По полученным данным о государственных услугах с портала «Госуслуги», удалось сформировать 12 самых востребованных услуг и 3 онлайн услуги.

Следующий шаг в работе состоит из сбора date set, по ключевым словам, эмоциональной оценке постов и комментариев относящимся к услугам из таблиц 2, 3.

Разрабатываемые сервисы для концепции «Умный город», оцениваются по предложенной модели

$$Y_k = M(X_1^k, X_2^k, \dots, X_n^k) \quad (1)$$

Для оценки сервисов, используются параметры, полученные из накопленных data sets за определенное количество времени. В качестве примера, для оценки сервиса «налоговая задолженность» используются параметры, отображенные на рисунке 1.

Таблица 1

Ключевые слова и фразы относящиеся к государственным услугам в ХМАО

№	Ключевое слово	№	Ключевое слово	№	Ключевое слово
1	запись на прием	11	дневник	21	работа
2	замена	12	успеваемость	22	уход за ребенком
3	кредит	13	семья	23	государственные услуги
4	налог	14	заявления о регистрации	24	собственность
5	электронной вид	15	электронное заявление	25	оформить пособие
6	очередь в детский сад	16	район	26	социальный
7	доступный	17	денежный	27	закон
8	участник	18	профессиональный	28	счет
9	паспорт	19	справка	29	оружие
10	услуга	20	автомобиль	30	образование

Ограничение data sets основывается на географическом положении, для получения корректных результатов. Выделяемые параметры для оценки каждого сервиса, исходя из экспертной оценки ранжируются по приоритетам.

Пропущенные данные в шкалах data set так же важны для создания общей картины состояния удовлетворенности граждан. В связи с чем удаление данных с «пустыми» или «пропущенными» значениями не целесообразно.

В рамках данного метода при расчете оценки удовлетворенности граждан ИТ сервисами есть два ключевых варианта замены «пустых» данных. Первый вариант -замена «пропущенных» данных на максимальный. Это связано с тем, что неудовлетворенный услугами гражданин в большинстве случаев осведомит свое отношение к ИТ сервису и взаимосвязи с ним. А вот удовлетворенные полученной услугой граждане, чаще всего оставят без внимания успешное взаимодействие, тем самым считая, что сервис работает правильно и в соответствии с потребностями.

Другой вариант, который был выбран для замены «пустых» значений в данном исследовании, модели оценки сервисов – замена пустого значения на зна-

Таблица 2

Востребованных услуг с портала «Госуслуги»

№	Название услуги	Количество обращений за 2018 год
1	Информирование застрахованных лиц о состоянии их индивидуальных лицевых счетов в системе обязательного пенсионного страхования согласно федеральным законам «Об индивидуальном (персонифицированном) учете в системе обязательного пенсионного страхования» и «Об инвестировании средств для финансирования накопительной пенсии в Российской Федерации»	415 274
2	Предоставление социальной поддержки отдельным категориям граждан в соответствии с принятыми нормативными правовыми актами Ханты-Мансийского автономного округа - Югры	342 812
3	Регистрационный учет граждан Российской Федерации по месту пребывания и по месту жительства в пределах Российской Федерации	165 507
4	Проведение экзаменов на право управления транспортными средствами и выдача водительских удостоверений	110 248
5	Регистрация автомототранспортных средств и прицепов к ним	99 323
6	Оформление и выдача паспортов гражданина Российской Федерации, удостоверяющих личность гражданина Российской Федерации за пределами территории Российской Федерации, содержащих электронный носитель информации	92 425
7	Выдача справок о наличии (отсутствии) судимости и (или) факта уголовного преследования либо о прекращении уголовного преследования	75 726
8	Выбор избирательного участка на выборах Президента Российской Федерации	33 500
9	Зачисление в образовательное учреждение	32 989
10	Выдача гражданину Российской Федерации разрешения на хранение и ношение охотничьего огнестрельного длинноствольного оружия, спортивного огнестрельного длинноствольного гладкоствольного оружия, охотничьего пневматического оружия или огнестрельного оружия ограниченного поражения и патронов к нему	25 193
12	Оформление и выдача паспортов гражданина Российской Федерации, удостоверяющих личность гражданина Российской Федерации за пределами территории Российской Федерации	24 554

Таблица 3

Востребованные онлайн услуги с портала «Госуслуги»

№	Тип услуги	Название онлайн услуги	Количество обращений за 2018 год
1	Региональная	Электронный дневник	11 377 341
2	Федеральная	Запись на прием к врачу	387567
3	Региональная	Очередь в детский садик	218 754

Рис. 1. Пример схемы взаимосвязанных компонентов модели по оценке сервисов

чение среднее по моде. Из ряда других значений на данный параметр выбирается наиболее популярный ответ граждан.

Многие IT-сервисы со временем подвергаются модернизации и расширением дополнительными подуслугами. Например, услуга на портале «Госуслуги» под название «Регистрация расторжения брака» состоит из 3-х различных подуслуг: «Регистрация расторжения брака на основании решения суда два заявителя», «Регистрация расторжения брака на основании решения суда один заявитель», «Регистрация расторжения брака по обоюдному согласию, заполняется двумя заявителями». В таких случаях косвенные оценки, относящиеся к подуслуги сложно отследить в каких-либо источниках, собирается только статистическая информация. В таких случаях предлагается мерить оценку удовлетворительности только к родительской услуге, но при добавлении дополнительной подуслуги или модернизации какой-либо из существующих, необходимо отследить общее изменение удовлетворенности услуги и внести этот показатель, как критерий модернизации, выраженный в разнице от предыдущей и текущей удовлетворенностью.

Выводы

Работа была направлена на создание модели оценки IT-сервисов, которые способны обеспечить устойчивый рост качества жизни граждан и благоприятных условий ведения предпринимательской и иной деятельности за счет использования цифровых технологий.

В работе представлена модель оценки IT-сервисов «Умного города», которая направлена на создание и развитие необходимых сервисов для «умного города». В основу модели заложена концепция «максимально возможная эффективность для граждан, при оптимальных затратах», которая должна быть реализована при создании и поддержании IT-сервисов для граждан.

В работе выделены средства поиска и получения данных из масс медия ресурсов, социальных сетей и государственных порталов услуг, такого как «Госуслуги». Сформирован список ключевых слов для сбора data set по государственным услугам. Произведен статистический анализ обращений к IT услугам на портале «Госуслуги» в ХМАО.

Ограничение дата сетов основано на географическом положении, для получения корректных результатов. Для решения проблемы с пропущенными данными в шкалах data set, принято решения о замене «пустых» данных на значение среднее по моде.

На основе предложенной модели будут проведены оценки для формирования обучающего набора, по результатам которых будет проведены эксперименты по оценки IT-сервисов при помощи рандомизированного машинного обучения.

Список литературы

1. Moscow «Smart City – 2030» https://www.mos.ru/2030/

2. Who Uses Smart City Services and What to Make of It: Toward Interdisciplinary Smart Cities Research / Miltiadis D. Lytras, Anna Visvizi.

3. Рандомизированное машинное обучение Popkov Y.S., Dubnov Y.A., Popkov A.Y. https://esstu.ru/library/free/Konf/Maths_2017/69_Попков_Дубнов_Попков.pdf.

4. AI-Powered Social Media Analytics For Businesses https://lucidya.com/en/

5. Hong Kong Smart City Blueprint https://www.smartcity.gov.hk/doc/HongKongSmartCityBlueprint(EN).pdf.

6. Оценка и измерение ИТ-процессов https://www.osp.ru/itsm/2012/05/13015272.html

7. Расчет стоимости ИТ-сервисов: определение, моделирование и расчет стоимости ИТ-сервисов http://www.inframanager.ru/library/about-methodology/raschet-stoimosti-it/

8. Оценка качества и уровня сервиса Качество сервиса https://studfiles.net/preview/3638652/page:48/

9. Системный анализ и исследование операций. Оценивание характеристик рандомизированных статических моделей данных (энтропийно-робастный подход) Ю.С. Попков, А.Ю. Попков, Ю.Н. Лысак http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=at&paperid=6164&option_lang=rus

10. A Comparative Evaluation of Log-Based Process Performance Analysis Techniques Authors and affiliations: Fredrik MilaniFabrizio M. Magg https://link.springer.com/chapter/10.1007/978-3-319-93931-5_27

11. Lawrence, T.M., Watson, R.T., Boudreau, M.-C., Mohammadpour, J.: Data flow requirements for integrating smart buildings and a smart grid through model predictive control. Procedia Eng. 180, 1402–1412 (2017).

12. Siano P.: Demand response and smart grids—a survey. Renew. Sustain. Energy Rev. 30, 461–478 (2014).

13. Hu Q., Li, F.: Hardware design of smart home energy management system with dynamic price response. IEEE Trans. Smart Grid. 4, 1878–1887 (2013).

14. О пользе обратной связи http://www.inframanager.ru/library/about-methodology/about-feedback/

15. Jagadeesh Chandra Bose, R.P., Gupta, A., Chander, D., Ramanath, A., Dasgupta, K.: Opportunities for process improvement: a cross-clientele analysis of event data using process mining. In: Barros A., Grigori D., Narendra N.C., Dam H.K. (eds.) ICSOC 2015. LNCS, vol. 9435, pp. 444–460. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-48616-0_31

16. Hompes, B.F.A., Buijs, J.C.A.M., van der Aalst, W.M.P.: A generic framework for context-aware process performance analysis. In: Debruyne, C., et al. (eds.) On the Move to Meaningful Internet Systems: OTM 2016 Conferences. OTM 2016. LNCS, vol. 10033, pp. 300–317. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-48472-3_17.

17. de Leoni, M., van der Aalst, W.M.P., Dees, M.: A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs. Inf. Syst. 56, 235–257 (2016).

18. Premchaiswadi, W., Porouhan, P.: Process modeling and bottleneck mining in online peer-review systems. SpringerPlus 4(1), 441 (2015).

19. Технологии умных городов: что влияет на выбор горожан? https://www.mckinsey.com/~/media/McKinsey/Industries/Public%20Sector/Our%20Insights/Smart%20city%20solutions%20What%20drives%20citizen%20adoption%20around%20the%20globe/smartcitizenbook-rus.ashx.

20. Аджубей Р.В. Оценка эффективности предоставления государственных и муниципальных услуг населению // Молодой ученый. – 2017. – № 1. – С. 135-140. – URL https://moluch.ru/archive/135/37885/ (дата обращения: 21.12.2018).

21. Основные показатели эффективности государственных услуг 2012. М.С. Таболина. Московский государственный институт электроники и математики (технический университет) http://ecsocman.hse.ru/data/2012/11/09/1251375651/24.pdf.