



XXI



**НАЦИОНАЛЬНАЯ
КОНФЕРЕНЦИЯ
ПО ИСКУССТВЕННОМУ
ИНТЕЛЛЕКТУ
С МЕЖДУНАРОДНЫМ
УЧАСТИЕМ**



КИИ-2023



ТРУДЫ КОНФЕРЕНЦИИ

Том 1

16-20 октября 2023 г.
Смоленск

XXI

Российская ассоциация
искусственного интеллекта

Федеральный
исследовательский центр
«Информатика
и управление» РАН

Филиал Национального
исследовательского
университета «МЭИ»
в г. Смоленске

НАЦИОНАЛЬНАЯ
КОНФЕРЕНЦИЯ
ПО ИСКУССТВЕННОМУ
ИНТЕЛЛЕКТУ
С МЕЖДУНАРОДНЫМ
УЧАСТИЕМ

КИИ-2023

ТРУДЫ КОНФЕРЕНЦИИ

Том 1

16-20 октября 2023 г.
Смоленск,
филиал НИУ «МЭИ» в г. Смоленске

Смоленск
2023

УДК 004.8+004.89+004.82+004.032.26(045)+004.9

ББК 32.813

Д 22

Организаторы конференции:

Российская ассоциация искусственного интеллекта

Федеральный исследовательский центр «Информатика и управление» РАН

Филиал Национального исследовательского университета «МЭИ»

в г. Смоленске

При поддержке Администрации Смоленской области

Д 22 Двадцать первая Национальная конференция по искусственному интеллекту с международным участием, КИИ-2023 (Смоленск, 16-20 октября 2023 г.). Труды конференции. В 2-х томах. Т.1. – Смоленск: Принт-Экспресс, 2023. – 410 с.

ISBN 978-5-91812-231-0

Двадцать первая Национальная конференция по искусственному интеллекту с международным участием КИИ-2023 продолжает традицию советских (российских) конференций, организуемых Российской ассоциацией искусственного интеллекта.

В первом томе трудов публикуются пленарные доклады и доклады участников конференции, представленные на следующих секциях:

Секция 1 «Инженерия знаний»,

Секция 2 «Интеллектуальный анализ данных»,

Секция 3 «Интеллектуальные агенты, роботы, интеллектуальное управление, компьютерное зрение»,

Секция 4 «Машинное обучение, нейросетевые методы».

ББК 32.813

Рецензенты: академик РАН, ИПУ РАН *С.Н. Васильев*,
д.т.н., доц., ФИЦ ИУ РАН *В.Н. Захаров*

ISBN 978-5-91812-231-0

© Авторы, 2023

© Российская ассоциация искусственного интеллекта, 2023

ПРОГРАММНЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

Сопредседатели Программного комитета

Грибова В.В., член-корр. РАН, ИАПУ ДВО РАН, Владивосток
Кобринский Б.А., д.м.н., проф., ФИЦ ИУ РАН, Москва

Заместители председателя Программного комитета

Забейхайло М.И., д.ф.-м.н., проф., ФИЦ ИУ РАН, Москва
Тельнов Ю.Ф., д.э.н., проф., РЭУ им. Г.В. Плеханова, Москва

Ответственный секретарь Программного комитета

Колоденкова А.Е., д.т.н., доц., СамГТУ, Самара

Члены программного комитета

Аверкин А.Н., к.ф.-м.н., доц., ФИЦ ИУ РАН, Москва
Афанасьева Т.В., д.т.н., РЭУ им. Г.В. Плеханова, Москва
Болодурина И.П., д.т.н., проф, Оренбургский ГУ, Оренбург
Боргест Н.М., к.т.н., доц., Самарский НИУ им. акад. С.П. Королева
Борисов В.В., д.т.н., проф., филиал НИУ «МЭИ», Смоленск
Васильев С.Н., акад. РАН, ИПУ РАН, Москва
Визильтер Ю.В., д.ф.-м.н., проф., ГосНИИ АС, Москва
Гаврилова Т.А., д.т.н., проф., СПбГУ, Санкт-Петербург
Городецкий В.И., д.т.н., проф., АО «Эврика», Санкт-Петербург
Дли М.И., д.т.н., проф. филиал НИУ «МЭИ», Смоленск
Еремеев А.П., д.т.н., проф., НИУ «МЭИ», Москва
Желтов С.Ю., акад. РАН, ГосНИИАС, Москва
Загорулько Ю.А., к.т.н., ИСИ СО РАН, Новосибирск
Каляев И.А., акад. РАН, ЮФУ, Ростов-на-Дону
Ковалев С.М., д.т.н., проф., РГУПС, Ростов-на-Дону
Колесников А.В., д.т.н., проф., БФУ им. И. Канта, Калининград
Козенко И.В., д.т.н., проф., СПИИРАН, Санкт-Петербург
Кузнецов С.О., д.ф.-м.н., проф., НИУ ВШЭ, Москва
Курейчик В.В., д.т.н., проф., ЮФУ, Таганрог
Лебедев О.Б., д.т.н., доц., ИКТИБ ЮФУ, Таганрог
Лукашевич Н.В., д.т.н., проф., МГУ им. М.В. Ломоносова, Москва
Мисник А.Е., к.т.н., доц., БРУ, Республика Беларусь, Могилев
Михеенкова М.А., д.т.н., проф., ФИЦ ИУ РАН, Москва
Мошкин В.С., к.т.н., доц., УлГТУ, Ульяновск
Пальчунов Д.Е., д.ф.-м.н., доц., ИМ СО РАН, Новосибирск
Палюх Б.В., д.т.н., проф., ТвГТУ, Тверь
Панов А.И., к.ф.-м.н., доц., ФИЦ ИУ РАН, Москва
Подвесовский А.Г., к.т.н., доц., БГТУ, Брянск
Редько В.Г., д.ф.-м.н., проф., НИИСИ РАН, Москва
Ройзензон Г.В., к.т.н., доц., ФИЦ ИУ РАН, Москва

Рыбина Г.В., д.т.н., проф., НИЯУ МИФИ, Москва
Смирнов И.В., к.ф.-м.н., доц., ФИЦ ИУ РАН, Москва
Соколов И.А., акад. РАН, ФИЦ ИУ РАН, Москва
Стефанюк В.Л., д.т.н., проф., ИППИ РАН, Москва
Сулейманов Д.Ш., акад. АН РТ, ИПС АН РТ, Казань
Федунов Б.Е., д.т.н., проф., ГосНИИ АС, Москва
Финн В.К., д.т.н., проф., ФИЦ ИУ РАН, Москва
Хачумов В.М., д.т.н., проф., ФИЦ ИУ РАН, Москва
Яковлев К.С., к.ф.-м.н., ФИЦ ИУ РАН, Москва
Ярушкина Н.Г., д.т.н., проф., УЛГТУ, Ульяновск

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

Председатель Организационного комитета

Борисов В.В., д.т.н., проф., филиал НИУ «МЭИ», Смоленск

Заместитель председателя Организационного комитета

Федулов А.С., д.т.н., проф., филиал НИУ «МЭИ», Смоленск

Члены организационного комитета

Гладков Л.А., к.т.н., доц., ИКТИБ ЮФУ, Таганрог

Кириллова Е.А., к.э.н., доц., филиал НИУ «МЭИ», Смоленск

Королева М.Н., к.т.н., доц., МГТУ им. Н.Э. Баумана, Москва

Луферов В.С., к.т.н., ООО «Сбер бизнес софт», Смоленск

Рысина Е.И., к.т.н., филиал НИУ «МЭИ», Смоленск

Синявский Ю.В., к.т.н., доц., филиал НИУ «МЭИ», Смоленск

Солопов Р.В., к.т.н., доц., филиал НИУ «МЭИ», Смоленск

Тютюнник А.А., к.э.н., доц., филиал НИУ «МЭИ», Смоленск

ИНФОРМАЦИОННАЯ ПОДДЕРЖКА КОНФЕРЕНЦИИ

ООО «Лаборатория информационных технологий», Смоленск

ПРЕДИСЛОВИЕ

Двадцать первая Национальная конференция по искусственному интеллекту с международным участием КИИ-2023 продолжает традицию советских (российских) конференций, организуемых Российской ассоциацией искусственного интеллекта (РАИИ).

Федеральный проект «Искусственный интеллект» направлен на решение междисциплинарных проблем, сформулированных в Национальной стратегии развития искусственного интеллекта, включая как исследования в области фундаментальной и прикладной науки, так и в образовательной сфере. Эти проблемы постоянно находятся в фокусе внимания членов РАИИ и нашли отражение в докладах конференции, организаторами которой являются Российская ассоциация искусственного интеллекта, Федеральный исследовательский центр «Информатика и управление» РАН, филиал Национального исследовательского университета «МЭИ» в г. Смоленске.

Тематика конференции охватывает следующие основные направления искусственного интеллекта: инженерия знаний; интеллектуальный анализ данных; интеллектуальные агенты, роботы, интеллектуальное управление; компьютерное зрение; машинное обучение, нейросетевые методы; нечеткие модели, мягкие измерения и вычисления; фреймворки для создания интеллектуальных систем; технологии проектирования и разработки интеллектуальных систем, прикладные интеллектуальные системы, включая биомедицинские системы.

Представлены пленарные доклады видных ученых и специалистов в области искусственного интеллекта. Из поданных на конференцию 90 секционных докладов Программным комитетом после рецензирования отобраны 64 доклада.

В конференции принимают участие более 160 ученых, исследователей и специалистов из академических институтов, университетов, отраслевых институтов и организаций Москвы, Санкт-Петербурга, Белгорода, Брянска, Владивостока, Владимира, Волгограда, Донецка, Иркутска, Казани, Калининграда, Красноярска, Новосибирска, Ростова-на-Дону, Самары, Смоленска, Таганрога, Твери, Ульяновска, Ханты-Мансийска, Челябинска, Могилёва (Республика Беларусь).

Программный комитет КИИ-2023

ПЛЕНАРНЫЕ ДОКЛАДЫ

УДК 004.82

ОНТОЛОГИЯ ПРОЕКТИРОВАНИЯ: ГЕНЕЗИС И РАЗВИТИЕ

Н.М. Боргест (*borgest@yandex.ru*)

Самарский национальный исследовательский университет
им. академика С.П. Королева, Самара

В докладе рассмотрены вопросы генезиса и развития онтологии проектирования, как научного направления, включающего в себя: исследование понятийного аппарата предметной области и разработку на его основе тезауруса, анализ принципов формирования критериев и моделей проектируемого объекта, выбор методов и сценариев проектирования, сбор и обработку информации об объекте. В онтологии проектирования исследуются и формализуются знания субъектов проектирования о процессе проектирования новых или модернизируемых артефактов.

Ключевые слова: онтология проектирования, предметная область, информационная система.

*«Онтологизация должна всегда
предшествовать цифровизации»*

В век информатизации и цифровизации многих видов деятельности наблюдается отставание в семантическом описании предметных областей (ПрО) и процессов в них. Для автоматизации процессов требуется передать соответствующие знания в информационные системы (ИС), интероперабельность которых трудно представить без построения онтологий ПрО, без применения семантических моделей и технологий.

Процесс формализации и компьютерного представления знаний долгие годы шёл без использования термина «онтология». Но не термин решает дело, а осознание необходимости зафиксировать видение ПрО, которая является предметом автоматизации. Необходимость явного выделения и создания этого особого рода компьютерных ресурсов есть эволюционный шаг в развитии информационных технологий.

Проектные практики в различных ПрО имеют разные степени сложности, проработанности, формализованности процессов и участия человека в решении проектных задач, а наличие разных подходов и терминологий по схожим процедурам, терминологических и понятийных разрывов затрудняет трансляцию опыта передовых отраслей и проектных практик в другие ПрО. Необходимо обобщить накопившийся проектный опыт лучших практик, разработать и предложить теоретические основы онтологии проектной деятельности, онтологии проектирования.

Онтологии ПрО, как модели этой ПрО, отражают содержание ПрО и происходящих в ней процессов. Согласованные, однозначно понимаемые пользователем и ИС сущности, их атрибуты и отношения определяют успех формализации, наряду с точным воспроизведением тщательно обоснованных критериев процессов в ПрО.

Для решения задач автоматизации в различных видах деятельности принципиальным является формализация процессов, объектов, критериев и методов принятия решений. И здесь онтологии ПрО стоят в начале создания ИС, обеспечивая их интероперабельность, подбор и применение систем поддержки принятия решений (СППР).

Основой для разработки онтологии проектирования явились накопленные многочисленных практик формализации знаний в различных ПрО, отбор лучших практик и стремление к дальнейшей автоматизации процессов. Основоположителем онтологии как учения о сущем, о бытии, как концептуальной модели мира признан философ Аристотель, в то время как первым онтологом в проектировании можно считать архитектора, механика и учёного-энциклопедиста Витрувия. Отдавая должное другим учёным, следует высоко оценить вклад Витрувия в выработку проектных приёмов и процедур в конкретных ПрО, в формирование основополагающих системных принципов архитектуры, включая изучение материалов, конструкций, среды обитания и пр.

С появлением компьютеров активно разрабатывались методы, технологии и программные реализации выбора и обоснования решений, а в дальнейшем и СППР. Онтологию, как способ компьютерного описания мира, использовали Дж. Мили (1967 г.) и Д. Пауэрс (1983 г.), опираясь на работы философов. Новый импульс этому термину дал Т. Грубер, который ввёл «онтологию» в научный оборот в информатике (1993 г.). Этот

термин приобрёл новое наполнение в информационной среде, образованной компьютерами, программами, словарями, базами данных (БД) и базами знаний (БЗ). Онтологии и их языки стали основой при разработке ИС. Работы учёных в области проектирования, информационных технологий, оптимизации, автоматизации в различных ПрО внесли неоценимый вклад в подготовку исходных предпосылок для выработки теоретических основ нового научного направления «онтология проектирования» (рис. 1).

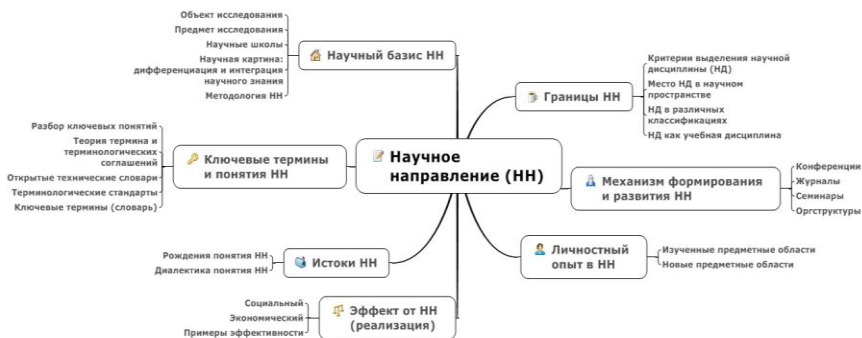


Рис. 1. Научное направление

Метафора: «Онтология проектирования – это взгляд сверху, как проекция взглядов снизу» в обобщённом виде иллюстрирует схему формирования этого научного направления на рис. 2.

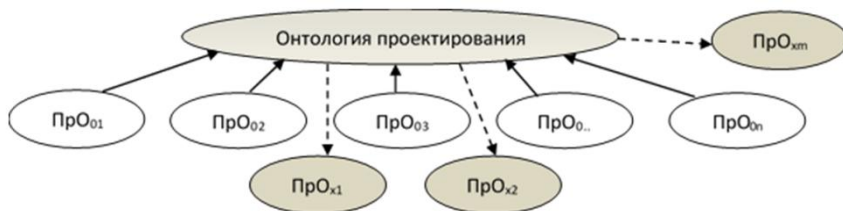


Рис. 2. Схема формирования НН «Онтология проектирования» на основе проектного опыта в передовых $PrO_{01}, \dots, PrO_{0n}$ и «трансляция» его в другие $PrO_{x1}, \dots, PrO_{xm}$

В онтологии проектирования рассматриваются и исследуются вопросы формализованного описания знаний субъектов проектирования о процессе проектирования артефактов, знания об объекте проектирования и близких к нему по свойствам артефактов, а также тезаурус ПрО. При этом методология проектирования включает в себя: исследование понятийного аппарата субъ-

екта проектирования и разработки на его основе тезауруса, анализ критериев и моделей объекта, методов и сценариев, сбор и обработку информации об объекте как системе и составляющих её элементах (рис. 3).



Рис. 3. Методология проектирования объекта

Ключевым фактором проектной деятельности является формализованная потребность, а ключевым принципом – учёт всех имеющиеся ресурсов (временных, материальных, финансовых и др.). Формирование в будущем «единого» языка проектанта на базе разрабатываемого тезауруса обеспечит и облегчит взаимопроникновение полученных результатов созидательной деятельности в различные ПрО.

Проблема границ научного знания напрямую связана с проблемой классификации науки, обусловленной отсутствием целостной онтологии науки, её понятийного аппарата, модели её развития. Разработка согласованного тезауруса и онтологии науки позволит создать «правовой механизм» актуализации номенклатуры научных специальностей.

ПрО состоят из сущностей, их связей и процессов, и именно процессы определяют границы ПрО, а сущности могут быть акторами в разных ПрО. В исследуемой ПрО важны роли (модели поведения), которые выполняют сущности в этой ПрО.

ПрО состоят из сущностей, их связей и процессов, и именно процессы определяют границы ПрО, а сущности могут быть акторами в разных ПрО. В исследуемой ПрО важны роли (модели поведения), которые выполняют сущности в этой ПрО.

Инфраструктура знаний научного направления включает в себя механизм формирования и развития, верификации и аккумуляции знаний, который, в свою очередь, содержит атрибуты добычи, фиксации, апробации и трансляции знаний.

Для систем, использующих математические модели объектов, важным является создание параметризованных моделей на основе онтологии ПрО, позволяющих избежать повторного их построения за счёт изменения значений параметров. Трудоёмкость построения параметризованных моделей «окупается» многократным их использованием и значительным уменьшением доли рутинного труда специалистов высокой квалификации.

Рассмотрим примеры применения научного направления «онтология проектирования» с участием автора.

Робот-проектант. Рассматривается пилотный проект интероперабельной ИС, реализующий методологию (рис. 3) в ПрО «проектирование самолёта».

В качестве ПрО робота-проектанта принят этап предварительного проектирования самолётов различного назначения для некоторого набора признаков схем. Основу робота-проектанта составляет тезаурус ПрО, объединяющий все компоненты интеллектуального помощника: БД прототипов, базу расчётных процедур, параметризованные 3D модели частей самолётов, базу экспертных правил, формирующих семантику действий, человеко-машинный интерфейс.

Формальная запись проектной задачи имеет вид:

$X = \{x_1, x_2, \dots, x_n\}$ – множество параметров;

$Z = \{z_1, z_2, \dots, z_m\}$ – множество ограничений;

$Y = \{y_1, y_2, \dots, y_k\}$ – множество критериев.

$M_{opt} = \{X_{opt}\}$ – результат решения проектной задачи.

Для поиска рационального проекта необходимо на основе данных технического задания (ТЗ), известных ограничений, назначенных критериев и их приоритетов (могут быть заложены в БЗ в виде экспертных знаний на основе прецедентов) обратиться к ИС, которая содержит:

$Pr = \{M_1, M_2, \dots, M_p\}$ – множество прецедентов (БД);

$R = \{m_1, m_2, \dots, m_r\}$ – множество расчётных (математических) моделей;

$K = \{k_1, k_2, \dots, k_r\}$ – множество экспертных правил (включая ассоциативные правила, импликации на признаках объектов в рассматриваемой ПрО и БЗ о сценариях) и связывающий их тезаурус ПрО.

Информационной основой для синтеза являются БД и данные, полученные в результате выполнения логических и расчётных модулей. Заполнение «матрицы» проекта начинается с занесения параметров из ТЗ. Незаполненные ячейки «матрицы», являясь «агентами» соответствующих параметров, в автоматическом режиме инициируют работу надлежащих модулей. Как только в «матрице» появляется достаточно входных данных для некоторого модуля, он начинает выполняться.

Выполнение независимых модулей может идти параллельно. Расчётные модули, вычислительные или сценарные правила или информационные запросы могут рассматриваться как программные агенты, имеющие свои потребности (входные данные) и возможности в виде полученных выходных данных (потенциальные ресурсы). То есть в ИС реализуется известная *сеть потребностей и возможностей*, которая реализуется средствами мультиагентной технологии.

На рис. 4 представлены различной сложности параметризованные 3D модели самолётов, которые показывают развитие робота-проектанта. На рис. 5 показаны фрагменты компоновки и варианты загрузки самолёта, полученные автоматически на основе заложенных правил. На рис. 6 показан выставочный образец робота-проектанта в научном корпусе Самарского университета.

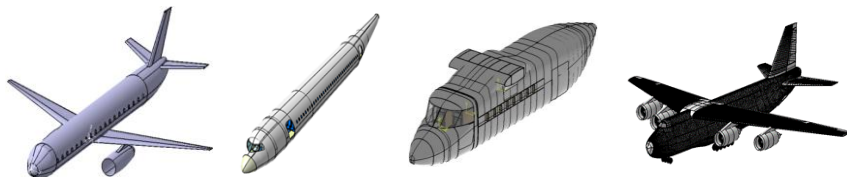


Рис. 4. Развитие параметризованных 3D моделей самолётов (2011-2021 гг.), используемых в роботе-проектанте

Онтология машиностроительного предприятия. Рассматривается онтология планирования производственных операций и процедур на предприятии на основе мультиагентного подхода. В ИС производственного планирования базовая онтология машиностроения использована в качестве прикладного инструмента и обеспечила разработчиков ИС информацией о предприятии и протекающих в нём бизнес-процессах.

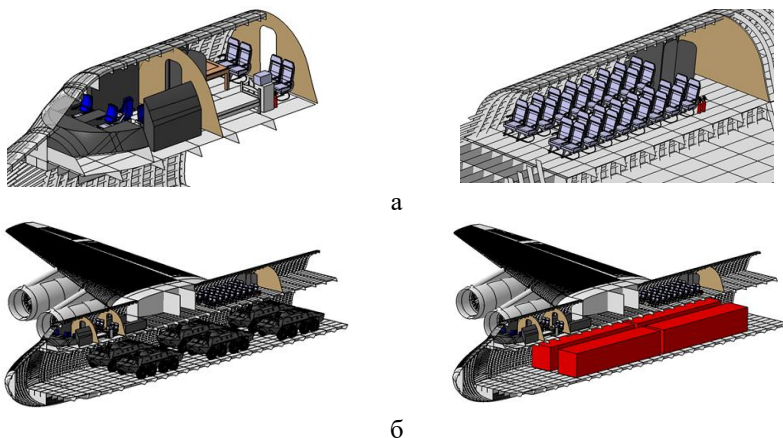


Рис. 5. Фрагменты компоновки (а) и варианты загрузки (б) транспортного самолёта, полученные роботом-проектантом



Рис. 6. Выставочный образец робота-проектанта самолёта

Проектирование музея. Рассмотрена концепция современного профильного музея на примере музея авиации и космонавтики имени С.П. Королёва Самарского университета.

Точное земледелие. Рассмотрены различные предметные онтологии (онтологии растений, онтологии признаков растений, онтологии экспериментальных условий и др.) и онтологии задач, решаемых в области точного земледелия. Представлены оригинальные результаты онтологического моделирования исследуемой области на основе конструктора БЗ. Основу решения задач в точном земледелии составляют онтологии ПрО и мультиагентные технологии.

Интеллектуальный жилой дом для условий Арктики. Предпроектный анализ интеллектуального жилого дома, как и любой проектной задачи, начинается с анализа проектной ситуации.

Нештатная ситуация на космической станции. Основой разрабатываемой СППР для устранения нештатной ситуации (НШС) на космической станции является онтология ПрО, включающей в себя все элементы и системы космической станции. Для построения плана парирования НШС за основу взяты алгоритмы, описанные в бортовой документации.

Онтологическое моделирование ПрО и решаемых в них задач позволяет разработчикам ИС учесть реальную картину процессов в ПрО, согласовать связи между сущностями в ней, обоснованно подойти к составу атрибутов сущностей и в итоге получить адекватный результат работы ИС.

Выводы

Онтология проектирования обобщает опыт проектной деятельности, имеющий долгую историю своего развития. Существенной проблемой её становления является бесконечное разнообразие ПрО созидательной проектной деятельности, большое количество участников этой деятельности, внесших свой вклад в формализацию знаний, в построение онтологий изучаемых и исследуемых областей, в создание артефактов разной природы от простых орудий до сложных человеко-машинных, социально-технических и общественных систем.

Онтология проектирования, как НН в рамках компьютерных наук и инжиниринга, включает в себя: исследование понятийного аппарата ПрО и разработку на его основе тезауруса, анализ принципов формирования критериев и моделей проектируемого объекта, методов и сценариев проектирования, сбор и обработку информации об объекте как системы и составляющих его элементах. В онтологии проектирования исследуются и формализуются знания субъектов проектирования о процессе проектирования новых или модернизаций уже известных артефактов.

УДК 007.52:004.896: 004

ОТ ИНЖЕНЕРИИ ЗНАНИЙ К НАУКЕ О ЗНАНИЯХ*

В.И. Городецкий (*vladim.gorodetsky@gmail.com*)
АО “Эврика”, Санкт-Петербург

Целью работы является краткий анализ современных источников знаний для приложений искусственного интеллекта и анализ перспектив их дальнейшего развития. Показано, что в настоящее время достаточно зрелыми и широко востребованными являются методы, алгоритмы и технологии получения знаний из данных на основе машинного обучения, а также различные варианты получения знаний с использованием технологии цифровых двойников. Однако, что касается экспертных знаний, то в настоящее время этот источник на практике используется редко, хотя он востребован в ряде критически важных, актуальных и часто уникальных классов приложений нового поколения, для которых эксперты являются единственным доступным источником знаний. Для решения проблемы эффективного доступа к экспертным знаниями необходимы интенсивные исследования и разработки в области инженерии знаний, способные поднять ее до уровня Науки о знаниях, которая будет в состоянии решать задачи обработки знаний того же масштаба и сложности, которые в настоящее время под силу Науке о данных по отношению к большим данным. Приведен анализ потенциальных направлений исследований и разработок в гипотетической Науке о знаниях, а также некоторых методов, моделей, алгоритмов и технологий обработки больших объемов сырых фрагментов экспертных знаний для обслуживания интеллектуальных приложений нового поколения.

Ключевые слова: знания, машинное обучение на основе данных, цифровые двойники, экспертные знания, машинное обучение на основе экспертных знаний, непротиворечивость экспертных знаний.

Введение

"Движущей силой возможностей интеллектуальной системы являются знания, которыми обладает эта системы о своей предметной области, а не изощренность процесса рассуждений, который она использует [Augier et

* Работа выполнена в инициативном порядке.

al., 2002]¹. Эти слова принадлежат Э. Фейгенбауму (E. Feigenbaum), одному из основателей искусственного интеллекта (ИИ). Эта мысль, впервые высказанная им еще более 30 лет тому назад, оказалась наиболее точной и краткой характеристикой существа ИИ и его движущей силы и как науки, и как информационной технологии (ИТ). И хотя текущий всплеск интереса ИТ-индустрии к технологиям ИИ обычно объясняется рядом недавних достижений, среди которых фундаментальную роль отводят большим данным, успехам Науки о данных, и технологиям глубокого обучения, глубинной причиной этого успеха является то, что эти достижения сделали доступными для разработчиков приложений ИИ *новый источник знаний* в виде больших данных и надежные методы и средства извлечения этих знаний из данных, а именно методы и алгоритмы машинного обучения, в частности, на основе нейросетевых технологий глубокого обучения.

Появление нового источника знаний привело практически к революционному взрыву возможностей ИИ. Его выдающиеся результаты, представленные инструментами типа ChatGPT и разнообразные приложения в области обработки естественного языка (ЕЯ), построенные с их помощью, например, предметно-настраиваемые голосовые помощники, системы машинного перевода с одного ЕЯ на другой, аннотирование и генерация текстов и т.п. в настоящее время многими воспринимаются как существо и синоним ИИ.

Этой же причиной обусловлен и рост популярности технологии цифровых двойников (ЦД) [Asch, 2022]. Цифровые двойники, с одной стороны, являются мощным источником процедурных знаний, которые в настоящее время составляют основу предсказательных моделей поведения сложных объектов, а с другой – источниками синтетических данных, в которых неявные трудно доступные процедурные знания, скрытые в ЦД, преобразованы в знания, которые представлены выборкой данных, т.е. в более доступную форму благодаря тем же технологиям машинного обучения.

Еще один хорошо известный источник знаний в ИИ – это прикладные предметно-ориентированные знания экспертов, их опыт решения задач, которые предполагается возложить на приложение ИИ. Получение, представление и использование экспертных знаний является предметом научного направления в ИИ, которое называют *инженерией знаний* [Гаврилова и др., 2016].

¹ В оригинале “The driver of the power of intelligent system is the knowledge the systems have about their universe of discourse, not the sophistication of the reasoning process the systems employ”.

В относительно недавнем прошлом (примерно до начала 2000-х годов) экспертные знания играли основную роль в ИИ. Однако примерно полтора – два десятилетия тому назад инженерия знаний остановилась в своем развитии как наука. Она была практически похоронена под натиском технологий больших данных и машинного В итоге приходится фиксировать, что в настоящее время роль экспертных знаний в различных приложениях ИИ минимальна.

Следует отметить, однако, что проблемы и модели использования экспертных знаний продолжают активно развиваться в области когнитивных исследований, см, например, [Горелова, 2021], [Авдеева и др., 2011], [Авдеева, 2018] и др., в которых исследуются проблемы моделирования на основе модели когнитивных карт. Однако в этих работах полагается, что сырые экспертные знания уже приведены к форме, диктуемой приложением, пространство факторов определено, и задачи исследований модели когнитивных карт и разработки приложений – это задачи когнитивного моделирования. Другими словами. в этих работах полагается, что все задачи инженерии экспертных знаний уже успешно преодолены.

Однако в настоящее время появляются актуальные ИТ-приложения нового поколения, которые принципиально отличаются от привычных нам приложений прошлого. Ярким примером таких приложений являются новые приложения в области политических, экономических, социальных и других подобных процессов регионального и межгосударственного уровня, связанных, например, с общественной безопасностью. Эти приложения обычно уникальны, и по ним нет и не может быть исторических данных, а модели этих процессов критически необходимы для управления ими. Это особенно важно в задачах предсказания и оценки напряженности конфликтов разного происхождения, например, на религиозной национальной и территориальной основе, которые обычно имеют более глубокий контекст, например, политический, экономический, социальный, военный и т.д. Яркими примерами таких конфликтов являются многочисленные цветные революции последних двух десятилетий. Можно в качестве примера привести и другие проблемы безопасности общества, обусловленные вредоносным и враждебным поведением того или иного происхождения:

Приложения, примеры которых приведены выше, характеризуются обычно большим масштабом и сложностью процессов, происходящих в них. Соответственно, для создания таких приложений потребуется использовать совместно знания большого числа экспертов разной специализации и квалификации. Сырые экспертные знания в этой ситуации будут иметь огромный объем и содержать разные и противоречивые мнения экспертов об одном и том же, некоторые мнения будут просто ошибоч-

ными. В этом множестве фрагментов сырых знаний неизбежно будет содержаться большое число противоречий и ошибок, которые потребуются локализовать и устранить.

Однако инженерия знаний современного уровня ее развития, включая результаты когнитивных исследований, принципиально не готова к тому, чтобы эффективно обслуживать потребности в знаниях, которые необходимы для приложений нового поколения. В этих приложениях уже невозможно полагаться на ручное преобразование экспертных знаний к форме знаний, пригодных для принятия решений, и это требует принципиального пересмотра отношения к проблеме сбора, очистки, приведения к непротиворечивой форме, представления в машиночитаемом виде и использования больших объемов, возможно, слабо связанных и противоречивых фрагментов экспертных знаний. С другой стороны, как отмечалось выше, знания являются основной движущей силой развития ИИ, и потому полноценное использование экспертных знаний будет означать новый скачок в развитии ИИ и его промышленных приложений.

Целью данной работы является краткий анализ современных источников знаний и прогноз развития методов, моделей, алгоритмов и технологий обработки больших объемов сырых фрагментов экспертных знаний для обслуживания интеллектуальных приложений нового поколения. В дальнейшем материале работы показывается, что для решения названных проблем необходимо поднять современную инженерию знаний до уровня полноценной науки фундаментального характера, которую естественно назвать Наукой о знаниях, и которая будет в состоянии эффективно решать задачи полноценной обработки сырых массивов экспертных знаний и задачи машинного обучения на их основе.

Далее в разделе 1 приводится краткий анализ современных источников знаний и задач, которые необходимо решать для преобразования исходного материала каждого источника в знания на уровне структур (поддержки) принятия решений. В разделе 2 представляется взгляд автора на предмет исследований и основные научные направления Науки о знаниях как наследника инженерии знаний и новые возможности, которые следует ожидать от нее. В разделе 3 приводятся некоторые примеры конкретных задач и методов их решения, которые следует отнести к Науке о знаниях и которые требуют первоочередного внимания исследователей. В заключении резюмируются основные результаты работы.

1. Источники знаний в ИИ

В ИИ в настоящее время используется три типа источников знаний:

а) Наборы данных об объекте, явлении или процессе, накопленные на основе предыдущего опыта работы с ними, например:

- сенсорные данные или значения атрибутов объектов, явлений, процессов или иных сущностей, зафиксированные в процессе их реального функционирования тем или иным способом;
- данные, накопленные в ходе научных экспериментов или наблюдений;
- большие и разнообразные массивы данных в Интернете, в частности, тексты на естественном языке, данные социальных сетей, СМИ, новостных каналов, данные социальных сетей, СМИ и т.д.
- исторические и статистические данные, зафиксированные на длительном временном интервале в официальных документах, базах данных, и т.д.

Современная эпоха Больших Данных сделала этот источник знаний доминирующим и важнейшим. Рис. 1 (левая колонка) демонстрирует основные этапы технологии получения знаний из данных, Для этих технологий фундаментальной основой служат методология, методы и алгоритмы Науки о данных, в частности различные алгоритмы и технологии извлечения знаний из данных на основе машинного обучения, среди которых наибольшей популярностью пользуются нейросетевые технологии глубокого обучения.

б) *Математические модели (цифровые двойники, ЦД) объектов, явлений или процессов*, опирающиеся на предметно-независимые экспертные знания, которые содержатся в книгах, научных статьях, ранее разработанных математических моделях и компьютерных программах.

в) *Прикладные (предметно-зависимые) знания экспертов*, которыми они обладают благодаря своему опыту работы в той предметной области, к которой относится разрабатываемое приложение ИИ.

Как уже отмечалось выше, успехи и популярность ИИ в ИТ-индустрии, отмечаемые в течение последних полутора десятков лет, принято связывать с появлением Больших Данных и технологий глубокого обучения нейросетей. Действительно, именно они сделали возможным создание таких качественно новых инструментов ИИ, как Трансформеры, ChatGPT и «постановку на поток» создание с их помощью предметно-ориентированных голосовых помощников и других классов приложений ИИ, связанных с обработкой естественного языка. Благодаря этим технологиям игровые системы ИИ превзошли в очных соревнованиях мировых лидеров в шахматах и Го. Именно Большие данные и технологии глубокого обучения нейросетей стали источниками многих других достижений в области ИИ-приложений.

Однако глубинной причиной всех этих достижений является то, что они сделали доступными практически неисчерпаемых объемы знаний, которые содержатся в Больших данных, и предложили надежный и относительно простой механизм извлечения этих знаний из данных.

Заметим, что растущая популярность ЦД как источников знаний во многом имеет те же самые причины. Действительно, ЦД как источник знаний используется в двух основных сценариях (рис. 1):

- как источник процедурных знаний, заданных формальной моделью (например, в формализме алгебраических или дифференциальных уравнений, сети машин состояний, графов переходов состояний, моделей типа "если..., то..." и др.) и ее программной реализации. Обратим внимание на то, что весьма популярные генеративные сети в основе своей имеют подстраивающийся цифровой двойник некоторого реального объекта или процесса.
- как генератор синтетических данных, осуществляющий так называемую "датафикацию" (англ. *datafication*) процедурных знаний путем их преобразования в набор данных с экземплярами, удовлетворяющими исходной формальной модели ЦД. Это делается для того, чтобы упростить обнаружение знаний, скрытых в сложной модели ЦД, и потому недоступных разработчику приложения напрямую.

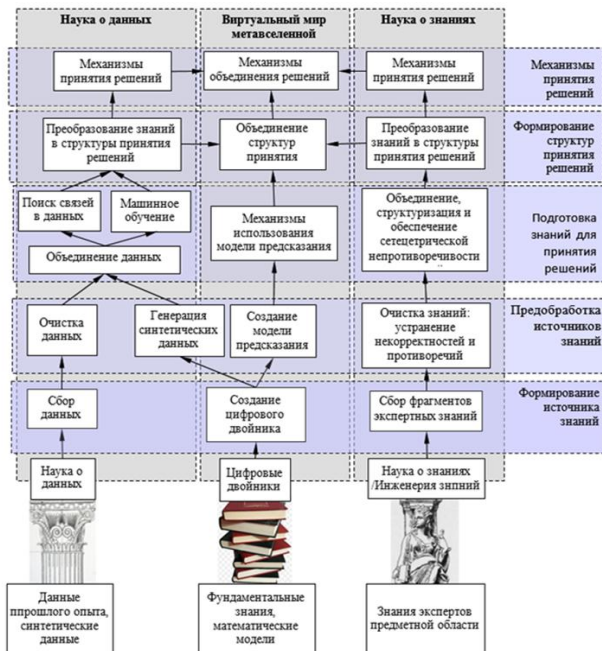


Рис. 1. Источники знаний, теоретические основы их получения и стандартные этапы технологий формирования и слияния знаний для принятия решений

Многие компоненты технологии ЦД опираются на технологии извлечения знаний на основе данных, что демонстрируется на рис. 1 (центральная колонка). Заметим, что прикладная значимость ЦД как источника знаний в настоящее время повышается за счет интеграции его технологий с технологиями метавселенной [Lee et al., 2021].

Что касается экспертных знаний, т.е. третьего источника знаний в ИИ, то, как отмечено выше, в настоящее время его роль сведена практически к нулю. До появления Науки о данных в течение нескольких десятилетий развитие ИИ шло без акцента на ведущую роль данных, ИИ рассматривали как науку о знаниях, а приложения ИИ называли системами, основанными на знаниях. Основными задачами ИИ в этот период были задачи получения, представления и использования знаний, причем в первую очередь – *экспертных* знаний. В период до появления больших данных и Науки о данных специалисты ИИ полагали, что эксперты являются основными источниками знаний, хотя методы машинного обучения на основе данных развивались параллельно инженерии знаний еще с начала 1960-х годов. Но с появлением больших данных и технологий глубоких нейросетей появилось мнение, что экспертные знания практически бесполезны, и это заметно затормозило развитие инженерии знаний.

Однако, как показано во Введении, в настоящее время можно утверждать, что ситуация меняется в пользу повышения роли экспертных знаний как следствие появления актуальных ИТ-приложения нового поколения, которые принципиально отличаются от привычных нам приложений прошлого. Примеры таких приложений, перечисленные выше во Введении, существенно отличаются от тех, с которыми ранее имела дело инженерия знаний. Приложения ИИ нового поколения, как правило, являются крупномасштабными и очень сложными, нужны экспертные знания для них могут быть получены от большого числа экспертов разных специализаций и разной квалификации. Современные технологии инженерии знаний не имеют алгоритмической основы и программных средств, способных обрабатывать огромные объемы фрагментов сырых экспертных знаний, содержащих разные мнения и представления большого числа экспертов и, как правило, противоречивых. При этом особенно трудными являются задачи локализации противоречий в разных моделях формализации экспертных знаний и выработка рекомендаций по их устранению.

В современной ситуации с обработкой экспертных знаний можно заметить достаточно много аналогий с тем, что происходило в области обработки данных в период, предшествующий появлению больших данных. Аналогично прошлой ситуации с большими данными, в настоящее время имеется четкое понимание высокой ценности экспертных знаний в при-

ложениях нового поколения, о которых здесь идет речь, и отсутствие эффективных методологий, методов, алгоритмов и инструментов доступа к этой ценности в нужной форме со стороны приложений.

Аналогию современной ситуации в ИИ, которая обусловила полтора десятилетия тому назад появление Науки о данных, и современной ситуации с экспертными знаниями можно рассматривать как тренд к возрождению инженерии знаний в новом качестве. Действительно, сырые экспертные знания в большом числе критически важных современных и перспективных приложений являются практически единственным источником знаний, а инженерия знаний оказывается не в состоянии обеспечить доступ приложений к этим знаниям. Системы (поддержки) принятия решений работают с очищенными и непротиворечивыми знаниями, которые представлены в определенной форме, диктуемой интерфейсом решателя (*англ.* actionable knowledge), и чтобы привести сырые знания экспертов к нужной форме, требуются непростые алгоритмы и программные инструменты.

Теоретическая разработка и программная реализация новых технологий работы с экспертными знаниями – это фундаментальная задача. Сходные по содержанию научные проблемы с доступом к знаниям, скрытым в больших данных, породили в свое время Науку о данных. Аналогичные, но все же более сложные проблемы и задачи в области обработки *больших объемов экспертных знаний* и их преобразование в форму, которая диктуется используемыми механизмами принятия решений требуют исследований и разработок в рамках науки, специально предназначенной для решения указанных проблем работы с экспертными знаниями. Эта наука в данной работе названа Наукой и знаниях (*англ.* Knowledge Science). Ее потенциальное содержание и основные направления исследований и разработок очерчиваются в следующем разделе.

2. От инженерии знаний к Науке о знаниях

Гипотетическая (пока) Наука о знаниях, как и всякая другая наука, претендующая на самостоятельную роль, должна иметь свой специальный объект (предмет) исследований, а также базовые направления исследований и разработок. В качестве специального объекта исследований в Науке знаниях естественно рассматривать сырые предметно-ориентированные экспертные знания и их множества, а в качестве основных научных направлений следует рассматривать те направления, научные результаты которых смогут обеспечить преобразование исходных сырых множеств экспертных знаний в структуры знаний для принятия решений.

Анализ технологии работы с экспертными знаниями приводит к выводу о том, что основные этапы их преобразования в готовый продукт и существо выполняемых при этом операций по смыслу совпадают с теми,

которые на рис. 1 приведены для двух других источников знаний. Названия этих этапов, общие для все трех технологий получения знаний приведены на этом рисунке справа и отделены друг от друга горизонтальными пунктирными прямыми. Их содержание применительно к экспертным знаниям можно рассматривать как грубую оценку будущего содержания исследований и разработок в Науке о знаниях. Дадим их краткое описание. При этом результаты инженерии знаний рассматриваются как начальные условия для развития Науки о знаниях.

1. *Сбор экспертных знаний.* Эта часть общей теории и технологии получения знаний для принятия решений достаточно хорошо исследована в инженерии знаний, и многое, что было в ней достигнуто в этой части, должно использоваться в Науке о знаниях.

2. *Предварительная обработка собранных знаний.* Задача этого уровня в части работы с множеством сырых фрагментов экспертных знаний формулируется как «Очистка и устранение противоречий в множестве фрагментов экспертных знаний». В инженерии знаний этот технологический этап слабо поддержан алгоритмической базой и программными средствами даже для пространств экспертных знаний низкой размерности и множеств фрагментов экспертных знаний небольшой мощности. На текущем этапе эту задачу, возможно, следует считать первоочередной задачей исследований в Науке о знаниях.

3. *Подготовка знаний для принятия решений.* Для данного источника знаний эту задачу можно формулировать как задачу машинного обучения на основе множества фрагментов экспертных знаний. Она имеет целью построение пространства принятия решений (на основе анализа зависимостей между факторами), непротиворечиво описывающими собранные знания, отбор непротиворечивых фрагментов знаний, относящихся построенному пространству принятия решений и представление этих знаний в машиночитаемой форме, например, в виде спецификации на формальном языке графовых баз данных и знаний с использованием некоторой стандартной программной платформы,

4. *Формирование структур принятия решений.* Существо этого этапа трансформации экспертных знаний состоит в том, чтобы преобразовать знания базы знаний в форму, в которой эти знания можно загружать в готовую систему принятия решений.

5. *Разработка конкретного механизма принятия решений.* Механизм принятия решений заранее выбранной структуры формируется на этом этапе технологии работы с экспертными знаниями. При наличии знаний из других источников, на этом этапе могут решаться также задачи интеграция знаний из разных источников и/или объединение решений, полученных на основе знаний разных источников.

3. Устранение противоречий и формирование пространств знаний – ключевые задачи Науки о знаниях

Задачи, которые рассматриваются в данном разделе, относятся к этапу 2 технологии работы с массивом сырых экспертных знаний и к этапу 3 (см. правую колонку на рис. 1). Проведенный предварительный анализ содержания задачи очистки экспертных знаний и выявления и устранения противоречий в них, а также относительно небольшой опыт ее практического решения автором позволяют выделить на этом этапе три ключевых проблемы, которые имеют аналоги в технологиях машинного обучения на основе данных. Последнее легко объяснимо, поскольку в задачах обработки экспертных знаний, как и в аналогичных задачах обработки данных, приходится работать с информационными структурами, но в случае экспертных знаний эти структуры сложнее, и потому в этом случае потребуются более фундаментальные исследования, и использование более сложных алгоритмов.

Среди задач обработки сырых экспертных знаний, стандартно решаемых на этом этапе, можно выделить три, которые требуется решать почти всегда. Они рассматриваются ниже.

а) *Задача обнаружения связей между координатами пространства факторов, представляющих экспертные знания.* В любой системе различные факторы, определяющие состояние системы, обычно взаимосвязаны, и реальная размерность факторного пространства может быть значительно меньше общего числа факторов, которые описывают модель экспертных знаний. К тому же, снижение размерности этого пространства всегда ведет к ускорению вычислительных процессов и повышению устойчивости самих вычислений.

Снижение размерности пространства данных является хорошо известной и практически полностью формализованной задачей технологии машинного обучения на основе данных. Взаимосвязи данных обычно задаются различными формальными моделями, например, корреляционными матрицами факторов, моделями "причина – следствие", регрессионными моделями или другими формальными моделями.

В когнитивных исследованиях для формального представления связей факторов на основе экспертной информации используются различные модели на основе графов, которые называются в общем случае картами знаний. При этом наиболее часто используются модели этого типа, называемые когнитивными картами ([Горелова, 2021], [Авдеева и др., 2011], [Авдеева, 2018]). В указанных работах можно найти много примеров конкретных когнитивных карт для разных приложений когнитивного моделирования. Кратко суть этой модели формализации экспертных знаний описывается ниже.

Когнитивная карта задает модель парных связей между факторами заданного множества в виде ориентированного взвешенного графа, в котором вершины взаимно однозначно соответствуют факторам состояния системы, а дуги представляют бинарные отношения между факторами. Эти связи являются направленными, и в стандартном случае они интерпретируются для пары факторов-вершин графа как причинно-следственное влияние одного фактора на другой. При этом каждой дуге этого отношения ставится в соответствие число из интервала $[-1,+1]$, которое интерпретируется как «сила» причинно-следственной связи, положительной или отрицательной. Обычно предполагается, что пара вершин и направленная связь между ними формально представляют правило "если..., то...", причем это влияние может быть как усиливающим (положительным), так и тормозящим (отрицательным).

Однако иногда эта связь интерпретируется иначе. Иногда связь между парой вершин графа имеет для каждой из них входящую и исходящую дугу, и тогда такую связь уже нельзя интерпретировать как причинно-следственную. Поэтому первое, о чем нужно договориться в этой модели, это о вариантах формальной интерпретации меры связи, заданной на дугах графа когнитивной карты и определить аксиоматику этой меры. Если вершины представляют случайные события, то эту меру можно интерпретировать, например, как коэффициент регрессии одного события (ему соответствует вершина, из которой дуга исходит) на другое, отвечающее входу дуги. Тогда всю дальнейшую формальную обработку фрагментов сырых знаний экспертов следует вести в контексте аксиоматики вероятностной меры связи, называемой коэффициентом регрессии и модели вероятностей на структурах, представленных графами. Если факторы являются случайными величинами, то тогда связи между ними можно интерпретировать как коэффициенты корреляции, а это уже совсем другая мера вероятностной связи, для которой есть свои закономерности, например, неотрицательность матрицы корреляций и другие, которые можно строго сформулировать и проверять формально. Возможны и другие формальные интерпретации меры связи между вершинами графа когнитивной карты и их исследование – это одна из важных задач науки о знаниях. На основе каждой формализации можно будет построить различные алгоритмы проверки непротиворечивости и локализации пар противоречивых связей при их наличии. Например, в самом простом случае причинных связей условием непротиворечивости множества фрагментов знаний, представленных такой когнитивной картой, будет отсутствие циклов в ее графе, а противоречия нужно будет искать в обнаруженных циклах.

Напомним, что описанная задача решается с целью последующего уменьшения размерности пространства факторов или его выбора по другим показателям, например, для обеспечения робастности прини-

маемых решений (за счет избыточности, например). Заметим, что в описанной модели и подходах к ее использованию в задачах обработки экспертных знаний просматривается аналогия с некоторыми моделями обработки данных.

Подчеркнем, что текущие разработки в области когнитивных исследований, к сформулированной задаче обнаружения и устранения противоречий в экспертных знаниях не относятся. Это задача является задачей будущих исследований гипотетической Науки о знаниях.

б) *Задача формального описания важности (весов) различных координат пространства экспертных знаний.* Условно под пространством знаний здесь понимается множество утверждений о различных свойствах объектов, явлений, событий или процессов, к которым относятся сырые экспертные знания. Например, в задачах оценки напряженности региональных конфликтов имеются такие измерения, как экономические противоречия, политические противоречия, религиозные и другие противоречия конфликтующих сторон. Оценка важности каждого аспекта рассматриваемого конфликта при оценке напряженности конфликта в целом – это пример прикладной задачи того типа, о которой идет речь в данном пункте. Обычно для решения таких задач используются различные подходы и приемы, развитые в теории многокритериального выбора. Среди них наиболее теоретически обоснованным и чаще всего используемым на практике является Метод Анализа Иерархий (МАИ) Т. Саати [Саати, 1993]. Этот метод широко известен, в особенности, специалистам в области теории систем, поэтому здесь он не рассматривается. Метод МАИ хорошо себя зарекомендовал на практике в самых разных классах приложений, причем не только в области бизнеса. Отметим, что в этом методе задача анализа непротиворечивости матрицы парных сравнений является частью более общей задачи, которая называется в МАИ задачей согласования матрицы парных сравнений.

Однако в текущем состоянии алгоритмизации этого метода он не в состоянии работать с матрицами парных сравнений большой размерности и требуется его существенное развитие для того, чтобы он был в состоянии локализовать противоречия в таких матрицах и давать рекомендации экспертам по устранению обнаруженных противоречий.

в) *Субъективные вероятностные оценки.* Одна из форм экспертных знаний в моделях с неопределённостью – это субъективные оценки вероятностей тех или иных событий, случайных величин, их вероятностных зависимостей, аналогичных зависимостей на множестве правил и т.п. Собственная практика автора показала, что подходящей формальной мо-

делью описания экспертных знаний такого типа может быть вероятностная модель с интервальной вероятностной мерой ([Городецкий и др., 1997], [Городецкий, 1997]). Чтобы стало понятней, о чем идет речь и каковы возможности этой модели, рассмотрим простой численный пример.

Одна из часто используемых форм экспертной информации в моделях знаний с неопределенностью – это правила, которые в посылке содержат некоторую формулу на множестве пропозициональных переменных, а в заключении – имя класса из заданного множества. При этом правилу в целом ставится в соответствие значение вероятности истинности. Обычно каждой такой формуле можно поставить в соответствие некоторую содержательную интерпретацию.

Предположим, что экспертам задаются вопросы о вероятностях истинности некоторых таких утверждений, и от экспертов не требуется знать, какая формула ставится в соответствие тому или иному вопросу. При этом от эксперта не требуется также формулировка ответа в виде числа; он может задать свое представление об истинности того или иного утверждения из таблицы рис. 2 в виде интервала, причем этот интервал может совпадать с интервалом $[0,1]$, если ответ эксперта «не знаю». Полагается также, что эксперт иногда может

Интервальные оценки истинности формул, как известно, отвечают вероятностной модели Демпстера – Шафера ([Dempster, 1966], [Shafer, 1976]), в которой нижняя граница интервала называется мерой доверия, а верхняя – мерой правдоподобия), или модели Халперна-Фагина-Мигиддо ([Fagin et al., 1988], [Fagin et al., 1989]), в которой эти границы трактуются в терминах теории меры как внутренняя и внешняя мера, соответственно. При некоторых условиях обе эти меры имеют одинаковую формальную основу.

Символ	Содержание высказывания	Символ	Содержание высказывания
Φ_1	x_1 характерно для x_2	Φ_{10}	x_1 обычно сопровождается x_2
Φ_2	x_1 наблюдается при x_2	Φ_{11}	при x_1 , как правило, x_2
Φ_3	x_1 отмечается при x_2	Φ_{12}	при x_1 обычно x_2
Φ_4	x_1 есть проявление x_2	Φ_{13}	x_1 иногда сопровождается x_2
Φ_5	x_1 есть признак x_2	Φ_{14}	x_1 часто сопровождается x_2
Φ_6	x_1 сопровождает x_2	Φ_{15}	x_1 исключает x_2
Φ_7	x_1 часто сопровождает x_2	Φ_{16}	x_1 приводит x_2
Φ_8	при x_1 часто присутствует x_2	Φ_{17}	при x_1 возникает x_2
Φ_9	x_1 может наблюдаться при x_2	Φ_{18}	x_1 может привести к x_2

Рис. 2. Примеры вопросов, которые инженер по знаниям задает экспертам о вероятностях тех или иных утверждений (Пример заимствован из работы [Городецкий и др., 1997])

Предположим, что после интервью с экспертами инженер по знаниям заполнил некоторые позиции в таблице, представленной на рис. 3. В ней каждому вопросу, который был задан эксперту, в первом столбце этой таблицы поставлена в соответствие пропозициональная формула, отвечающая содержательной формулировке вопроса, заданного эксперту. Во втором столбце с наименованием *Оценка эксперта* записан ответ эксперта в виде интервала истинности соответствующей формулы. Пусть дополнительно эксперт сообщил, что имеют место, по его мнению, неравенства

$$p(x_1) \geq p(x_2), p(x_1 \wedge \bar{x}_2) \geq p(x_1 \wedge x_2), \text{ и } p(x_1) \geq 2p(x_1 \wedge x_2). \quad (1)$$

Очевидно, что эти ограничения достаточно сложные, чтобы для них можно было найти решения вручную.

Далее, была произведена обработка этих данных в контексте вероятностной аксиоматики. Результаты представлены в третьем столбце таблицы, приведенной на рис. 3. Заметим, что задача оценки непротиворечивых интервалов вероятностей для заданных их начальных довольно грубых экспертных оценок сводится к задаче линейного программирования [Городецкий и др., 1997]. Результаты решения такой задачи для полученных от эксперта исходных оценок вероятностей с учетом неравенств (1) представлены в третьем столбце таблицы, приведенной на рис. 3. Очевидно существенное повышение точности экспертной информации.

Приведенный пример наглядно демонстрирует возможности формализации экспертных знаний в условиях неопределенности с привлечением вероятностных моделей, оперирующих с интервальными оценками.

Formulae	Expert's estimate	Consistent estimate	Formulae	Expert's estimate	Consistent estimate
$y_1 = p(x_1)$	[0.3, 0.9]	[0.3, 0.4]	$y_7 = p(x_1 \& \bar{x}_2)$	[0, 1]	[0, 0.1]
$y_2 = p(x_2)$	[0, 1]	[0.3, 0.4]	$y_8 = p(\bar{x}_1 \& \bar{x}_2)$	[0, 1]	[0.5, 0.7]
$y_3 = p(\bar{x}_1)$	[0, 1]	[0.6, 0.7]	$y_9 = p(\bar{x}_1 \supset x_2)$	[0.2, 0.9]	[0.2, 0.9]
$y_4 = p(\bar{x}_2)$	[0.1, 0.8]	[0.6, 0.7]	$y_{10} = p(x_1 \supset x_2)$	[0, 1]	[0.9, 1]
$y_5 = p(x_1 \& x_2)$	[0, 1]	[0.3, 0.33]	$y_{11} = p(\bar{x}_1 \supset \bar{x}_2)$	[0, 1]	[0.9, 1]
$y_6 = p(\bar{x}_1 \& x_2)$	[0, 1]	[0, 1]	$y_{12} = p(x_1 \supset \bar{x}_2)$	[0.2, 0.7]	[0.67, 0.7]

Рис. 3. Пример исходной экспертной информации, представленной распределением вероятностей с интервальной вероятностной мерой и результатов ее обработки с учетом экспертных ограничений и ограничений, заданных вероятностной аксиоматикой (Пример заимствован из работы [Городецкий, 1997])

Заключение

Качество функционирования любого приложения ИИ определяется, прежде всего, качеством и полнотой знаний, которые используются им для решения различных его частных задач и, прежде всего, задач принятия решений. Простое доказательство тому – это революционный рост интереса ИТ-индустрии к технологиям ИИ в последнее десятилетие за счет того, что технологии больших данных и глубокого обучения сделали доступным новый богатый источник знаний. Очередной прорыв в этом направлении станет возможным только тогда, когда появится новый источник знаний и эффективные технологии доступа к ним.

В работе аргументируется, что таким источником знаний могут стать экспертные знания, которые в настоящее время практически не используются, хотя они очень востребованы в ряде критически важных, актуальных и, зачастую, уникальных классов приложений, для которых эксперты являются единственным доступным источником знаний.

В работе показано, что для решения проблемы эффективного доступа к экспертным знаниями необходимы интенсивные исследования и разработки в области инженерии знаний, способные поднять ее до уровня *Науки о знаниях*, которая будет в состоянии решать задачи обработки больших объемов сырых экспертных знаний того же масштаба и сложности, которые в настоящее время под силу Науке о данных по отношению к большим данным.

Список литературы

- [Авдеева, 2018] Авдеева З.К. Лекция по теории и практике когнитивных карт. https://mipt.ipu.ru/sites/default/files/page_file/Авдеева.pdf.
- [Авдеева и др., 2011] Авдеева З.К., Коврига С.В. Формирование стратегии развития социальноэкономических объектов на основе когнитивных карт. Saarbrücken: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2011. – 184 с.
- [Гаврилова и др., 2016] Гаврилова Т.А., Кудрявцев Д.В., Муромцев Д.И. Инженерия знаний. Модели и методы: учебник. – СПб.: Изд-во «Лань», 2016.
- [Горелова, 2021] Горелова Г.В. Когнитивное моделирование сложных систем: Состояние и перспективы. Системный вплив в проектировании и управлении. Сборник научных трудов XXV Международной научной и учебно-практической конференции. В 3-х ч. Ч. 1. – СПб., 2021.
- [Городецкий и др., 1997] Городецкий В.И., Тулупьев А.Л. Формирование непротиворечивых баз знаний с неопределенностью // Известия РАН «Теория и системы управления». – 1997. – № 5.
- [Городецкий, 1997] Городецкий В.И. Интервальные вероятностные меры неопределенности в инженерии знаний // Юбилейный сборник трудов СПИИРАН. – 1997.

- [Саати, 1993] Саати Т. Принятие решений. Метод анализа иерархий: пер. с англ. Р.Г. Вачнадзе. – М.: Радио и связь, 1993.
- [Augier et al., 2002] Augier M. and Vendelo M.T. An Interview with Edward A. Feigenbaum. – https://www.researchgate.net/publication/236904576_An_Interview_with_Edward_A_Feigenbaum.
- [Asch, 2022] Asch M.A Toolbox for Digital Twins: From Model-Based to Data-Driven. Society for Industrial and Applied Mathematics. – Philadelphia, 2022. – 826 p. – <https://doi.org/10.1137/1.9781611976977.fn>.
- [Lee et al., 2021] Lee L., Braud T., Zhou P., et al, All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda, 2021, arXiv:2110.05352.
- [Dempster, 1966] Dempster A.P. Upper and Lower Probabilities induced by a Multi-valued Mapping // Annals of Mathematical Statistics. – 1966. – 36.
- [Fagin et al., 1988] Fagin R., Halpern J.Y., Megiddo N.A. Logic for Reasoning about Probabilities // In Proceedings 3th IEEE Symposium on Logic and Computer Science. – 1988. (The extended version of the paper is published as Report RJ6190(60900), 4/12/1988).
- [Fagin et al., 1989] Fagin R., Halpern J.Y. Uncertainty, Belief, and Probability // In Proceedings of 11th International Joint Conference on Artificial Intelligence. – Morgan Kaufman Publ. Detroit, Michigan, USA, 1989.
- [Shafer, 1976] Shafer G. A Mathematical Theory of Evidence. – Princeton University Press, 1976.

АППАРАТ ПОНЯТИЙ ДСМ-МЕТОДА АВТОМАТИЗИРОВАННОЙ ПОДДЕРЖКИ ИССЛЕДОВАНИЙ

В.К. Финн (*ira.finn@gmail.com*)
Федеральный исследовательский центр
«Информатика и управление» РАН, Москва

В докладе рассмотрен аппарат понятий (определения, аксиомы, описание эмпирических данных, методологические принципы и эвристики «knowledge discovery») ДСМ-метода автоматизированной поддержки исследований для конструирования интеллектуальных систем, реализующих специальный класс правдоподобных рассуждений таких, что они порождают гипотезы о причине, гипотезы о предсказании исследуемых эффектов и эмпирические закономерности.

Ключевые слова: ДСМ-метод, аппарат понятий, автоматизированная поддержка исследований.

Термин «аппарат понятий» ввел К. Айдукевич в [Ajdukiewicz, 1978]. Аппарат понятий есть множество релевантных понятий, посредством организации которых характеризуется некоторая область исследований (определения, аксиомы, описание эмпирических данных, методологические принципы и эвристики «knowledge discovery»).

ДСМ-метод автоматизированной поддержки исследований (ДСМ-метод АПИ) [ДСМ-метод, 2009], [Финн, 2023] есть система понятий, методологические принципы, средства представления знаний для конструирования интеллектуальных систем (ИС-ДСМ), реализующих специальный класс правдоподобных (амплиативных в смысле Ч.С. Пирса [Kapitan T., 1997]) рассуждений таких, что они порождают гипотезы о причине, гипотезы о предсказании исследуемых эффектов и эмпирические закономерности.

Необходимыми условиями применимости ДСМ-метода АПИ является существование в базах фактах отношений причинности таких, что они сохраняются при расширениях баз фактов [Финн, 2023]. Сохранение гипотез о причинах и гипотез о предсказаниях посредством этих причин порождают *эмпирические закономерности* (ЭЗК). Посредством ЭЗК определяются эмпирические модальности (необходимости, возможности и слабой возмож-

ности). Порождение ЭЗК может быть продолжено расширениями баз фактов, что породит последовательность модальностей, характеризующих ЭЗК ранга r , где $r \geq 1$, а r – число периодов порождения ЭЗК.

Средством порождения ЭЗК являются *ДСМ-рассуждения* такие, что они образованы синтезом индукции (порождение причин), аналогии (порождение предсказаний) и абдукции (принятие гипотез о причине). ДСМ-рассуждения реализуются на достаточном основании [Челпанов, 1994], которым является операциональность порождения причин, обоснованность порождаемых гипотез как носителей ЭЗК, монотонность неубывания качества характеристик ДСМ-рассуждения и верификация результатов. ДСМ-рассуждения являются средством *логики рассуждений* ДСМ-метода.

Методологическими основаниями логики рассуждений являются *концепция эмпирической качественной причинности* (КЭКП) (она – онтология ДСМ-метода АПИ), *конструктивная эволюционная эпистемология* (КЭЭ), содержащая *концепцию истин ДСМ-метода* (КИ-ДСМ). В КИ-ДСМ (она выразима в метаязыке) определяются три вида истин – корреспондентная (верификация результатов ДСМ-метода), когерентная (оценка гипотез о причине и предсказании), прагматическая (оценка ЭЗК с учетом точки зрения интерпретатора).

Логика рассуждений есть средство *эвристики* ДСМ-метода АПИ, которая осуществляется посредством подсистемы *когнитивного интерфейса интеллектуальной системы* (ИС-ДСМ), реализующей ДСМ-метод АПИ.

ИС-ДСМ имеют открытые базы фактов и базы знаний, решатель задач и когнитивный интерфейс. ИС-ДСМ осуществляет *интеллектуальный анализ данных* (ИАД), результатом которого является семейство *открытых квазиаксиоматических теорий* (КАТ). КАТ состоит из множества аксиом структуры данных, процедурных и дескриптивных (последние характеризуют предметную область), открытого множества фактов, множества гипотез, множества обнаруженных ЭЗК и правил вывода, содержащих правила индуктивного вывода, вывода по аналогии и правил абдуктивного вывода. Результаты ИАД получают присвоение трех видов *доверия* – максимального, допустимого, минимального. Возможны присвоения неопределенного доверия и недоверия. Все виды доверия и недоверия реализуются в модуле когнитивного интерфейса ИС-ДСМ и используют прагматические истинностные значения (прагматические истину, неопределенность и ложь). Перечисленные понятия ДСМ-метода образуют *остов* аппарата понятий ДСМ-метода. Эти понятия в свою очередь характеризуются исходными релевантными понятиями.

Аппарат понятий ДСМ-метода АПИ является возможным ответом на проблемы, сформулированные в статье Д. Маккарти и Д. Хейеса [Маккарти и др., 1972]: необходимость уточнения идеи «интеллекта», создание принципов эпистемологии для «knowledge discovery» и эвристики его порождения.

Список литературы

- [Ajdukiewicz, 1978] Ajdukiewicz K. The World-Picture and the Conceptual Apparatus. В кн.: K. Ajdukiewicz The Scientific World-Perspective and other essays, 1931-1963. D. Reidel Publishing Company Dordrecht, Holland, 1978. – P. 67-89.
- [ДСМ-метод, 2009] ДСМ-метод автоматического порождения гипотез. Логические и эпистемологические основания / под общей ред. О.М. Аншакова. – М.: ЛИБРОКОМ, 2009. – 432 с.
- [Финн, 2023] Финн В.К. Интеллект, информационное общество, гуманитарное знание и образование. – М.: ЛЕНАНД, 2023.
- [Kapitan, 1997] Kapitan T. Peirce and the Structure of Abductive inference // В кн.: Studies in the Logic of Charles Sanders Peirce / Eds. N. Houser, Don D. Roberts and James Van Evra. – Indiana University Press, 1997. – P. 477-496.
- [Челпанов, 1994] Челпанов Г.И. Учебник логики. – М.: Прогресс, 1994.
- [Маккарти и др., 1972] Маккарти Дж., Хейес Р. Дж. Некоторые философские проблемы в задаче построения искусственного интеллекта // В кн.: Кибернетические проблемы бионики. – М.: МИР, 1972. – С. 40-88.

Секция 1 | ИНЖЕНЕРИЯ ЗНАНИЙ

УДК 004.048

РАЗРАБОТКА ОНТОЛОГИИ ПРЕДМЕТНОЙ ОБЛАСТИ ДЕЯТЕЛЬНОСТИ НАУЧНО-ОБРАЗОВАТЕЛЬНОЙ ОРГАНИЗАЦИИ

Н.К. Андриевская (*nataandr@yandex.ru*)

Т.А. Васяева (*vasyaeva@gmail.com*)

Т.В. Мартыненко (*umbralada@gmail.com*)

Донецкий национальный технический университет, Донецк

В работе описывается подход к разработке онтологической модели, при котором на различных этапах существования онтологии были использованы различные способы ее формирования. На начальном этапе «базовая» онтологическая модель реализована в экспертном режиме. Дальнейшее расширение онтологии выполнялось в полуавтоматическом режиме. Разработаны следующие полуавтоматические способы: с помощью корпуса документов, на базе электронных словарей, с помощью NLP обработки текстовых источников, а также с использованием универсальной междоменной онтологии DBpedia.

Ключевые слова: онтология, Ontology Learning, SOAT, WordNet, Wiktionary, DBpedia.

Введение

В последние десятилетия для структурирования, формализации и унификации представления знаний с целью их многократного и гибкого использования в информационных системах применяются онтологии.

Проблемы создания и использования онтологий отражены в работах известных зарубежных и российских ученых, например, Т. Груббера [Gruber, 1995], Н. Гуарино [Guarino, 1998], Р. Крухтена [Kruchten, 2004], С. Стааба [Staab, 2010], Л.С. Козачкова [Козачков, 1985], Г.С. Осипова [Осипов, 1990], А.В. Смирнова [Смирнов и др., 2002], Т.А. Гавриловой [Гаврилова и др., 2006], А.Ф. Тузовского, В.З. Ямпольского [Тузовский и др., 2005] и др.

Разработке онтологий предметной области деятельности научно-образовательной организаций и порталов научных знаний посвящено достаточно большое количество работ, например: [Загорулько и др., 2007], [Олейник и др., 2016], [Фролова, 2016], [Садовничий и др., 2018], [Навроцкий и др., 2018].

Часть работ и материалов конференций описывают информационные системы, построенные на онтологическом подходе. Например, система проектирования Web-порталов с помощью разработанного инструментария разработки PORTO (PORTal Ontology) базируется на онтологиях [Gavrilova et al., 2003]. В другой работе российских авторов описана разработка онтологий и методов их наполнения для системы дистанционного обучения ECOLE (Enhanced Course Ontology for Linked Education) [Стафеев и др., 2015].

Следовательно, разрешение проблем формирования и наполнения онтологий является весьма перспективной задачей, в том числе и для предметной области научной и образовательной деятельности.

1. Этапы создания онтологической модели

Онтологическая модель проектируемой системы состоит из онтологии верхнего уровня и онтологий предметных областей.

В целом она служит для представления понятий, необходимых для описания процессов научной, методической и учебной деятельности, а также описания тематик конкретных научных исследований, дисциплин и непосредственно самих информационных ресурсов.

При разработке онтологической модели использовался комбинированный способ формирования онтологии [Андриевская, 2020]:

Подбор и использование онтологий верхнего уровня для формирования “базовой” онтологии.

“Экспертное” создание и редактирование каркаса “базовой” онтологии.

Использование при создании онтологии среднего уровня знаний, полученных в результате полуавтоматической обработки существующих корпусов документов.

Подбор и использование существующих глоссариев, тезаурусов для выбора имен понятий и населения экземплярами при полуавтоматическом формировании онтологий нижних уровней.

Использование знаний, полученных из известных междоменных и тематических онтологий, а также из открытых источников Интернет при полуавтоматическом формировании онтологий нижних уровней.

Итерационный процесс создания онтологий состоит из определенных этапов, каждый из которых позволяет извлекать из текста знания, изображен на рис. 1.

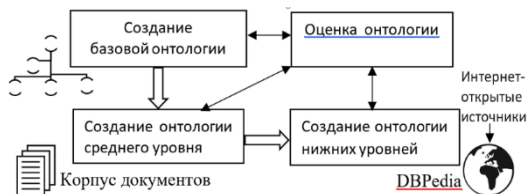


Рис. 1. Итерационный процесс создания онтологии

После экспертного построения первоначального варианта «базовой» онтологии, представляющей собой относительно неизменную часть и являющуюся каркасом, происходит оценка ее адекватности. В случае необходимости возможен переход к ее модификации и доработке. Процесс повторяется и для других уровней онтологии до тех пор, пока не будет достигнута определенная наполненность терминами конкретной предметной области.

2. Разработка «базовой» онтологической модели

С целью повышения непротиворечивости модели был разработан каркас прикладной онтологической модели на базе предметной области, называемый «базовой» онтологией.

На начальном этапе создания были изучены и адаптированы в «базовую» онтологическую модель следующие онтологии верхнего уровня: Dublin Core, FOAF, VIVO, BIBO, DBpedia, TEACH, VCard.

Для более наглядного представления онтологии применялся модульный принцип построения. Верхний уровень «базовой» онтологии, разработанной в редакторе Protégé, приведен на рис. 2.

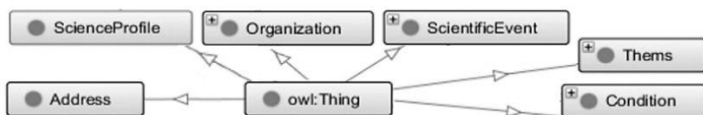


Рис. 2. Верхний уровень «базовой» онтологии

В табл. 1 описаны классы верхнего уровня «базовой» онтологии.

Классы верхнего уровня «базовой» онтологии

Имя класса	Класс в онтологии	Описание класса
Address	vCard:Address	Адрес
Condition		Состояние/Статус
Organization	foaf:Organization	Учреждение
ScientificEvent		Событие
ScienceProfile		Наукометрический профиль
Thems	Vivo:ReseachArea	Тематика ИП

В дальнейшем «базовая» модель была откорректирована, улучшена и расширена за счет процедур полуавтоматического пополнения знаний.

3. Способы полуавтоматического пополнения онтологии

Методология наполнения онтологий была и остается самым слабым звеном онтологической инженерии [Gomez-Perez et al, 2004]. Со времени публикации этого суждения (2004 г.) по тематике SOAT (Semi-automatic domain Ontology Acquisition Tool) выполнено значительное количество исследований и написано немало работ.

Рамаяя и др. предложили полуавтоматический подход для заполнения онтологии [Ramaiah et al., 2007]. Термины, связанные с тематикой архитектуры программного обеспечения и их экземпляры автоматически идентифицировались и извлекались путем разбора статей Википедии.

В одной из работ сбор данных для уровня предметных областей науки и образования производился из сторонних баз знаний и таксономий [Стафеев и др., 2015].

Задача пополнения онтологий на основе уже существующих знаний находится и сегодня в фокусе внимания, что говорит одновременно и о ее актуальности, и о ее нерешенности в целом.

Поскольку формирование качественной онтологии является довольно сложной и длительной задачей, то она была упрощена с целью скорейшего получения практических результатов, в частности, пока не рассматривался вопрос оценки качества построенной онтологии.

На первом этапе была решена задача извлечения терминов и концептов. При расширении онтологии терминами и концептами реализован комбинированный подход с использованием следующих способов:

- пополнение онтологии с помощью шаблонизированных и семантически размеченных документов;
- пополнение на базе словарных определений существующих словарей;
- пополнение на базе онтологии DBpedia, которая является производной от Википедии;

- пополнение способом, основанным на обработке произвольных источников с помощью NLP-технологий с целью обнаружения контекстно-зависимых слов, относящихся к одной предметной области.

Следующие этапы после нахождения терминов-кандидатов – распознавание новых концептов и сохранение данных в онтологию. В случае небольших баз знаний фиксированных предметных областей обычно считают, что все специфичные термины несут в себе новые концепты и считаются однозначно определенными [Федоренко и др., 2013].

В полуавтоматическом режиме в программном модуле эксперта реализована возможность принятия решения по целому ряду вопросов: является ли слово-кандидат концептом или термином? подлежит ли добавлению в онтологию? какой из предложенных алгоритмом семантически близких существующих концептов в наибольшей определяется словом-кандидатом. Таким образом на данном этапе реализации системы выполняется контроль процесса расширения и качества создаваемой онтологии.

3.1. Способ пополнения онтологии с помощью шаблонизированных документов

Большая часть методов, предназначенных для извлечения заданного количества терминов любой длины из коллекции документов и не различающих вхождения одного термина, укладывается в общую схему. Согласно этой схеме, метод извлечения терминов состоит из трех перечисленных ниже этапов [Астраханцев, 2014]:

- Предобработка текстового документа и сбор кандидатов: фильтрация слов и словосочетаний, извлеченных из коллекции документов, по статистическим и лингвистическим критериям.
- Подсчет признаков: перевод каждого кандидата в вектор признакового пространства (в простейшем случае – простая скалярная оценка). После этого может проводиться проверка – сравнение слов, которые могли бы быть емкими лексическими единицами, с эталонным списком ключевых слов.
- Вывод на основе признаков: оценка вероятности быть термином для каждого кандидата на основе значений признаков, последующая сортировка всех кандидатов, по этой оценке, и взятие заранее определенного числа кандидатов.

Модуль автоматического разбора документов обрабатывал такие типовые документы, как ВКР, научные статьи факультетских сборников и журналов, отчеты по курсовым работам, рефераты и т.п.

В первую очередь была решена задача извлечения текстовой информации из файлов различных форматов, таких как doc, pdf, docx, rtf. При извлечении из текста неструктурированной информации использована

библиотека морфологического анализа RHPMorphy. Далее был произведен парсинг текста с помощью регулярных выражений PCRE – библиотеки RHP, реализующий работу регулярных выражений в стиле Perl.

В результате обработки документа был сформирован предварительный список ранжированных по частоте слов-кандидатов в термины, для которых в дальнейшем будет проведена процедура распознавания новых концептов и непосредственно внесение в онтологию.

На рис. 3 изображен фрагмент программного окна, демонстрирующий результат извлечения ключевых слов, информации о преподавателях и персонах из текста статьи для дальнейшего помещения в онтологию.

Преподаватели
Андриевская x
Ключевые слова
УНИВЕРСИТЕТ x ТЕХНОЛОГИЯ x ОБРАБОТКА x ДАННЫЙ x ОТБОР x ФАКТОР x ЗАДАЧА x
КЛАССИФИКАЦИЯ x ПРОГНОЗИРОВАНИЕ x
ФИО
А.О. Петце x Петце А.О. x А.А. Ежов x С.А. Шумский x Чубукова И.А. x Барсегян А.А. x Куприянов М.С. x

Рис. 3. Результат обработки текста статьи

В случае ошибочного извлечения возможно вручную удалить слово из набора слов, а затем сохранить полученные данные.

3.2. Способ, базирующийся на использовании электронных словарей

Одним из самых распространенных способов пополнения онтологий является способ, основанный на использовании словарных терминов, полученных из электронных словарей и тезаурусов.

Реализованы способы расширения онтологии, базирующиеся на использовании слов, полученных из электронных словарей WordNet, RussNet, Wikipedia, Wiktionary.

Технология Wiki – современная интеллектуальная технология, позволяющая использовать модели знаний и выдавать поисковые запросы пользователям. MediaWiki является свободной программой и распространяется на условиях общественной лицензии GNU. В Semantic MediaWiki существует свой собственный встроенный инструмент для экспорта в RDF, позволяющий извлекать размеченную семантическую информацию из списка страниц. При разработке программного модуля, использующего ресурс Wikipedia для пополнения предметной онтологии системы, были использованы следующие библиотеки:

- библиотека Python WikipediaBot Framework, которая использует MediaWiki API (специальный интерфейс прикладного программирования) для взаимодействия с MediaWiki-системой с целью автоматизации, получения данных и внесения изменений;
- словарь Wiktionary – это свободно пополняемый многофункциональный многоязычный онлайн словарь и тезаурус. Wiktionary, как и все проекты MediaWiki имеет API.
- библиотека RDFLib сохраняет в файл типа OWL;

Среди достоинств Вики-технологий стоит отметить многоязычность платформы и единый API для всех Wiki-проектов. Кроме этого, фонд Викимедиа регулярно публикует дампы каждого из своих проектов бесплатно. Эти дампы доступны в виде больших XML-файлов, которые можно импортировать в базу данных SQL с помощью специального программного обеспечения «MWDumper», что позволит локально использовать ресурс и значительно сократить время поиска по сравнению с веб-версией.

Словарь Wiktionary кроме словарных входов с толкованиями, семантическими отношениями и переводами, содержит словарные пометы. По сравнению с WordNet, наиболее близким аналогом среди электронных словарей, включающих словарные пометы и находящийся в открытом доступе, Wiktionary содержит около 370 помет против 170 в WordNet [Крижановский и др., 2014]. Wiktionary превосходит WordNet по следующим параметрам: больший объем и более быстрое обновление материала (последнее особенно актуально); большее количество редакторов (сотни редакторов в Английском Викисловаре, десятки – в Русском Викисловаре).

Редакторами Викисловарей разработана система категорий словарных помет, призванная упорядочить и систематизировать словарные пометы. Наиболее детально на данный момент в Английском Викисловаре проработаны категории помет предметных областей, например, категории topical (см. рис. 4).

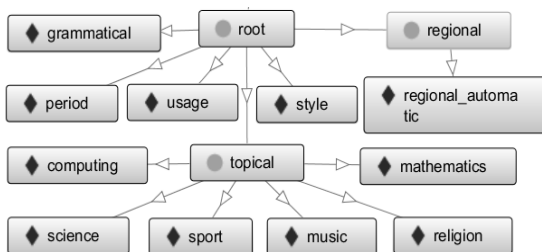


Рис. 4. Категории помет предметных областей для словаря Wiktionary

Программный модуль использует следующие библиотеки:

- cURL – библиотека функций, которая позволяет взаимодействовать с множеством различных серверов по различным протоколам;
- PHPQuery – библиотека для парсинга, аналог jQuery на PHP;
- MediaWikiApi – веб-служба, обеспечивающая доступ к таким функциям Wiki, как аутентификация, операции над страницами и поиск по Wiki;
- PHP Text Analysis – это библиотека PHP для выполнения задач поиска информации (IR) и обработки естественного языка (NLP);
- phpMorphy – библиотека для морфологического анализа, реализованная на платформе PHP.

Фрагмент программного модуля, содержащего описание некоторых классов, приведен на рис. 5.

```
class WikiSemantic implements ISemanticParsable {
    const TERMS_TAG = 'div.mw-parser-output ol li';
    private $parser;
    private $wikiApi;
    public function __construct()
    {
        $this->parser = new PHPQueryParser();
        $this->wikiApi = new WikipediaApi();
    }
    public function getTByW(string $word): string
    {
        return $this->parser->ParseText($this->wikiApi->
        GetWikiPage($word),self::TERMS_TAG);
    }
}

class WordSemantic implements ISemanticParsable {
    private $wordnet_api;
    public function __construct()
    {
        $this->wordnet_api = new WordNetApi();
    }
    public function getTByW(string $word): string
    {
        return $this->wordnet_api->
        getSynsetsGloss($word);
    }
}
```

Рис. 5. Фрагмент программного модуля

Целесообразность использования словарей оценивалась следующим образом. Эксперт предварительно разметил тестовые файлы, наполненные множеством слов по тематике информационных технологий, программирования и вычислительной технике. Программный модуль вернул результаты поисковых запросов по различным ключам для каждого конкретного словаря. Затем эксперт подсчитал количество полученных релевантных слов, количество полученных нерелевантных слов, количество релевантных слов, не выданных по запросу, количество нерелевантных слов, не выданных по запросу [Андриевская и др., 2020].

По результатам исследований рассчитанные значения показателя «F-мера» для различных словарей и различных ключей поиска помещены в табл. 2.

Таблица 2

Значения показателя F-мера для различных словарей

№	Ключ	WordNet	Wiki Pedia	Wiktionary	Wiktionary с topical
1	operating system	0,33	0,80	0,86	0,78
2	computer languages, programming language	0,67	0,92	0,91	0,93
3	server	0,82	0,91	0,9	0,89
4	protocol	0,80	0,84	0,86	0,88
..
20	database	0,40	0,22	0,86	0,86
	В среднем	0,69	0,68	0,8	0,87

Как показали эксперименты, словарь Wiktionary с уточнением предметной области topical, для которого показатель качества, выраженного F-мерой, достиг значения 0,87, обладает наибольшей терминологической полнотой для тестируемой предметной области и в среднем возвращает большее количество релевантных результатов.

3.3. Способ пополнения с использованием DBpedia

Существует много работ по тематике SOAT, в которых Википедия выступает в качестве ресурса для пополнения онтологии [Бочаров и др., 2010], [Астраханцев, 2014], [Кравцов, 2015].

Опираясь на подход, описанный в работе [Bhat et al., 2017], будем в качестве ресурса использовать онтологию DBpedia, которая в последнее десятилетие развивалась быстрыми темпами благодаря усилиям открытого сообщества. DBpedia – это универсальная онтология, основанная на наиболее часто используемых информационных блоках в статьях Википедии, содержит более 685 классов, 2795 различных свойств и более 4,2 миллиона экземпляров.

На рис. 6 представлен фрагмент онтологии DBpedia, содержащий необходимые понятия и их взаимосвязи для извлечения данных по предметной области, связанной с программным обеспечением.

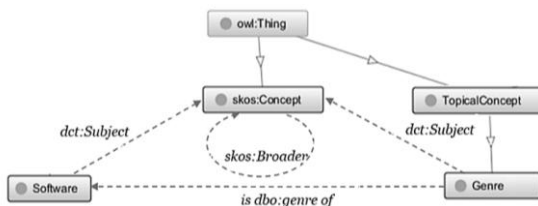


Рис. 6. Фрагмент онтологии DBpedia

Для извлечения концептов были использованы следующие свойства: `dbo:abstract`, `dct:subject`, `owl:sameAs`, `rdfs:comment`. Эти свойства в онтологии DBpedia (`dbo`) являются производными от верхних онтологий, включая Dublin core (`dct`), Web Ontology Language (`owl`), Resource Description Framework Schema (`rdfs`) и Simple Knowledge Organization System (`skos`), на что указывают их префиксы.

В качестве инструментального средства использован UIMA framework [UIMA, 2016]. Для работы с онтологией DBpedia применялся плагин DBpedia Spotlight, который выполняет именованное извлечение сущностей, включая их обнаружение и устранение неоднозначности [DBpedia Spotlight, 2020]. С его помощью выделены предложения, фразы, взвешены слова-кандидаты по метрике TF-IDF.

Для количественной оценки возможности использования данного способа извлечения терминов было выполнено эмпирическое исследование. В рамках тестового примера выбрали десять объёмных англоязычных статей, подготовленных в рамках научных конференций, связанных с информационными технологиями.

Поисковые запросы формулировались с помощью SPARQL и выполнялись встроенным движком ARQ Apache Jena. Запрос, извлекающий термины, представляющие собой по типу программные ресурсы для некоторого значения экземпляра `dbo:genre`, приведен на рис. 7.

```
SELECT ?software WHERE { ?software dbo:genre *variable .  
  ?software ns:type dbo:Software }
```

Рис. 7. Пример запроса на извлечение терминов

В запросе параметр `*variable` заменялся соответствующим значением, например, «URI: [http://dbpedia.org/resource/Relational database](http://dbpedia.org/resource/Relational_database)». В список результатов для ключа поиска «Relational database» такие реляционные БД, как «SQLite», «MySQL», «MariaDB». Остальные запросы строятся подобным образом.

Эксперт проанализировал тексты и вручную посчитал количество соответствующих запросам терминов (Expert). Документы были загружены в веб-клиент системы. Далее экспертом было определено количество полученных релевантных слов (TP), количество полученных нерелевантных слов (FP), количество релевантных слов, не выданных по запросу (TN) и рассчитаны показатели качества поиска концептов такие, как точность, полнота и F-мера. В табл. 3 приведены результаты ручной и автоматической разметки по всей группе документов, а также показатели качества поиска концептов.

Показатели качества поиска концептов

Expert	TP	FP	TN	Точность	Полнота	F-мера
266	241	35	67	0,873	0,782	0,825

Таким образом, 266 элементов были вручную размечены экспертом. Аннотатор DBpedia с точностью 0,87 (доля автоматически извлекаемых элементов, которые являются релевантными), полнотой 0,78 (доля релевантных элементов, которые были успешно извлечены) и F-мерой 0,83 (гармоническое среднее точности и полноты) автоматически аннотировал 241 элемент. Приведенные выше результаты тестирования показывают, что в общем концепты могут быть извлечены из документов с использованием общедоступной онтологии DBpedia.

Заключение

На различных этапах создания онтологии целесообразно использовать различные способы ее создания. На начальном этапе использовались существующие онтологии «верхнего уровня», а также способ «экспертного создания». При формировании онтологий нижних уровней использовались полуавтоматические способы обработки корпусов документов, словарные способы пополнения, а также ранее разработанные универсальные онтологии.

Исследования показали, что электронные словари-тезаурусы WordNet и его русскоязычный аналог RussNet можно рассматривать как элементы для построения требуемой лингвистической онтологии. Для пополнения предметной онтологии нижних уровней предпочтительнее использовать Wiki-ресурсы, в частности словарь Wiktionary, для которого показатель качества поиска, выраженного F-мерой, достиг самого высокого среди тестируемых словарей значения 0,87 для данной предметной области.

Также был разработан способ пополнения онтологии с использованием междоменной онтологии DBpedia. Исследования показали, что концепты могут быть извлечены из документов с точностью 0,87 (доля автоматически извлекаемых элементов, которые являются релевантными), полнотой 0,78 (доля релевантных элементов, которые были успешно извлечены) и F-мерой, равной 0,83.

Список литературы

- [Gruber, 1995] Gruber T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing // International Journal of Human-Computer Studies. – November 1995. –Vol. 43, Issues 5–6.
- [Guarino, 1998] Guarino N. Formal Ontology in Information Systems // Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Italy, June 6–8, 1998 / Ed. N. Guarino. – Amsterdam : IOS Press, 1998.

- [**Kruchten, 2004**] Kruchten P. An ontology of architectural design decisions in software intensive systems // in 2nd Groningen workshop on software variability. – 2004.
- [**Staab et al., 2010**] Staab S., Studer R. Handbook on ontologies. – Springer Science & Business Media, 2010.
- [**Козачков, 1985**] Козачков Л.С. Категориальные Тезаурусы в Базах Данных // Научно-Техническая Информация. – 1985. – Серия 2. – № 5.
- [**Осипов, 1990**] Осипов, Г.С. Построение модели предметных областей. Неоднородные семантические сети // Известия АН СССР. Техническая кибернетика. – 1990. – № 5.
- [**Смирнов и др., 2002**] Смирнов А.В., Пашкин М.П., Шилов Н.Г. Онтологии в системах искусственного интеллекта: возможности построения и организации // Новости искусственного интеллекта. – 2002. – № 1.
- [**Гаврилова и др., 2006**] Гаврилова Т.А., Кудрявцев Д.В. Горовой В.А. Модели и методы формирования онтологий // Научно-технические ведомости СПбГПУ. – 2006. – № 4.
- [**Тузовский и др., 2005**] Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии) / под общ. ред. В.З. Ямпольского. – Томск: НТЛ, 2005.
- [**Gavrilova et al., 2003**] Gavrilova T., Gorovoy V. Ontological Engineering for Corporate Knowledge Portal Design In "Processes and Foundations for Virtual Organisations", Eds. L. Camarinha-Matos and H. Afsarmanesh. – Kluwer Academic Publishers, 2003.
- [**Загоруйко и др., 2007**] Загоруйко Ю.А., Боровикова О.И. Технология построения онтологий для порталов научных знаний // Вестник НГУ. Серия: Информационные технологии. – 2007. – № 2.
- [**Олейник и др., 2016**] Олейник А.Г., Ломов П.А. Разработка онтологий интегрированного пространства знаний // Онтология проектирования. – 2016. – Т. 6, № 4 (22).
- [**Фролова, 2016**] Фролова, Н.Б. Разработка OWL-онтологии образовательных ресурсов СГУ // Вестник ВГУ, Серия: Системный анализ и информационные технологии. – 2016. – № 3.
- [**Садовничий и др., 2018**] Садовничий, В.А., Васенин В.А. Интеллектуальная система тематического исследования наукометрических данных: предпосылки создания и методология разработки. Ч. 1 // Программная инженерия. – 2018. – Т. 9, № 2.
- [**Навроцкий и др., 2018**] Навроцкий М.А., Жукова Н.А., Муромцев Д.И. Онтология проектирования, применения и сопровождения порталов научно-технической информации // Онтология проектирования. – 2018. – № 1 (27).
- [**Стафеев и др., 2015**] Стафеев С.К., Муромцев Д.И., Козлов Ф.А. ECOLLE - семантический агрегатор открытых образовательных ресурсов // КИО. – 2015. – № 1.
- [**Андриевская, 2020**] Андриевская Н.К. Разработка прикладной онтологии в системах обработки данных научных и научно-образовательных организаций // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2020. – № 3.

- [Gomez-Perez et al, 2004] Gomez-Perez A., Fernando-Lopez M., Corcho O. *Ontology Engineering*. – Springer – Ferlag, 2004.
- [Ramaiah et al., 2007] Ramaiah M.S., Prabhakar T.V., Rambabu D., and others. ArchVoc-Towards an ontology for software architecture // in *Sharing and Reusing Architectural Knowledge-Architecture, Rationale, and Design Intent*, 2007. SHARK/ADI'07: ICSE Workshops 2007.
- [Федоренко и др., 2013] Федоренко Д.Г., Астраханцев Н.А. Автоматическое извлечение новых концептов предметно-специфичных терминов // *Труды ИСП РАН*. – 2013.
- [Астраханцев, 2014] Астраханцев Н. Автоматическое извлечение терминов из коллекции текстов предметной области с помощью Википедии // *Труды Института системного программирования РАН*. – 2014. – Т. 26, № 4.
- [Крижановский и др., 2014] Крижановский А.А., Смирнов А.В., Круглов В.М., Крижановская Н.Б., Кипяткова И.С. Автоматическое извлечение словарных помет из Русского Викисловаря // *Тр. СПИИРАН*. – 2014. – 33.
- [Андриевская и др., 2020] Андриевская Н.К., Секирин А.И., Канатуш С.В. Анализ возможностей использования существующих словарей для пополнения онтологии // *Informatics & Cybernetics*. – 2020. – № 2 (20).
- [Бочаров и др., 2010] Бочаров В.В., Пивоварова Л.М. Википедия как ресурс для пополнения онтологии // *Новые информационные технологии в автоматизированных системах*. – 2010. – № 13.
- [Кравцов, 2015] Кравцов Д.В. Метод извлечения из Википедии предметно-ориентированных лингвистических онтологий // *Современные наукоемкие технологии*. – 2015. – № 12-3.
- [Bhat et al., 2017] Bhat M., Shumaiev K., Biesdorf A., Hohenstein U., Hassel M., Matthes F. An Ontology-based Approach for Software Architecture Recommendations // *23rd Americas Conference on Information Systems (AMCIS)*. – At: Boston, MA, USA. – 2017.
- [UIMA, 2016] *Unstructured information management applications*. Apache, UIMA. – 2016. – URL: <http://uima.apache.org>. Data access: 10.10.2023.
- [DBpedia Spotlight, 2020] *DBpedia Spotlight. Shedding light on the web of documents*. – URL: dbpedia-spotlight.org. Data access: 10.10.2023.

СОЗДАНИЕ ГРАФА ЗНАНИЙ НА ОСНОВЕ ОНТОЛОГИИ СЕМАНТИЧЕСКОЙ БИБЛИОТЕКИ*

О.М. Атаева (*oli.ataeva@gmail.com*)

С.А. Серебряков (*serebrvas@gmail.com*)

Н.П. Тучкова (*natalia_tuchkova@mail.ru*)

Федеральный исследовательский центр
«Информатика и управление» РАН, Москва

В работе рассматривается подход к построению графа знаний на основе онтологического представления научных предметных областей в рамках семантической библиотеки. Онтология цифровой семантической библиотеки определяет структуру данных контента библиотеки. Каждому элементу данных, загруженных в библиотеку, можно сопоставить вершину онтологии, определяющую положение элемента данных в онтологии. На основании связей онтологии и связей, определенных на этапе проектирования, можно построить граф данных. На примере онтологии семантической библиотеки LibMeta обсуждается технология формирования графа знаний современных приложений в математике. Обсуждаются задачи наполнения графа, встраивания в граф, извлечения связей и узлов графа.

Ключевые слова: онтология, семантические отношения, предметная область, граф знаний.

Введение

В работе рассматривается подход к построению графа знаний на основе онтологического представления научных предметных областей. На примере математических предметных областей и их приложений в математической физике показывается, как формируются узлы и ребра графа знаний.

Технические проблемы информационного поиска в научных текстах накапливаются как снежный ком, как только выясняется, что одни и те же термины встречаются в контекстах, относящихся к различным предметным областям. Например, термины «ансамблевое моделирование», «построение ансамбля данных» применяется для описания метода построения прогнозов во всех областях исследований, где затруднительно дать точ-

* Работа выполнена при финансовой поддержке РФФ, грант № 22-21-00449.

ные количественные оценки рассматриваемого процесса или явления, а можно только оценить вероятность отдельных результатов моделирования. Хотя сам термин относится к математической статистике, и объяснить его смысл можно только обращаясь к этому разделу математики. И так происходит с большинством математических терминов, что естественно, поскольку математические методы исследования существуют не сами по себе, как одна из вершин человеческого разума, а как средство для моделирования различных явлений природы и общества.

Попытки описать «всю математику» предпринимались научным сообществом многократно [Александрова, 2008]. Целью таких проектов всегда было предоставление научному сообществу наиболее полного описания математических знаний для обучения и работы. Наиболее полное описание терминов и понятий на русском языке в доцифровую эпоху было дано в математической энциклопедии под редакцией Российского советского академика Ивана Матвеевича Виноградова [Виноградов, 1977]. Это уникальное издание, которое переведено и частично оцифровано в новое время. Естественно, что цифровая эпоха породила множество проектов по созданию наиболее полного описания «мировой математики» в интернет. Один из наиболее важных - проект Глобальной цифровой математической библиотеки [GDML, 2016], объединивший в свое время мировое научное сообщество, сформулировал свою цель, как создать, наконец, технологию быстрого доступа к экспертным знаниям для образования и работы. Этот проект, потребовал очень много ресурсов, пока он не осуществлен, но стал отправной точкой для обоснования многочисленных национальных проектов. Краткий обзор этих проектов есть в публикации авторов статьи [Tuchkova, 2022].

Математические знания, как и знания других предметных областей, отражаются в научных публикациях. Особенности научных публикаций известны, они предполагают определенную структуру изложения, требуемую в инструкциях для авторов. В математических текстах присутствуют разделы, связанные с определениями, доказательствами и формулами. Одна и та же формула может использоваться в различных приложениях, например, волновое уравнение встречается в различных задачах математической физики [Владимиров, 2006]. Проблемы работы с формулами во многом решаются, если для их представления использован вариант TeX [Miller, 2002]. Для работы с формулами появились также оцифрованные справочники, например, NIST [NIST], развиваются также инструменты работы с формулами, такие, как MATHEMATICA [MATHEMATICA], MAPLE [MAPLE] и другие. Тем не менее, остаются актуальные задачи обработки публикаций из доцифровой эпохи, когда, например, в русскоязычной литературе не было стандартов оформления во многих журналах.

Отметим, что главная, задача, которая стоит перед разработчиками математических ресурсов, а именно поиск по математическим публикациям и использование формул при поиске, стала рассматриваться только при появлении онтологий, которые позволяют задействовать связи формул с контекстом и тем самым определить их тематическую, предметную принадлежность [Елизаров, 2022].

Построение графов знаний составляет еще один шаг в использовании семантики научных статей для организации поиска. Этой проблеме посвящена данная работа.

Изложение опирается на понятия, связанные с информацией и интеллектуальным анализом данных такие, как знание, извлечение знаний, онтология предметной области, научная предметная область, тезаурус, семантическая цифровая библиотека, информационная потребность пользователя, метод онтологического проектирования и, собственно, граф знаний.

Эти понятия встречаются в многочисленных источниках, но для определенности, перечислим их в контексте этого изложения [Rabunal, 2009].

Определения

Знание – это структурированные данные, извлеченные из различных источников. В частности, рассматриваемые в этой работе словари, энциклопедии и др., [Rabunal, 2009], [Ataeva, 2019] относятся к знаниям (то есть уже структурированным данным).

Извлечение знаний – это получение полезной информации из структурированных и неструктурированных источников [Rabunal, 2009].

Онтология предметной области – это «спецификация концептуализации» [Gruber, 1991] предметной области, где под концептуализацией подразумевается абстрактное представление предметной области. Спецификация определяет набор понятий предметной области и отношений между ними. Понятия предметной области связаны семантическими отношениями (понятия, их свойства и отношения) [Vrandečić, 2009], [Semantic, 2023].

Научная предметная область – в нашем случае математическая предметная область [Tuchkova, 2022], [Serebryakov, 2022].

Тезаурус (информационно-поисковый) – это словарь предметной области, где термины связаны иерархическими и ассоциативными связями [Dextre, 2012]. Статьи тезауруса могут содержать дополнительные поля такие, как ссылки на первоисточники, формулы, ссылки на публикации и др. [Ataeva, 2019].

Семантическая цифровая библиотека – это цифровая библиотека, элементы контента которой связаны иерархическими и ассоциативными отношениями в соответствии с онтологией предметной области [Tuchkova, 2022].

Управление данными – это все манипуляции с данными в цифровой библиотеке, включая накопление и извлечение полезной информации (знаний) [Rabunal, 2009], [Allahyari, 2017].

Информационная потребность пользователя – это субъективное представление пользователя, основанное на совокупности его знаний и представлений о запрашиваемой информации [Gilyarevskij, 1993].

Метод онтологического проектирования – это метод, конечной целью которого является представление предметной области в виде онтологии [Tuchkova, 2022], [Serebryakov, 2022]. Основная последовательность шагов, необходимых для построения онтологии, – это определение цели онтологии, выделение основных верхнеуровневых понятий и их свойств и выделение связей между ними. По сути, после определения целей и задач формирования онтологии выделяются три основных процесса: (1) – это сбор данных из источников и их онтологизация, (2) представление знаний в виде таксономий и (3) – реализация и построение на основе экспертных знаний формальной модели, пригодной для машинного использования в конкретной задаче [Allemang, 2020].

Граф знаний (далее ГЗ) по определению [Paulheim, 2017], [Kroetsch, 2016], [Blumauer, 2014], [Faerber, 2018] представляет собой базу знаний, которая использует модель данных или топологию, структурированную графом, для интеграции данных. При этом нам ближе определение из [Faerber, 2018], где граф знаний определяется как RDF (W3C Resource Description Framework), который состоит из троек (s, p, o). Каждая тройка – это упорядоченное множество RDF терминов: субъект $s \in U \cup B$, предикат $p \in U$, и объект $o \in U \cup B \cup L$. Также RDF термином является URI $u \in U$, «пустой» узел $b \in B$, или литерал $l \in L$. То есть граф знаний наилучшим способом представляется на основе RDF – схем, но не всякое RDF – представление данных трактуется как граф знаний и наоборот.

Графы знаний часто используются для хранения взаимосвязанных описаний сущностей – объектов, событий, ситуаций или абстрактных понятий, а также для кодирования семантики, лежащей в основе используемой терминологии.

Этот вид представления знаний давно используется в логике и искусственном интеллекте [Davis, 1993], например, в семантических сетях [Sowa, 2006] и фреймах [Minsky, 2023]. Совсем недавно он был использован в сообществе Semantic Web с целью создания «сети данных», которая может быть прочитана машинами [Berners-Lee, 2001]. Хотя это видение Semantic Web еще предстоит полностью реализовать, некоторые его части уже реализованы. В частности, концепция связанных данных [Berners-Lee, 2006], [Bizer, 2009] получила распространение, поскольку она облегчает публикацию и связывание данных в Интернете с использованием RDF

[Davis, 1993], [Klyne, 2004], [Cyganiak, 2023], [Brachman, 2004], [Sowa, 2000]. Наилучшее представление ГЗ средствами RDF обеспечивается за счет того, что могут быть выполнены условия интеграции данных, стандартного их представления, повторного использования и связей, для построения графа.

ГЗ сочетают в себе характеристики нескольких парадигм управления данными:

- *База данных*, поскольку данные можно исследовать с помощью структурированных запросов;
- *Графы*, потому что их можно анализировать как любую другую сетевую структуру данных;
- *База знаний*, поскольку они несут формальную семантику, которую можно использовать для интерпретации данных и вывода новых фактов.

Назначение графа знаний состоит в извлечении знаний из источников данных [Wang, 2019], [Wang, 2022] и представлении этих знаний в виде модели (графа), которая определяет сущности и отношения между ними. К числу актуальных задач относятся *построение графа знаний* научной предметной области и *выбор технологии* его построения. Новая вершина тезауруса предметной области появляется, если при анализе (предобработке) первичных данных (статей) выявлен новый термин, обозначающий понятие этой предметной области, но ранее он не встречался (или отсутствовал) в тезаурусе. Задача пополнения задача может трансформироваться в задачу встраивания в граф [Yang, 2021], когда один из узлов заменяется подграфом.

Постановка задачи

В этой работе будем исследовать проблемы достраивания и наполнения графа на примере семантической библиотеки LibMeta и использовать подход онтологического проектирования для построения графа знаний. При достраивании могут появляться как новые связи, так и узлы, но типы связей при этом фиксированы. При наполнении графа не предполагается добавление новых типов узлов, но список типов связей может меняться.

Для реализации технологии построения ГЗ в цифровой библиотеке в рамках определенной тезаурусом предметной области требуется провести исследование целого ряда задач:

- *Определение узлов ГЗ*. При проектировании графа необходимо с одной стороны предусмотреть типы его узлов, а с другой – иметь возможность добавлять новые типы узлов в процессе. Проектирование включает в себя онтологическое моделирование предметной области.

- *Определение связей.* Связи между узлами также определяются при проектировании модели графа, но в процессе подключения новых источников нужно предусмотреть возможность определять новые типы связей.
- *Достраивание ГЗ.* При развитии знаний о предметной области возникает необходимость достраивания графа., но предварительно выполняются задачи онтологического моделирования для описания новых типов узлов графа и возможности расширения модели тезауруса предметной области и связей между узлами. Существует две связанные между собой задачи при достраивании ГЗ: *добавление связей и добавление понятий.*
- *Наполнение ГЗ.* При наличии модели данных, которые соответствуют топологии графа, наполнение графа сводится к извлечению данных и размещению их в соответствии с типами узлов и выявлению связей между ними. Наличие разветвленного тезауруса предметной области ограничивает набор используемых понятий и связей при его наполнении. При наполнении графа обычно используются разные подходы и алгоритмы обработки естественного языка для извлечения данных, среди них такие, как распознавание именованных сущностей и извлечение отношений между ними, нейросетевые подходы, например, семейство алгоритмов word2vec, автоматический перевод и генерация текста и т.д.
- *Поиск по ГЗ:*
 - *поиск в глубину,* когда при фильтрации результатов поиска изначальный поисковый запрос не меняется пользователем, а лишь корректируется тип искомых вершин графа;
 - *поиск в ширину,* когда, в отличие от предыдущей ситуации (при поиске в глубину) для удовлетворения информационной потребности пользователя меняется исходный запрос, могут использоваться автоматически сформированные синонимы или альтернативные запросы.

Поиск по графу относится к задачам интеллектуального анализа данных, который отличается от “популярных” задач искусственного интеллекта тем, что выявляет скрытые взаимосвязи между различными узлами, используя верифицированные экспертами связи тезауруса предметной области.

- *Интеграция данных.* Интеграция данных из разных источников позволяет шире взглянуть на рассматриваемую предметную область, найти новые взаимосвязи и обогатить знания, представленные в графе знаний. Воплощением идеи интеграции данных в графе является

представление данных в парадигме Linked Open Data (LOD) [Schmachtenberg, 2014]. Основные задачи, решаемые на этом этапе, – это устранение проблем, возникающих при объединении данных из разных источников как на уровне данных, так и на уровне модели данных.

Методы и подходы, используемые в LibMeta

Роль онтологий

При разработке семантических библиотек особое внимание уделяют модели данных содержимого библиотеки. При этом контент цифровых библиотек может быть описан различными форматами и представлен различными способами. Библиотека, определяемая с помощью системы LibMeta, рассматривается как хранилище структурированных разнообразных данных с возможностью их интеграции с другими источниками данных и предполагает возможность специфицирования своего контента за счет описания предметной области. В качестве средства формализации выступает онтология контента семантической библиотеки [Serebryakov, 2022].

Решая задачу проектирования онтологии, приходим к необходимости использования метаданных разного уровня:

- метаданные как универсальные понятия структуры данных цифровой библиотеки;
- метаданные как часть описания объектов прикладной области или подмножества прикладной области;
- метаданные прикладной области как таковые.

Онтологическое проектирование LibMeta включает в себя весь процесс создания цифрового описания предметной области (математика) и расширения этого описания при добавлении новых публикаций из смежных и/или ранее не представленных в этой библиотеке областей.

В такой онтологии на верхнем уровне используются понятия, по сути своей относящиеся к высокоуровневым онтологиям и не связанные со спецификой какой-то конкретной предметной области. На втором уровне используются понятия, описывающие предметную область, при этом являющиеся экземплярами классов, определенных на первом уровне, но при этом используемые как определения классов для описания данных третьего уровня уже в конкретной предметной области. Определение предметной области задается тезаурусом [Ataeva, 2019], [Tuchkova, 2022], [Serebryakov, 2022], который содержит основные термины этой предметной области, связанные иерархическими и горизонтальными связями между собой. Содержимое библиотеки задается типами ресурсов, описание которых задает множество допустимых объектов, возможно объединенных в разнообразные коллекции, составляющие вместе с тезаурусом ее контент.

Таким образом, в онтологии цифровой библиотеки можно выделить «системную» часть, так или иначе описывающую структуру самой библиотеки, и «пользовательскую», описывающую структуру данных предметной области, загруженных в библиотеку. Тезаурус является в каком-то смысле верхней частью пользовательской онтологии, полностью и явно видимой пользователю. Еще одной частью онтологии, влияющей на зависимость данных предметной области, является определение связей в пользовательском интерфейсе на этапе проектирования библиотеки.

Роль онтологии в процессе проектирования и эксплуатации цифровой библиотеки можно резюмировать следующим образом:

- на основе онтологии строится база данных библиотеки;
- онтология может использоваться для создания точки доступа LOD к данным библиотеки;
- на основе онтологии строится граф знаний библиотеки.

Граф знаний LibMeta

Как было сказано выше, граф знаний представляет собой базу знаний, которая использует модель данных или топологию, структурированную графом, для интеграции данных. Граф знаний – это граф, свойства вершин которого, дуги (отношения между вершинами) и методы работы с ними определяются приложениями, характерными для обработки «знаний», что в свою очередь определяется семантикой этого понятия и может меняться от приложения к приложению. Граф знаний – это просто граф с определенной семантической интерпретацией вершин и дуг (бинарных отношений).

Онтология цифровой библиотеки определяет *структуру данных* контента библиотеки. Каждому элементу данных, загруженных в библиотеку, можно сопоставить вершину онтологии, определяющую положение элемента данных в онтологии. На основании связей онтологии и связей, определенных на этапе проектирования, можно построить граф данных. Данные предметной области можно представить как граф, структура которого определяется онтологией, вершины (статьи) – экземпляры элементов онтологии, связи – связи тезауруса. Это граф знаний цифровой библиотеки. Используются следующие типы связей элементов графа знаний LibMeta:

- иерархические связи (род, вид);
- горизонтальные связи (синоним, ассоциация);
- семантические связи между ресурсами определяемые при описании предметной области;
- связи определяемые на уровне пересечения метаданных (смотри также, совпадает с).

Каждый элемент графа состоит из двух частей – собственно данных (в случае научных приложений, как правило, статьи) и метаданных. К числу метаданных относятся, в частности:

- индексы рубрикаторов, для предметной области Математика это MSC, УДК, ГРНТИ и т.д.;
- ключевые слова;
- авторы;
- упоминаемые в статье теоремы, формулы, имена, понятия тезауруса.

Наполнение графа

Перед размещением нового объекта в граф происходит его преобразование, которая включает формирование метаданных, а именно подготовка текстов к включению в уже подготовленную инфраструктуру графа знаний:

- препроцессинг и очистка этих данных;
- разметка структурных элементов в соответствии с онтологией документа;
- разметка документа по терминам тезауруса;
- выделение связей в документе между терминами предметной области/ключевыми словами;
- сжатие текста: извлечение метаданных, извлечение основных структурных элементов (аннотация, теорема, формула, ...), извлечение списка терминов и ключевых слов, извлечение связей, извлечение формул (только тех формул, которые относятся к «важным» фрагментам текста) и связанных с ними терминов.

Сжатый текст используется для установления связей с тезаурусом и с имеющимися объектами, уже погруженными в предметную область. Пересечения по метаданным (ключевые слова, классификаторы, понятия тезауруса) между новыми и объектами предметной области позволяют оценить степень соответствия предметной области, степень семантической близости между объектами и выделить новые связи и термины, которые не встречались ранее.

Типы основных связей, которые используются для выявления новых связей между элементами графа знаний *объект – объект, понятие – понятие, объект – понятие – объект, понятие – объект – понятие, классификатор – понятие – классификатор, понятие – классификатор – понятие, объект – классификатор – объект, классификатор – объект – классификатор*.

Все эти связи задействованы при формировании интерфейса пользователя и дают возможность как дополнения и редактирования, так и навигации по ГЗ в системе.

В качестве объекта могут выступать разные типы узлов такие, как публикации, авторы, формулы, теоремы и т.д. В качестве элементов классификатора могут использоваться элементы разных классификаторов. В качестве понятий на концах связей могут также выступать понятия разных тезаурусов (частей тезауруса, кустов тезауруса).

На основе такого подхода в предметной области «математика» были выделены дополнительные связи между (*MSC – УДК*), (*Формулы – MSC*), (*Формулы – УДК*), (*Мат.Энциклопедия – MSC*), (*Мат.Энциклопедия – ОДУ*), (*Мат.Энциклопедия – Публикации*), (*Мат.Энциклопедия – Спец.функции*), (*Мат.Энциклопедия – Формулы*), (*Формулы – Ключевые слова*), (*Публикации – Мат.Энциклопедия*), (*Публикации – Спец.Функции*).

При масштабировании поиска мы предлагаем пользователю связанные объекты, оцениваемые по степени семантической близости, где семантическая близость определяется количеством связей, которые устанавливаются между объектами на основе перечисленных типов связей и их комбинаций.

Достраивание графа

Научные знания, представленные публикациями, для размещения в семантическую библиотеку должны подвергнуться предварительной обработке. Варианты предобработки зависят от исходного материала и могут варьироваться в зависимости от степени структурированности статей. При необходимости, могут быть задействованы алгоритмы машинного обучения.

Далее, если предобработка показала наличие признаков принадлежности *данных* определенной *предметной области*, то в онтологию предметной области они и помещаются. Иначе возникает вопрос о новой (например, смежной) предметной области и появлении новой вершины графа знаний в онтологии.

Для классических предметных областей (таких как математика) это, как правило, термин из новых задач или приложений в междисциплинарных исследованиях. Один из таких характерных примеров – это приложения в уравнениях математической физики. Каждая новая задача связана с экспериментами и/или физическими процессами, для которых сформулированы определенные краевые условия для известных уравнений. Множество исследований, связанных с новой постановкой задачи, отраженных в публикациях, будет содержать не только термины, которые пересекаются с классическим описанием уравнения, но и новые термины, связанные с основной предметной областью. Таким образом узел в междисциплинарном случае превращается в подграф. Как пример можно рассмотреть узел «смешанные уравнения теории упругости», который связан со узлами «уравнения математической физики» из ГЗ «математика» и разделов

«теория упругости» из ГЗ «Механика сплошной среды», которые раньше не имели связей. Это узел узкой предметной области, данные которого встраиваются в наш ГЗ «Математика» [Serebryakov, 2022]. В этом узле появляются связанные термины тезауруса образующие подграф ГЗ «Дифференциальные уравнения линейной теории упругости»: *Смешанные уравнения теории упругости, Основные уравнения линейной теории упругости, Дифференциальные уравнения линейной теории упругости, Теория упругости, Решения задач теории упругости, Механика сплошной среды, Уравнения статики в объеме, Уравнения равновесия, Определение линейного тензора деформации через вектор перемещения, Закон состояния линейно-упругого тела, Краевые задачи линейной теории упругости, Смешанные уравнения УЧП.*

Заключение и выводы

Основная цель библиотеки состоит в интеграции математических знаний как в ретроспективном направлении (использование научного наследия классиков математики, которое отсутствует в энциклопедии в силу ограниченного периода ее создания), так и в перспективном, путем добавления публикаций из различных актуальных новых предметных областей математики и приложений (также отсутствующих в энциклопедии) для дальнейшего использования LibMeta в образовательных и научных проектах.

Граф знаний может быть построен автоматически, программным путем и знания получены автоматически. Реализация идеи «автоматического» извлечения научных знаний базируется на том, что если все результаты размещены в базах данных и цифровых библиотеках, то при обращении к ним пользователь может получить обзор современных исследований и некоторые знания на заданную тему. Для математических знаний аналогичная идея была заложена еще задолго до появления технологии построения графа знаний как способ обеспечения доступа исследователей к научным результатам во Всемирной математической библиотеке (WDML [Iop, 2017]). При этом про математические предметные области известно, что они до сих пор недостаточно семантически описаны. При продолжающемся росте публикаций, основная идея извлечения таких знаний остается актуальной.

Дальнейшие исследования посвящены применению алгоритмов искусственного интеллекта для извлечения знаний на русском языке из предметной области на основе энциклопедии и наполнения библиотеки.

Список литературы

- [Александрова, 2008] Александрова Н.В. История математических терминов, понятий, обозначений. Словарь – справочник. Издание третье, исправленное. – М.: Изд-во ЛКИ, 2008. – 248 с.
- [Виноградов, 1977] Виноградов И.М. (ред.). Математическая энциклопедия (в 5 т.). – М.: Советская энциклопедия. 1977–1985
- [GDML, 2016] AMS: Специальная сессия по математической информации в цифровую эпоху науки, Сизтл, 6–7 января 2016 г. – http://jointmathematicsmeetings.org/meetings/national/jmm2016/2181_program_ss65.html.
- [Владимиров, 2006] Владимиров В.С. Что такое математическая физика? Препринт. Математический институт им. В.А. Стеклова РАН. – М.: МИАН, 2006. – 20 с.
- [Miller, 2002] Miller B., Youssef A. Technical Aspects of the Digital Library of Mathematical Functions // *Annals of Mathematics and Artificial Intelligence*. – 2003. – 38. – P. 121-136. – <https://doi.org/10.1023/A:1022967814992>.
- [NIST] <https://dlmf.nist.gov/>.
- [MATHEMATICA] <http://www.wolfram.com>.
- [MAPLE] <https://www.maplesoft.com/products/Maple/>.
- [Елизаров, 2022] Елизаров А.М., Кириллович А.В., Липачёв Е.К., Невзорова О.А. Онтология математического знания ONTOMATH^{PRO} // Доклады российской академии наук. Математика, информатика, процессы управления. – 2022. – Т. 507, № 1. – С. 29-35.
- [Rabunal, 2009] Rabunal J.R., Dorado J., Sierra A.P. Encyclopedia of Artificial Intelligence (IGI Global,2009). – <https://doi.org/10.4018/978-1-59904-849-9>.
- [Ataeva, 2019] Ataeva O., Serebryakov V., and Sinelnikova E. Thesaurus and Ontology Building for Semantic Library Based on Mathematical Encyclopedia, (DAMDID/RCDL 2019), Kazan, Russia, October 15-18, 2019 // CEUR Workshop Proceedings, 148#157. – 2019.
- [Gruber, 1991] Gruber T.R. The role of common ontology in achieving sharable, reusable knowledge bases / In: J.A. Allen, R. Fikes, and E. Sandewell, ed. // *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning (KR'91)*. – 1991. – P. 601-602. – <https://doi.org/10.5555/3087158.3087222>.
- [Vrandecic, 2009] Vrandecic D. Ontology Evaluation, In *Handbook on Ontologies // International Handbooks on Information Systems*, edited by S. Staab, R. Studer. – 2009. – P. 293-313. – https://doi.org/10.1007/978-3-540-92673-3_13.
- [Semantic, 2023] Semantic Web. – <https://www.w3.org/standards/semanticweb>. Accessed 2023.
- [Tuchkova, 2022] Ataeva O., Serebryakov V., and Tuchkova N. Development of the semantic space 'Mathematics' by integrating a subspace of its applied area // *Lobachevskii J. of Mathematics*. – 2022. – 43 (12). – P. 3435-3446. – <https://doi.org/10.1134/S1995080222150069>.
- [Serebryakov, 2022] Ataeva O., Serebryakov V., and Tuchkova N. Creating the Applied Subject Area Ontology by Means of the Content of the Digital Semantic Library // *Lobachevskii J. of Mathematics*. – 2022. – 43 (7). – P. 1795-1804. – <https://doi.org/10.1134/S1995080222100043>.

- [Dextre, 2012] Dextre Clarke S.G., Zeng M.L. Standard Spotlight: From ISO 2788 to ISO 25964: The Evolution of Thesaurus Standards Towards Interoperability and Data Modeling // *Information Standards Quarterly* (ISQ). – 2012. – 24 (1). – P. 20-26. – <https://doi.org/10.3789/isqv24n1.2012.04>.
- [Allahyari, 2017] Allahyari M., et al., A brief survey of text mining: Classification, clustering and extraction technique // *arXiv preprint arXiv:1707.02919*. – 2017. – <https://doi.org/10.48550/arXiv.1707.02919>.
- [Gilyarevskij, 1993] Gilyarevskij R.S., Markusova V.A., Chernyj A. Nauchnye kommunikacii i problemy informacionnoj potrebnosti // *Nauchno-tehnicheskaya informaciya*. – 1993. – Ser. 1 (9). – P. 1-7
- [Allemang, 2020] Allemang D., Hendler J., Gandon F. *Semantic Web for the Working Ontologist: Effective Modeling for Linked Data, RDFS, and OWL* (Association for Computing Machinery, 2020).
- [Paulheim, 2017] Paulheim H. Knowledge Graph Re#nement: A Survey of Approaches and Evaluation Methods // *Semantic Web*. – 2017. – 8 (3). – P. 489-508. – <https://doi.org/10.3233/SW-160218>.
- [Kroetsch, 2016] Kroetsch M. and Weikum G. *Journal of Web Semantics: Special Issue on Knowledge Graphs*. – 2016. – <http://www.websemanticsjournal.org/index.php/ps/announcement/view/19>. Accessed 2023.
- [Blumauer, 2014] Blumauer A. From Taxonomies over Ontologies to Knowledge Graphs. – 2014. – <https://blog.semanticweb.at/2014/07/15/from-taxonomies-over-ontologiesto-knowledge-graphs>. Accessed 2023.
- [Faerber, 2018] Faerber M., Bartscherer F., Menne C., Rettinger A. Linked Data Quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO // *Semantic Web*. – 2018. – 9 (1). – P. 77-129. – <https://doi.org/10.3233/SW-170275>.
- [Davis, 1993] Davis R., Shrobe H., and Szolovits P. What is a knowledge representation? // *AI Magazine*. – 1993. – 14 (1). – P. 17-33.
- [Sowa, 2006] Sowa J.J.F. *Semantic Networks, Encyclopedia of Artificial Intelligence* (Wiley,1992,2006).
- [Minsky, 2023] Minsky M. A framework for representing knowledge // *MIT-AI Laboratory Memo*. – 1974. – 306. – <https://hdl.handle.net/1721.1/6089>. Accessed 2023.
- [Berners-Lee, 2001] Berners-Lee T., Hendler J., and Lassila O. *The Semantic Web*. – 2001. – <http://www.scientificamerican.com/article/the-semantic-web/>. Accessed 2023.
- [Berners-Lee, 2006] Berners-Lee T. *Linked Data–Design Issues*. – 2006. – <http://www.w3.org/DesignIssues/LinkedData.html>. Accessed 2023.
- [Bizer, 2009] Bizer C., Heath T., and Berners-Lee T. *Linked data the story so far, International Journal on Semantic Web and Information Systems*. – 2009. – 5 (3). – P. 1-22. – <https://doi.org/10.4018/jswis.2009081901>.
- [Klyne, 2004] Klyne G. and Carroll J.J. *Resource Description Framework (RDF): Concepts and Abstract Syntax*. – 2004. – <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>. Accessed 2023.
- [Cyganiak, 2023] Cyganiak R., Wood D., and Lanthaler M. *Resource Description Framework (RDF): Concepts and Abstract Syntax*. – 2014. – <https://www.w3.org/TR/rdf11-concepts/>. Accessed 2023.

- [Brachman, 2004]** Brachman R. and Levesque H. Knowledge Representation and Reasoning (San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2004). – <https://doi.org/10.1146/annurev.cs.01.060186.001351>.
- [Sowa, 2000]** Sowa J.J.F. Knowledge Representation: Logical, Philosophical and Computational Foundations (Pacific Grove, CA, USA: Brooks/Cole Publishing Co., 2000).
- [Wang, 2019]** Wang P., Jiang H., Xu J., and Zhang Q. Knowledge graph construction and applications for web search and beyond // Data Intelligence. – 2019. – 1 (4). – P. 333-349. – https://doi.org/10.1162/dint_a_00019.
- [Wang, 2022]** Wang J. Math-KG: Construction and Applications of Mathematical Knowledge Graph // arXiv:2205.03772. – 2022. – <https://doi.org/10.48550/arXiv.2205.03772>.
- [Yang, 2021]** Yang H., Zhang L., Wang B., Yao T., Liu J. Cycle or Minkowski: Which is More Appropriate for Knowledge Graph Embedding? // Proceedings of the 30th ACM International Conference on Information & Knowledge Management. – 2021. – P. 2301-2310. – <https://doi.org/10.1145/3459637.3482245>.
- [Schmachtenberg, 2014]** Schmachtenberg M., Bizer C., Paulheim H. State of the LOD Cloud, University of Mannheim // Data and Web Science Group. – 2014. – 30.
- [Ion, 2017]** Ion P.D.F., Watt S.M. The Global Digital Mathematics Library and the International Mathematical Knowledge / In: Geuvers, H., England, M., Hasan, O., Rabe, F., Teschke, O. (eds) // Intelligent Computer Mathematics. CICM 2017. Lecture Notes in Computer Science, Springer, Cham. – 2017. – 10383. – P. 56-69. – https://doi.org/10.1007/978-3-319-62075-6_5.

К ВОПРОСУ О РАЗРАБОТКЕ ОНТОЛОГИЙ НАУЧНОЙ И УЧЕБНОЙ РАБОТЫ*

Т.А. Гаврилова (*gavrilova@gsom.spbu.ru*)

А.В. Кузнецова (*menshikova@gsom.spbu.ru*)

И.А. Лещева (*leshcheva@gsom.spbu.ru*)

Санкт-Петербургский государственный университет,
Санкт-Петербург

Работа посвящена онтологическому инжинирингу учебной и научной работы преподавателей вузов. Разработанные онтологии легли в основу карт знаний, служащих для упрощения процессов поиска и обработки информации и позволяющих получить наглядный «портрет» конкретного преподавателя или кафедры вуза.

Ключевые слова: онтологии, карты знаний, управление знаниями.

Введение

Онтологии как концептуальные модели предметной области являются одним из наиболее перспективных подходов к формированию баз знаний и графов знаний. В информационных системах онтологии служат системообразующим теоретическим и методическим каркасом для моделирования сложных предметных областей.

К таким сложным областям относятся учебная и научная деятельность преподавателей вузов. Их моделирование необходимо при создании корпоративных систем, при поиске нужной информации, при инжиниринге процессов и других видов автоматизации. При этом современные системы информационного менеджмента частично помогают лишь в выполнении рутинных процедур, в то время как большую часть времени работы с такими системами занимает поиск информации и ее структурирование.

В последнее время многие исследователи делают акцент на преимуществах компрессии информации и ее визуализации [Erppler, 2019], [Kudravnsev et al., 2017]. При этом обнаружилась слабая взаимосвязь между потребностями университетов и новыми технологиями в области инженерии знаний и визуального онтологического инжиниринга. Модели и

* Работа выполнена при поддержке РФФИ (проект № 23-21-00168, <https://rscf.ru/project/23-21-00168/>).

методы визуализации знаний пока недостаточно зрелы для решения практических задач управления знаниями и информационного менеджмента. Например, информационные перегрузки и разобщенность создают множественные преграды для поиска партнеров и контрагентов организации научных исследований и разработок.

Данный доклад посвящен дизайну онтологий учебной и научной работы преподавателей вузов с последующей разработкой карт знаний для упрощения процессов поиска и обработки знаний. В учебной и научной работе мы делаем акцент на создание таксономий, позволяющих описать деятельность преподавателя с точки зрения его знаний и компетенций.

Онтология учебной работы базируется на визуализации активности по осуществлению аудиторной и внеаудиторной нагрузке на образовательных программах различного уровня и формата. Онтология научной работы включает описание проведения исследований по проектам, подготовки научных публикаций, выступлений на конференциях, а также научной экспертной работы. Разработанные онтологии легли в основу карт знаний, позволяющих получить наглядный «портрет» конкретного преподавателя или кафедры вуза.

Карты знаний представляют специальные инструменты для анализа информации, которые могут улучшить процесс принятия управленческих решений и снизить когнитивную нагрузку работников умственного труда. Такие диаграммы способствуют эффективному поиску знаний и принятию решений, указывая, ЧТО известно сотрудникам, ГДЕ эти знания находятся, КТО является носителем знаний. Так карта знаний устанавливает связь между предметом, носителем и местонахождением знаний.

Наш подход апробирован на одном из факультетов крупного классического университета Санкт-Петербурга. Разработанные карты знаний наглядно представляют информацию о том, какими знаниями обладает кафедра, факультет или вуз, и кто из сотрудников является владельцем этих знаний. Набор карт знаний позволяет формировать визуально интеллектуальный ландшафт научного сообщества, сравнивать научные потенциалы разных научных коллективов, а также находить коллег для сотрудничества и междисциплинарных исследований.

1. Об онтологическом инжиниринге

В нашей стране и за рубежом накоплен большой опыт методологических и технологических исследований по практическому дизайну и формированию онтологий. Онтологическому инжинирингу уже более 30 лет, он был начат в пионерских работах Грубера и Ушхольда [Gruber, 1993], [Uschold, 1998] и активно развивается до настоящего времени в публикациях [Patel et al., 2020], [Krieg-Brückner et al., 2021].

Активно в этой области работают и российские исследователи [Боргест и др., 2019], [Ворожцова и др., 2019]. Разработка онтологий представляет ядро современной инженерии знаний [Гаврилова и др., 2016], [Загорюлько, 2020].

Однако разработка практических онтологий в производстве, проектировании и менеджменте, остаётся скорее на уровне «искусства». Проектирование онтологий довольно слабо освещено в литературе по онтологическому инжинирингу, большинство авторов сосредоточили свои усилия на формализации и моделировании как таковом [Poveda-Villalón et al., 2022]. Известные методологии и технологии также ориентированы на организационные и технологические аспекты, не касаясь проблемы формирования понятий, их уровня абстракции, баланса отношений и других вопросов семантики.

В то время как онтологии стали стандартом де-факто в области разработки баз знаний, процессы извлечения и особенно структурирования знаний по-прежнему остаются некоторым «белым пятном» в современной литературе по инженерии знаний. Можно сказать, что в семиотическом треугольнике «синтаксис – семантика – прагматика» пока главенствуют синтаксис и прагматика.

Новый импульс онтологический инжиниринг получил в связи с мощным «взлетом» графов знаний на основе онтологий [Noy et al., 2019] и автоматизированным построением онтологий [AI-Aswadi et al., 2020].

В данном проекте мы использовали классические методики, доработав их под свои задачи.

2. Онтология учебной работы в области социальных наук

Существует множество дидактических онтологий предметных областей для преподавания. Несмотря на большой интерес, и попыткам создать онтологию учебной работы [Ulfa et al., 2020], [Куликов, 2020], [Хабаров и др., 2020], практически не существует общепринятой единой онтологии преподавания.

В данном проекте за основу взята учебная работа по преподаванию социальных наук (менеджмент, экономика, государственное управление), так как в естественных технических науках имеется своя специфика обучения.

Основными видами учебной работы, включенными в онтологию стали:

- 1) аудиторная нагрузка (чтение лекций и ведение семинаров);
- 2) внеаудиторная нагрузка (руководство курсовыми и выпускными работами);
- 3) подготовка и издание учебно-методических материалов и различные виды распространения педагогического опыта.



Рис. 1. Онтология учебной работы

При этом учитывались различия уровней программ подготовки (бакалавр/магистр, аспирантура, дополнительные образовательные программы и корпоративное обучение) и типы участия в разработке содержания курсов (обновление курса, разработка нового курса, разработка новой программы или бизнес-симуляции).

Эскизы онтологий были разработаны на основе инструментов майнд-мэппинга и концептуального моделирования, а затем перенесены в систему PROTÉGÉ. Методология перехода от интеллект-карт к онтологии обсуждалась ранее в работе [Гаврилова и др., 2020]. Рис. 1 представляет фрагмент онтологии учебной работы. На рис. 2 представлена одна из карт портфолио кафедры для учебной работы, дающая представление о курсах, читаемых преподавателями одной из кафедр факультета.

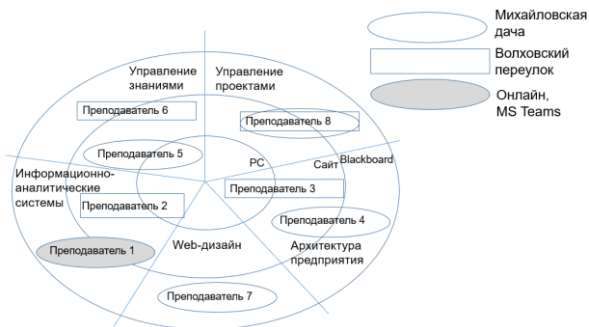


Рис. 2. Карта знаний о преподаваемых на кафедре дисциплинах

Карта знаний о преподаваемых на кафедре дисциплинах уточняет информацию о месте чтения лекций в одном из двух кампусов или онлайн, а также о способе хранения материалов по курсу (на персональном компьютере, на сайте или на платформе электронного обучения Blackboard).

3. Онтология научной работы

Научная работа как таковая в большой степени зависит от предметной области, и очевидно, что при работе работа по подготовке статьи по философии и для статьи по микробиологии, виды и формы научной активности будут существенно различаться. Однако общие формы научного познания схожи и всегда подразумевают творческую активность по увеличению суммы научных знаний о человеке, природе и обществе. С прагматической точки зрения можно говорить о различных видах фундаментальных и прикладных исследований, а также о разработке новых теорий, технологий, методик и рекомендаций.

Таким образом, встает вопрос о создании онтологии научных исследований. Существуют различные онтологии научной работы, например, Core Ontology for Scientific Research Activities (COSRA) [Campos et al., 2019], ЭМПИРИОН [Беглер и др., 2020], DINGO – онтология для научных проектов [Chialva et al., 2020]. Однако все они описывают природу и структуру эмпирических данных научных исследований.

Так, предметно-независимая онтология для описания исследовательской активности, COSRA основана на концептуальной модели наблюдений и измерений [ISO 19156, 2011], она подходит для тех областей, где стандарт ISO 19156 работает, что не всегда применимо для «мягких» предметных областей. Еще одной сильной стороной онтологии для научно-исследовательской деятельности COSRA является то, что она имеет сложные связи с единой фундаментальной онтологией высшего уровня (Upper Fundamental Ontology, UFO).

В данной задаче мы ставим более узкую и прагматическую цель – описать научную деятельность преподавателей, поэтому опираемся на традиционные формы исследовательской активности (например, проекты, выступления на конференциях и подготовка публикаций).

Онтология научной работы преподавателя включает

- 1) сведения о видах проектов (проекты с внешним финансированием от исследовательских фондов, проекты с внешним финансированием от компаний, проекты с внутренним финансированием от университетов);
- 2) роли в заявках на гранты и проекты, и в самих проектах (руководитель, исполнитель, ведущий исполнитель);
- 3) типы результатов проекта (например, теоретические модели, аналитические отчеты, методология исследования, методология управления и т.д.).

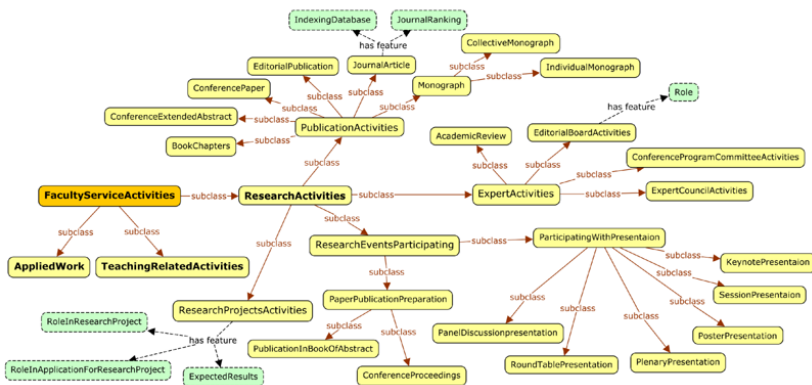


Рис. 3. Онтология научной работы

Разработанная онтология была положена в основу разработки научно-го портфолио. Портфолио содержит набор карт знаний, например, карта научных интересов, карта публикационной активности, карта грантовой активности и др. На рис. 4 – пример карты, дающей представление об основных областях научной активности преподавателей факультета.

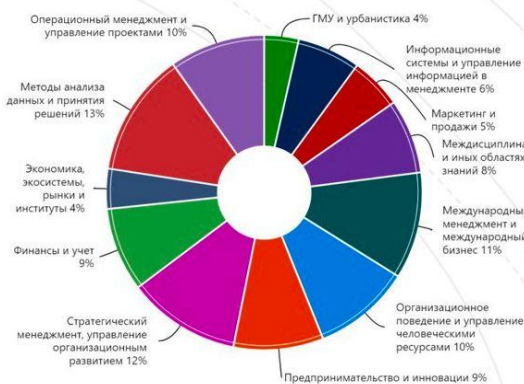


Рис. 4. Области научной активности преподавателей факультета

Представленная на рис. 4 карта отражает и фиксирует имеющиеся у преподавателей знания в различных областях научно-исследовательской деятельности, в которых специализируются научно-педагогические работники факультета, тем самым упрощая организацию процессов обмена и передачи знаний на факультете.

Заключение

Современные образовательные и научно-исследовательские организации функционируют в гигантских информационных пространствах. В таких условиях становится критичным поиск релевантной информации. Реальные активы знаний компании, а также их доступность и структурированность, становятся все более важными.

В данном докладе рассмотрены вопросы практической разработки карт знаний вузов на основе онтологий. Разработанные онтологии позволяют строить карты знаний, отражающие различные проекции интеллектуального капитала вуза и его подразделений на системной и хорошо структурированной основе. Эти проекции отражают коллективные и индивидуальные многомерные «портреты» знаний профессорско-преподавательского состава и научных подразделений.

Благодарности. Авторы выражают благодарность доценту Д.В. Кудрявцеву, который был инициатором данного проекта и активно участвовал в разработке дизайна исследования, а также Ольге Алкановой за организацию опроса.

Список литературы

- [Боргест, 2022] Боргест Н.М. Проблемы разработки и развития онтологии науки: анализ классификаций // Онтология проектирования. – 2022. –12 (3 (45)). – С. 278-98.
- [Беглер и др., 2020] Беглер А.М., Кудрявцев Д.В., Гаврилова Т.А. О применении онтологий для интеграции данных научных исследований // Труды 18 национальной конференции по искусственному интеллекту с международным участием КИИ-2020 (в рамках I Национального Конгресса по когнитивным исследованиям, искусственному интеллекту и нейроинформатике). – М., 2020. – С. 3-12.
- [Ворожцова и др., 2019] Ворожцова Т.Н., Макагонова Н.Н., Массель Л.В. Онтологический подход к проектированию базы данных для оценки влияния энергии на окружающую среду // Информационные и математические технологии в науке и управлении. – 2019. – 3, 15.
- [Гаврилова и др., 2016] Гаврилова Т.А., Кудрявцев Д.В., Муромцев Д.И. Инженерия знаний. Модели и методы. 2016. – СПб.: Лань.
- [Гаврилова и др. 2020] Гаврилова Т.А., Страхович Э.В. Визуально-аналитическое мышление и интеллект-карты в онтологическом инжиниринге // Онтология проектирования. – 2020. – 1. – С. 87-99.
- [Грибова и др., 2019] Грибова В.В., Шалфеева Е.А. Онтология диагностики процессов // Онтология проектирования. – 2019. – 9.4, 34. – С. 449-461.
- [Загорулько, 2020] Загорулько Ю.А., Загорулько Г.Б. Искусственный интеллект. Инженерия знаний. – 2020.

- [Хабаров и др., 2020] Хабаров В. И., Волежанина И. С. Онтологические форматы профессионального образования в области транспорта // Современный ученый. – 2020. – 2. – С. 95-101.
- [Куликов, 2020] Куликов Н.В. Исследование онтологий и возможных областей их применения // Методы и средства обработки и хранения информации. – 2020. – С. 16-23.
- [Al-Aswadi et al., 2020] Al-Aswadi F.N., Chan H.Y., Gan K.H., Automatic ontology construction from text: a review from shallow to deep learning trend // Artificial Intelligence Review. – 2020. – 53. – P. 3901-3928.
- [Eppler, 2019] Eppler M. Big Data meets Big Knowledge: Design Principles for the Combination of Visual Analytics and Knowledge Visualization in Collaborative Business Contexts. – 2019.
- [Chialva et al., 2020] Chialva D., Mugabushaka A. M. (2020). Dingo: an ontology for projects and grants linked data // ADBIS, TPD and EDA 2020 Common Workshops and Doctoral Consortium: International Workshops: DOING, MADEISD, SKG, BBIGAP, SIMPDA, AIMinScience 2020 and Doctoral Consortium, Lyon, France, August 25–27, 2020, Proceedings. 24, 183-194). – Springer International Publishing, 2020.
- [Gruber, 1993] Gruber T. A translation approach to portable ontology specifications // Knowledge acquisition. – 1993. – 5, 2. – P. 199-220.
- [ISO 19156, 2011] ISO 19156:2011: Geographic information – Observations and measurements. International Standard. 2011.
- [Krieg-Brückner et al., 2021] Krieg-Brückner B., Mossakowski T., Codescu M. Generic Ontology Design Patterns: Roles and Change Over Time // Advances in Pattern-Based Ontology Engineering. – 2021. – 51, 25.
- [Kudravytsev et al., 2017] Kudravytsev D., Gavrilova T. From anarchy to system: A novel classification of visual knowledge codification techniques // Knowledge and Process Management. – 2017. – 24 (1). – P. 3-13.
- [Noy et al., 2019] Noy N., Gao Y., Jain A., Narayanan A., Patterson A., Taylor J. Industry-scale Knowledge Graphs: Lessons and Challenges: Five diverse technology companies show how it's done // Queue. – 2019. Apr 1;17 (2). – P. 48-75.
- [Patel et al., 2020] Patel A., Jain S. A partition based framework for large scale ontology matching // Recent Patents on Engineering. – 2020. – 14 (3). – P. 488-501.
- [Ulfa et al., 2020] Ulfa S., Lasfeto D.B., Kurniawan C. Modelling the learner model based ontology in adaptive learning environment // Journal of Disruptive Learning Innovation (JODLI). – 2020. – 1 (1). – P. 34-45.
- [Uschold, 1998] Uschold M., King M., Moralee S., Zorgios Y. The enterprise ontology // The knowledge engineering review. – 1998. – 13 (1). – P. 31-89.
- [Campos et al., 2019] Campos P.M.C., Reginato C.C., Almeida, J.P.A. Towards a Core Ontology for Scientific Research Activities // Lecture Notes in Computer Science. – 2019. – 11787. – P. 3-12.
- [Poveda-Villalón et al., 2022] Poveda-Villalón M., Fernández-Izquierdo A., Fernández-López M., García-Castro R. Lot: An industrial oriented ontology engineering framework // Engineering Applications of Artificial Intelligence. – 2022. – Vol. 111.

ОПИСАНИЕ БАЗОВЫХ ЭЛЕМЕНТОВ МЕТАГРАФОВОЙ МОДЕЛИ С ИСПОЛЬЗОВАНИЕМ ГРАНУЛЯРНО-ПРОТОГРАФОВОГО ПОДХОДА*

Ю.Е. Гапанюк (*gapyu@bmstu.ru*)

Московский государственный технический университет
им. Н.Э. Баумана, Москва

Проблема, которая решается в данной работе, связана с тем, что в предыдущих вариантах формализации метаграфовой модели не рассматривался единый базис, на основе которого можно было бы описать такие различные по своей природе элементы модели, как вершины, метавершины и ребра. В качестве такого базиса в данной статье предложена гранулярно-протографовая модель. Рассмотрены предпосылки объединения гранулярного подхода и протографовой модели. Предложены основные элементы гранулярно-протографовой модели: фасеты гранул и гранулярные элементы протографа. Также предложены основные операции над элементами гранулярно-протографовой модели: операция сигнатурного прилегания и операция организации гранул. Рассмотрены основные положения аннотируемой метаграфовой модели. Предложено представление основных элементов метаграфовой модели (вершин, метавершин, ребер) с помощью элементов гранулярно-протографовой модели и введенных над ними операций. Рассмотрен комплексный пример описания метаграфа с помощью элементов гранулярно-протографовой модели.

Ключевые слова: метаграф, метавершина, ребро, протограф, гранулярно-протографовая модель, фасет гранулы, гранулярный элемент протографа.

Введение

В работе [Гапанюк, 2022] было предложено использовать протографовый подход как основу для конструирования метаграфовой модели. Эта статья развивает идеи работы [Гапанюк, 2022]. В данной статье протографовый подход развивается на основе идей гранулярных вычислений [Pedrycz et al., 2011].

* Работа выполнена при финансовой поддержке РФФИ (проект №20-07-00770).

В работе [Гапанюк, 2022] была предложена формализация операции прилегания элементов протографа друг к другу. В данной статье идея прилегания расширяется до операции сигнатурного прилегания со знаком на основе специализированных структур – фасетов гранул.

Также вводится понятие гранулярного элемента протографа, который, в отличие от традиционного элемента протографа, может иметь внутреннюю структуру. Для формирования внутренней структуры вводится операция организации гранул.

Предложенная гранулярно-протографовая модель позволяет на базе гранулярных элементов протографа построить унифицированное описание таких различных по своей природе элементов метаграфовой модели как вершины, метавершины и ребра.

Основной целью предлагаемого подхода является усовершенствование способов гранулярного представления знаний [Stalder at al., 2021]. Наиболее распространенными подходами к гранулярному представлению знаний являются гранулярные онтологии [Polkowski, 2023], Liu at al., 2014], а в последнее время также и гранулярные графы знаний [Liu at al., 2022]. При этом детальная структура элементов онтологии (гранул), как правило, авторами не рассматривается, рассматриваются лишь нечеткие отношения между элементами гранулярной модели.

В отличие от существующих подходов, предложенная гранулярно-протографовая модель позволяет детально рассмотреть внутреннюю структуру гранулы и способы взаимодействия между гранулами.

1. Предпосылки объединения гранулярного подхода и протографовой модели

Понятие протографа вводится в работах [Кручинин, 2017а] и [Кручинин, 2017б]. В соответствии с [Кручинин, 2017б]: «протограф P задается множеством элементов $\{p_i\}$, $i = 1, n$ и матрицей соседства (смежности) $M_{n \times n}$, состоящей из 0 и 1, где 1 означает соседство (смежность) элемента a элементу b ». Таким образом, как отмечается в [Сухобоков, 2022]: «протограф можно рассматривать как граф, не имеющий ребер; роль ребер выполняет прилегание вершин друг к другу». Необходимо отметить, что поскольку определение протографа основано на матрице смежности, то речь идет именно о попарном прилегании элементов протографа друг к другу.

Также в [Кручинин, 2017б] отмечается, что прилегание может быть направленным (ориентированным), но данный подход не будет использоваться в предлагаемой модели.

В [Гапанюк, 2022] была введена двухместная операция « $*$ » прилегания элементов протографа:

$$p_1 * p_2. \quad (1.1)$$

Исходя из определения неориентированного протографа, операция прилегания является коммутативной (так как порядок объединения двух соседних элементов протографа не важен) и ассоциативной (так как порядок объединения нескольких соседних элементов протографа также не важен).

В данной работе предлагается рассмотреть элемент протографа как *гранулу*, концепция гранулы является основополагающей в гранулярных вычислениях [Pedrycz et al., 2011].

С точки зрения работ [Кручинин, 2017a] и [Кручинин, 2017b], *элемент протографа* является *атомарным*. Согласно словарю [Словарь Merriam-Webster] гранулу можно определить как «маленькую частицу, в особенности как одну из многочисленных частиц, образующих большую единицу». Таким образом, существенным свойством гранулы, как и элемента протографа, также является *атомарность*.

Согласно [Yao et al., 2013], гранулы «могут быть дополнительно разложены на более мелкие гранулы, называемые субгранулами. В зависимости от сложности, уровня абстракции и размера, субгранулы могут быть расположены на разных уровнях. «*Универсум*» определяется на уровне наиболее общей (грубой) гранулы. Гранулы на самом низком уровне состоят из элементов конкретной используемой модели». Таким образом, существенным свойством гранулярности является *иерархическая организация гранул*, которая является частным случаем *холонической организации*. Холоническая организация часто используется в мультиагентном подходе. Согласно [Тарасов, 2002], холон – «целое, рассматриваемое в то же время как часть целого».

Согласно Л. Заде [Zadeh, 1996], существуют две основные операции по формированию и разделению гранул: «*грануляция* включает в себя разложение целого на части. И, наоборот, *организация* предполагает интеграцию частей в целое». Таким образом, существенным свойством грануляции являются операции *конструирования и деконструирования (разделения) гранул*. В то же время, согласно [Yao, 2005], «термин грануляция можно рассматривать более широко. Вместо того, чтобы вводить два понятия, т.е. грануляцию и организацию, грануляцию можно было бы определить как процесс в обоих направлениях: организация и декомпозиция».

Важными понятиями в теории грануляции являются разбиения и покрытия. Согласно [Yao et al., 2002], «разбиение состоит из непересекающихся подмножеств, а покрытие состоит из, возможно, пересекающихся подмножеств. Отношения между этими подмножествами называются отношениями разбиения и отношениями покрытия соответственно. С помощью операций разбиения или покрытия могут быть созданы деревья грануляции». Формальные определения разбиений и покрытий:

1. Для всех $i, X_i \neq \emptyset$, разбиения и покрытия не могут быть пустыми;
2. Для всех $i \neq j, X_i \cap X_j = \emptyset$, разбиение состоит из непересекающихся подмножеств, данное свойство выполняется только для разбиений;
3. $\bigcup X_i = UN$, объединение всех покрытий или разбиений образует «универсум», где X_i – разбиение или покрытие; UN – «универсум», полный набор гранул.

На основе рассмотренных предпосылок объединим гранулярный подход и протографовую модель в форме **гранулярно-протографовой модели**.

2. Основные элементы гранулярно-протографовой модели

2.1. Сигнатуры, заряды и фасеты гранул

Под *сигатурой* будем понимать произвольный элемент sig множества сигатур SIG . Удобнее всего воспринимать сигатуру как текстовый идентификатор, предназначенный для объединения элементов с одинаковыми сигатурами.

Под *зарядом* sn будем понимать элемент множества зарядов $SN = \{+, -\}$. Заряд может принимать только два значения «+» или «-».

Определим *фасет гранулы* как комбинацию сигатуры и знака:

$$fct_i = \langle sig_j, sn_k \rangle, \quad (2.1)$$

Для удобства будем использовать следующие обозначения фасетов:

$sig(123)+$ – положительно заряженный фасет с сигатурой «123».

$sig(123)-$ – отрицательно заряженный фасет с сигатурой «123».

$sig(123)?$ – фасет с сигатурой «123» и произвольным зарядом.

$sig+$ – положительно заряженный фасет с произвольной сигатурой.

$sig-$ – отрицательно заряженный фасет с произвольной сигатурой.

$sig?$ – фасет с произвольной сигатурой и произвольным зарядом.

Далее на основе фасетов гранул может быть определено понятие гранулярного элемента протографа.

2.2. Гранулярные элементы протографа

Гранулярный (гранулированный) элемент протографа (далее используется сокращение ГЭП) является аналогом элемента протографа в предлагаемой модели.

Определим гранулярный элемент протографа gp_i следующим образом:

$$\begin{aligned}
 gp_i &= \langle id_{gp}, FCT^{INNER}, FCT^{OUTER}, GP^{INNER} \rangle, \\
 FCT^{INNER} &= \{ fct_j^{INNER} \}, FCT^{OUTER} = \{ fct_k^{OUTER} \}, \\
 GP^{INNER} &= \{ gp_N \},
 \end{aligned}
 \tag{2.2}$$

где id_{gp} – уникальный идентификатор гранулярного элемента протографа; FCT^{INNER} – множество внутренних фасетов гранулярного элемента протографа; FCT^{OUTER} – множество внешних фасетов гранулярного элемента протографа; fct_j^{INNER} – фасет, принадлежащий множеству внутренних фасетов; fct_k^{OUTER} – фасет, принадлежащий множеству внешних фасетов; gp_N – вложенный гранулярный элемент протографа, принадлежащий множеству вложенных гранулярных элементов протографа GP^{INNER} , соединенный с gp_i через внутренние фасеты.

Для обоих множеств внешних и внутренних фасетов выполняются ограничения разбиения, при этом фасет не может принадлежать одновременно к двум множествам и внутренних, и внешних фасетов:

$$\begin{aligned}
 i \neq j, FCT_i^{INNER} \cap FCT_j^{INNER} &= \emptyset \\
 i \neq j, FCT_i^{OUTER} \cap FCT_j^{OUTER} &= \emptyset \\
 FCT^{INNER} \cap FCT^{OUTER} &= \emptyset.
 \end{aligned}
 \tag{2.3}$$

Таким образом, гранулярный элемент протографа можно представить как полую фигуру, наружная и внутренняя поверхности которой могут быть разделены на в общем случае бесконечное количество фасетов. Наружные и внутренние фасеты не могут пересекаться между собой. Также внутренние и внешние фасеты по отдельности являются разбиениями.

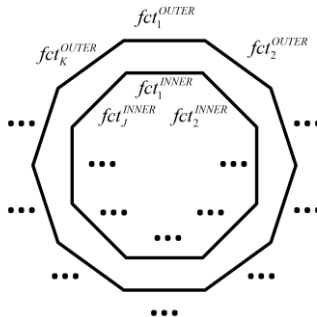


Рис. 1. Пример гранулярного элемента протографа

Пример гранулярного элемента протографа показан на рис. 1. На рисунке показаны внутренние фасеты fst_1^{INNER} , fst_2^{INNER} , ... fst_J^{INNER} , а также внешние фасеты fst_1^{OUTER} , fst_2^{OUTER} , ... fst_K^{OUTER} .

По умолчанию предполагается, что внешние фасеты обладают положительным зарядом, а внутренние – отрицательным.

Внутренние фасеты могут быть использованы для соединения с внутренними элементами, принадлежащими множеству GP^{INNER} , а внешние фасеты для соединения с внешними элементами протографовой модели.

3. Основные операции над элементами гранулярно-протографовой модели

3.1. Операция сигнатурного прилегания

Предложенная в [Гапанюк, 2022] операция прилегания «*» позволяла соединять произвольные соседние элементы протографа.

В отличие от нее, двухместная операция сигнатурного прилегания «A <=sig(s1)=> B» или сокращенно «A <=> B» позволяет соединять только фасеты с одинаковой сигатурой s1 и противоположными зарядами. Отметим, что сигнатура s1 не является параметром операции и используется в определении, только чтобы подчеркнуть совпадение сигнатур у фасетов A и B.

Следующий вариант вызова приведет к успешному выполнению операции:

$$\text{sig}(123)+ \text{<=sig}(123)\text{=>} \text{sig}(123)-.$$

В случае несовпадения сигнатур или зарядов операция не будет выполнена успешно:

- $\text{sig}(123)+ \text{<=sig}(123)\text{=>} \text{sig}(123)+$ – нельзя выполнить операцию для одноименных положительных зарядов;
- $\text{sig}(123)- \text{<=sig}(123)\text{=>} \text{sig}(123)-$ – нельзя выполнить операцию для одноименных отрицательных зарядов;
- $\text{sig}(123)+ \text{<=>} \text{sig}(321)-$ – нельзя выполнить операцию для несовпадающих сигнатур;
- $\text{sig}(123)+ \text{<=>} \text{sig}(321)+$ – нельзя выполнить операцию для одноименных зарядов и несовпадающих сигнатур.

3.2. Операция организации гранул

Необходимо отметить, что операция сигнатурного прилегания задает связь между элементами, но не объединяет автоматически эти элементы в единую сущность.

Для объединения элементов используется операция *организации графулы*, применяемая в гранулярном подходе. Она предполагает интеграцию гранулярных элементов протографа в единое целое.

В случае операции организации создается новый гранулярный элемент протографа, включающий организуемые элементы, как вложенные элементы множества $\{gp_N\}$. Операция организации записывается в форме:

$$gp_{RES} = gp_1 <+> gp_2, \quad (3.1)$$

где gp_1 и gp_2 – исходные ГЭП, объединяемые с помощью операции грануляции. Предполагается, что у каждого из ГЭП есть хотя бы один внешний положительно заряженный фасет.

В результате выполнения операции создается новый ГЭП gp_{RES} со следующими характеристиками:

- $gp_{RES}.GP^{INNER} += \{gp_1, gp_2\}$ – исходные ГЭП gp_1 и gp_2 добавляются к множеству вложенных гранулярных элементов протографа GP^{INNER} ;
- $gp_{RES}.FCT^{INNER} += \{\text{sig}(gp_1)-, \text{sig}(gp_2)-\}$ – к множеству внутренних фасетов результирующего элемента добавляются отрицательно заряженные фасеты, сигнатуры которых соответствуют сигнатурам фасетов объединяемых элементов;
- Внутренние фасеты результирующего элемента $\text{sig}(gp_1)-, \text{sig}(gp_2)-$ с использованием операции сигнатурного прилегания объединяются с внешними фасетами добавляемых элементов $\text{sig}(gp_1)+, \text{sig}(gp_2)+$, объединение происходит успешно, так как фасеты имеют противоположные заряды:
 - $gp_{RES}.FCT^{INNER}.\text{sig}(gp_1)- \Leftrightarrow gp_1.FCT^{OUTER}.\text{sig}(gp_1)+$;
 - $gp_{RES}.FCT^{INNER}.\text{sig}(gp_2)- \Leftrightarrow gp_2.FCT^{OUTER}.\text{sig}(gp_2)+$.

Отметим, что операция организации не требует, чтобы для организуемых элементов выполнялись операции сигнатурного прилегания между собой на внешних фасетах, но и не препятствует этому.

4. Основные положения аннотируемой метаграфовой модели

В данном разделе напомним основные положения аннотируемой метаграфовой модели в соответствии с [Гапанюк, 2021].

Аннотируемая метаграфовая модель расширяет идеи исходной модели А. Базу и Р. Блэннинга [Basu et al., 2007] и идеи работы [Астанин и др., 2012].

В аннотируемой метаграфовой модели метаграф определяется следующим образом:

$$MG = \langle V, MV, E \rangle, \quad (4.1)$$

где MG – метаграф; V – множество вершин метаграфа; MV – множество метавершин метаграфа; E – множество ребер метаграфа.

Вершина метаграфа характеризуется множеством атрибутов:

$$v_i = \{atr_k\}, v_i \in V, \quad (4.2)$$

где v_i – вершина метаграфа; atr_k – атрибут.

Ребро метаграфа характеризуется множеством атрибутов, исходной и конечной вершиной и признаком направленности:

$$e_i = \langle v_s, v_e, eo, \{atr_k\} \rangle, e_i \in E, eo = true \mid false, \quad (4.3)$$

где e_i – ребро метаграфа; v_s – исходная вершина (метавершина) ребра; v_e – конечная вершина (метавершина) ребра; eo – признак направленности ребра ($eo=true$ – направленное ребро, $eo=false$ – ненаправленное ребро); atr_k – атрибут.

Фрагмент метаграфа:

$$MG_i = \{ev_j\}, ev_j \in (V \cup E \cup MV), \quad (4.4)$$

где MG_i – фрагмент метаграфа; ev_j – элемент, принадлежащий объединению множеств вершин (метавершин) и ребер метаграфа.

Таким образом, фрагмент метаграфа в общем виде может содержать произвольные вершины (метавершины) и ребра.

Метавершина метаграфа:

$$mv_i = \langle \{atr_k\}, \{ev_j\} \rangle, mv_i \in MV, ev_j \in (V \cup E \cup MV), \quad (4.5)$$

где mv_i – вершина метаграфа; atr_k – атрибут, ev_j – элемент, принадлежащий объединению множеств вершин (метавершин) и ребер метаграфа. Таким образом, метавершина, в дополнение к свойствам вершины, включает вложенный фрагмент метаграфа.

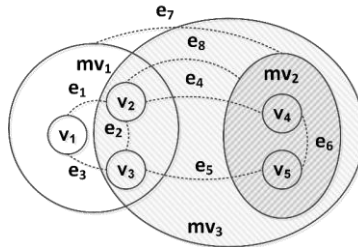


Рис. 2. Пример описания метаграфа в аннотируемой метаграфовой модели

Наличие у метавершин собственных атрибутов и связей с другими вершинами является важной особенностью метаграфов. Это соответствует принципу эмерджентности, то есть приданию понятию нового качества, несводимости понятия к сумме его составных частей. Фактически, как только вводится новое понятие в виде метавершины, оно «получает право» на собственные свойства, связи и т.д., так как в соответствии с принципом эмерджентности новое понятие обладает новым качеством и не может быть сведено к подграфу базовых понятий.

Таким образом, метаграф можно охарактеризовать как «сложный граф с эмерджентностью» или «сложную сеть с эмерджентностью», то есть фрагмент сети, состоящий из вершин и связей, может выступать как отдельное целое.

Пример описания метаграфа показан на рис. 2. Данный метаграф содержит вершины, метавершины и ребра. На рис. 2 показаны три метавершины: mv_1 , mv_2 и mv_3 . Метавершина mv_1 включает вершины v_1 , v_2 , v_3 и связывающие их ребра e_1 , e_2 , e_3 . Метавершина mv_2 включает вершины v_4 , v_5 и связывающее их ребро e_6 . Ребра e_4 , e_5 являются примерами ребер, соединяющих вершины v_2 - v_4 и v_3 - v_5 , включенные в различные метавершины mv_1 и mv_2 . Ребро e_7 является примером ребра, соединяющего метавершины mv_1 и mv_2 . Ребро e_8 является примером ребра, соединяющего вершину v_2 и метавершину mv_2 . Метавершина mv_3 включает метавершину mv_2 , вершины v_2 , v_3 и ребро e_2 из метавершины mv_1 а также ребра e_4 , e_5 , e_8 , что показывает холоническую структуру метаграфа.

Отметим, что описание различных вариантов представление атрибутов является отдельной задачей, которая не рассматривается в данной статье.

5. Представление основных элементов метаграфовой модели с помощью элементов гранулярно-протографовой модели

5.1. Представление вершин и метавершин

В терминах гранулярно-протографовой модели вершина может быть представлена как ГЭП:

$$v_i \equiv gp_i, GP^{INNER} = \emptyset, FCT^{INNER} = \emptyset, FCT^{OUTER} \neq \emptyset, \quad (5.1)$$

где v_i – вершина метаграфа; gp_i – гранулярный элемент протографа, соответствующий вершине.

Для вершины множество внутренних фасетов FCT^{INNER} и множество вложенных гранулярных элементов протографа GP^{INNER} являются пустыми, так как вершина не имеет внутренней структуры. Множество внешних фасетов FCT^{OUTER} не является пустым, что делает вершину способной к внешним соединениям.

В терминах гранулярно-протографовой модели метавершина также может быть представлена как ГЭП:

$$mv_i \equiv gp_i, GP^{INNER} \neq \emptyset, FCT^{INNER} \neq \emptyset, FCT^{OUTER} \neq \emptyset, \quad (5.2)$$

где v_i – вершина метаграфа; gp_i – гранулярный элемент протографа, соответствующий вершине.

Множество внешних фасетов FCT^{OUTER} не является пустым, что делает метавершину способной к внешним соединениям.

Множество внутренних фасетов FCT^{INNER} и множество вложенных гранулярных элементов протографа GP^{INNER} для метавершины также не являются пустыми и используются для описания внутренней структуры метавершины.

5.2. Представление ребер и соединений

В терминах гранулярно-протографовой модели ребро может быть представлено как ГЭП с двумя внешними фасетами, которые соответствуют входу и выходу ребра:

$$\begin{aligned} e_i &\equiv \langle gp_i, eo \rangle, fct_s \in FCT^{OUTER}, \\ fct_E &\in FCT^{OUTER}, FCT^{INNER} = \emptyset, \\ fct_s &= \text{sig(start)-}, fct_E = \text{sig(end)-}, \end{aligned} \quad (5.3)$$

где e_i – ребро метаграфа; gp_i – гранулярный элемент протографа, соответствующий ребру; eo – признак направленности ребра; fct_s – фасет, соответствующий исходной вершине (метавершине) ребра; fct_E – фасет, соответствующий конечной вершине (метавершине) ребра.

Для ребра множество внутренних фасетов FCT^{INNER} всегда пустое.

Фасету fct_s соответствует сигнатура $start$, а фасету fct_E соответствует сигнатура end . Предполагается, что фасеты ребра всегда отрицательно заряжены.

5.3. Комплексный пример описания метаграфа с помощью элементов гранулярно-протографовой модели

Пример описания представлен на рис. 3. В целях удобства все ГЭП обозначены как соответствующие элементы метаграфовой модели. Восьмиугольниками показаны вершины и метавершины, ромбами показаны ребра, а окружностями показаны соответствующие прилегания. Участие в операции прилегания показано пунктирными линиями.

Рассматриваемый пример содержит почти все основные случаи, характерные для метаграфовой модели. Метавершина mv_1 включает вложенные вершины v_1 и v_2 и соединяющее их ребро e_1 . При этом сама метавершина

mv_1 соединена с вершиной v_3 посредством ребра e_2 . Содержательно данный пример можно интерпретировать как фрагмент графа знаний, в котором вершины и метавершины метаграфа соответствуют объектам, а ребра соответствуют связям между объектами.

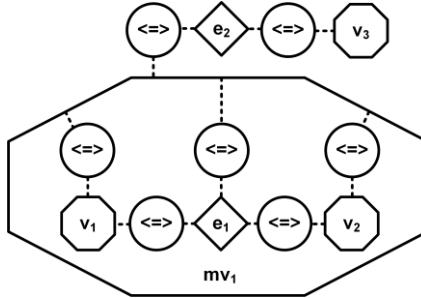


Рис. 3. Пример описания метаграфа с помощью элементов гранулярно-протографовой модели

Метавершина mv_1 включает вложенные вершины v_1, v_2 , объединенные ребром e_1 . Элементы v_1, v_2, e_1 через внешние фасеты соединены операцией сигнатурного прилегания с внутренними фасетами метавершины mv_1 :

- $mv_1.FCT^{INNER}.sig(v1)- \langle == \rangle v_1.FCT^{OUTER}.sig(v1)+$;
- $mv_1.FCT^{INNER}.sig(v2)- \langle == \rangle v_2.FCT^{OUTER}.sig(v2)+$;
- $mv_1.FCT^{INNER}.sig(e1)- \langle == \rangle e_1.FCT^{OUTER}.sig(e1)+$.

Ребро e_1 соединяет вершины v_1 и v_2 через внешние фасеты:

- $v_1.FCT^{OUTER}.sig(v1)+ \langle == \rangle e_1.FCT^{OUTER}.sig(v1)-$;
- $v_2.FCT^{OUTER}.sig(v2)+ \langle == \rangle e_1.FCT^{OUTER}.sig(v2)-$.

Ребро e_2 соединяет метавершину mv_1 и вершину v_3 через внешние фасеты:

- $mv_1.FCT^{OUTER}.sig(mv1)+ \langle == \rangle e_2.FCT^{OUTER}.sig(mv1)-$;
- $v_3.FCT^{OUTER}.sig(v3)+ \langle == \rangle e_2.FCT^{OUTER}.sig(v3)-$.

Таким образом, метаграф, содержащий метавершину, вершины и ребра, может быть описан с помощью комбинации элементов гранулярно-протографовой модели.

Заключение

Традиционная метаграфовая модель содержит три вида элементов: вершины, метавершины, ребра. Метавершина, в дополнение к свойствам вершины, включает вложенный фрагмент метаграфа.

Особенностью аннотированной метаграфовой модели является то, что одинаковый набор вершин и ребер может быть включен в несколько различных метавершин, которые могут представлять различные ситуации и быть аннотированы различными атрибутами.

На основе предложенной гранулярно-протографовой модели возможно сконструировать базовые элементы метаграфовой модели.

Предложенную гранулярно-протографовую модель можно рассматривать как базис для глубокой трансформации элементов метаграфовой модели.

Список литературы

- [Астанин и др., 2012] Астанин С.В., Драгныш Н.В., Жуковская Н.К. Вложенные метаграфы как модели сложных объектов // Инженерный вестник Дона. – 2012. – № 4. – URL: ivdon.ru/magazine/archive/n4p2y2012/1434.
- [Гапанюк, 2021] Гапанюк Ю.Е. Этапы развития метаграфовой модели данных и знаний // Интегрированные модели и мягкие вычисления в искусственном интеллекте (ИММВ-2021): Сборник научных трудов X-й Международной научно-технической конференции. В 2-х т. Т. 1. – Смоленск, 2021. – С. 190-200.
- [Гапанюк, 2022] Гапанюк Ю.Е. Расширение метаграфовой модели на основе протографового подхода // Двадцатая Национальная конференция по искусственному интеллекту с международным участием: Труды конференции. – М., 2022. – С. 62-73.
- [Кручинин, 2017а] Кручинин С.В. Протографы и архиграфы как обобщение графов // Научно-исследовательские публикации. – 2017. – № 3 (41). – С. 23-33.
- [Кручинин, 2017б] Кручинин С.В. О некоторых обобщениях графов: мультиграфы, гиперграфы, метаграфы, потоковые и портовые графы, протографы, архиграфы // Вопросы науки. – 2017. – № 3. – С. 48-67.
- [Словарь Merriam-Webster] Merriam-Webster. – <http://www.m-w.com/>, last accessed 2023/05/24.
- [Сухобоков, 2022] Сухобоков А.А. Метаграфово-табличная модель данных для систем управления активами // Искусственный интеллект в автоматизированных системах управления и обработки данных: Сборник статей Всероссийской научной конференции. В 2-х т. Т. 1. – М., 2022. – С. 93-99.
- [Тарасов, 2002] Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002. – 352 с.
- [Basu et al., 2007] Basu A., Robert W. Blanning. Metagraphs and their applications. – New York: Springer, 2007.
- [Pedrycz et al., 2011] Pedrycz W., Chen S.M. Granular Computing and Intelligent Systems. – Springer-Verlag Berlin Heidelberg, 2011.
- [Polkowski, 2023] Polkowski L. Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems / In: Lin TY., Liau CJ., Kacprzyk J. (eds) Granular, Fuzzy, and Soft Computing. Encyclopedia of Complexity and Systems Science Series. Springer, New York, (2023). – https://doi.org/10.1007/978-1-0716-2628-3_262.

- [Liu et al., 2014] Liu, Y., Zheng, X., Tang, F., Chen, X.: Ontology design with a granular approach // *Expert Systems with Applications*. – 2014. – Vol. 41, Issue 10. – P. 4867-4877.
- [Liu et al., 2022] Liu X., Song R., Wang Y., Xu H. A Multi-Granular Aggregation-Enhanced Knowledge Graph Representation for Recommendation // *Information*. – 2022. – 13 (229). – DOI: 10.3390/info13050229.
- [Stalder et al., 2021] Stalder F., Denzler A., Mazzola L. Towards Granular Knowledge Structures: Comparison of Different Approaches // In: 2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI), Herl'any, Slovakia, 2021. – P. 261-266. – DOI: 10.1109/SAMI50585.2021.9378677.
- [Yao et al., 2002] Yao J.T., Yao Y.Y. Induction of classification rules by granular computing // In: Proc. 3rd Int. Conf. Rough Sets Current Trends Comput. – 2002. – Vol. LNAI 2475. – P. 331-338.
- [Yao, 2005] Yao J.T. Information granulation and granular relationships // In: Proc. IEEE Conf. Granul. Comput. Beijing, China, 2005. – P. 326-329.
- [Yao et al., 2013] Yao J.T., Vasilakos A.V., Pedrycz W. Granular Computing: Perspectives and Challenges // *IEEE Transactions on Cybernetics*. – 2013. – 43 (6). – P. 1977-1989.
- [Zadeh, 1996] Zadeh L.A. Key roles of information granulation and fuzzy logic in human reasoning, concept formulation and computing with words // In: Proc. IEEE 5th Int. Conf. Fuzzy Syst. – 1996.

ЛИНГВИСТИЧЕСКИЕ ГРАФЫ ЗНАНИЙ ПОРТАЛА «ТЮРКСКАЯ МОРФЕМА»

А.Р. Гатиатуллин (*ayrat.gatiatullin@gmail.com*)

Н.А. Прокопьев (*nikolai.prokopyev@gmail.com*)

Д.Ш. Сулейманов (*dvdt.slt@gmail.com*)

Институт прикладной семиотики АН РТ, Казань

В работе описываются графы знаний, представленные в портале «Тюркская морфема», технологии их создания и их использование для обработки естественного языка. Особенность данных графов знаний в том, что, с одной стороны, они содержат лингвистические единицы разного языкового уровня, а с другой стороны, концепты, обозначающие значения этих лингвистических единиц, которые встроены в общую систему концептов. Данные графы знаний, используются для семантической разметки электронных корпусов, представленных в рамках портала, а также сами пополняются за счет информации, представленной в этих электронных корпусах.

Ключевые слова: граф знаний, интернет-портал, лингвистическая единица.

Введение

Количество данных, производимых и хранимых каждый день, растет практически экспоненциально, поэтому актуальной является задача их эффективной обработки. Большое разнообразие этих данных создает дополнительные сложности для их осмысленного анализа. Создается ситуация, когда методы представления и обработки знаний не успевают за постоянно растущим объемом обрабатываемых данных. Особенно это актуально для «малоресурсных языков». Понятие малоресурсных языков было введено в 2003 году нидерландским ученым С. Краувером [Krauwet, 2003]. Согласно его определению, малоресурсные языки – это естественные языки, для которых не хватает электронных лингвистических ресурсов для обработки языка и речи.

В настоящее время основная масса лингвистического программного обеспечения разработана для языков индоевропейской семьи, а они достаточно сильно отличаются по своей структуре от тюркских языков. Как следствие, компьютерные модели, описывающие индоевропейские языки, плохо подхо-

дят для описания тюркских языков. Несмотря на то, что в настоящее время активно развиваются технологии машинного обучения, которые можно обучить для любого типа языков, и при наличии больших массивов данных для этих языков в целом ряде случаев они способны решать поставленные задачи, они не способны решать все виды возникающих задач.

Таким образом, одной из важных и актуальных задач для семейства тюркских языков является создание лингвистических графов знаний и технологий их использования для эффективного решения задач обработки языка. Технологии, разработанные или адаптированные для тюркских языков, могут быть использованы для других языков агглютинативного типа, которые по другой классификации можно назвать языками с элементарно-комбинаторной грамматикой.

1. Обзор типов графов знаний

В последние годы в сфере обработки семантических данных большое внимание уделяется технологиям, называемым графы знаний [Hogan et al., 2020], [Fensel et al., 2020], [Ji et al., 2020], которые рассматриваются как системы представления знаний на основе графов, способных упорядочить информацию гибким и интуитивно понятным способом. Графы знаний активно используются в лингвистических сервисах таких крупных компаний как Google, Yandex, Facebook.

Несмотря на активное использование графов знаний, единого общепринятого их определения не существует [Ehrlinger et al., 2016]. Рассмотрим одно из определений, представленное в работе [Pan et al., 2017]:

Граф знаний – это структурированный набор данных, собранный из разнородных источников данных, совместимый с моделью данных RDF и имеющий (OWL) онтологию в качестве своей схемы. Граф знаний (ГЗ) не обязательно связан с внешними графами знаний; однако сущности в графе знаний обычно имеют информацию о типе, определенном в его онтологии, которая полезна для предоставления контекстной информации о таких сущностях.

Одни авторы видят в графах знаний некую реинкарнацию онтологий [Lawgunowicz, 2017], другие определяют их как системы, основанные на знаниях Knowledge-based systems (KBS) [Ahmed et al., 2019]. Хотя по своей сути они представляют комбинацию онтологий и тезаурусов с множествами именованных сущностей. Как показывает обзор литературы, базовыми компонентами являются такие ресурсы, как WordNet и FrameNet, остальные ресурсы объединяются вокруг них. В более выгодном положении оказываются те языки, для которых ранее были созданы подобные ресурсы. Среди тюркских языков к таким языкам относится только турецкий.

Существуют разные комбинации графов знаний с другими технологиями, представляются расширенные графы знаний (Extended knowledge graphs) [Adrian et al., 2020]. Например, Кшиштофом Яновичем разработан KnowWhereGraph, который связывает ГИС технологии с открытыми графами [Janowicz, 2020]. Конечная цель этого проекта – понять, как создавать значимые характеристики (независимые переменные) с помощью ГИС GeoKG на базе ГЗ для последующих моделей, таких как прогнозирование цепочки поставок или картографирование состояния почвы, путем включения пространственно-временной семантики.

Базовым элементом данного графа является географический концепт с соответствующим набором параметров. Пример фрагмента географического графа знаний, образуемого с помощью подобных элементов представлен на рис. 1.

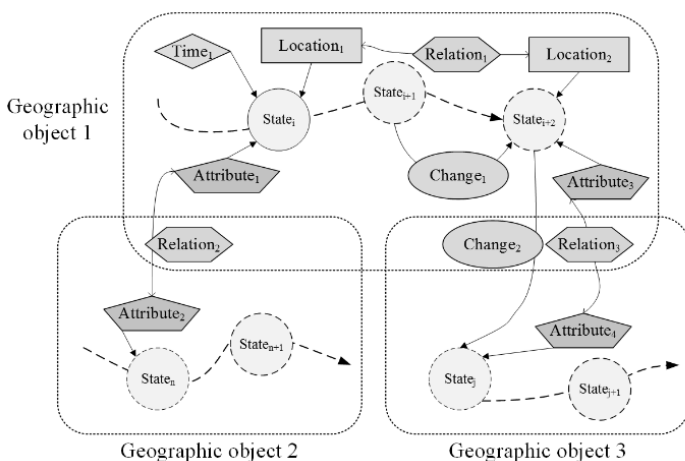


Рис. 1. Пример фрагмента GeoKG

Комбинируются графы знаний не только с географическими данными, но и с технологиями теории вероятностей, создавая отдельную разновидность графов знаний, называемыми вероятностными графами знаний (Probabilistic Knowledge Graphs, PKG). Эти ГЗ включают статистические модели для реляционных данных.

Еще одной комбинацией графов знаний являются комбинации таксономических графов знаний с данными фреймового типа, создавая так называемые Frame oriented Knowledge Graphs. Одним из таких разработок является ресурс FrameSter [Gangemi et al., 2016]. Framester является концентратором между, такими ресурсами, как FrameNet, WordNet, VerbNet,

BabelNet, DBpedia, Yago, DOLCE-Zero, а также другими ресурсы. Framester – это не только сильно связанный граф знаний, но в нем также применяется строгая формальная обработка семантики фреймов Филлмора, что позволяет выполнять полноценные запросы и рассуждения OWL на большом графе знаний на основе фреймов.

Каждый из описанных вариантов комбинирования графов знаний позволяет представлять разного типа лингвистическую информацию. Комбинирование с ГИС позволяет отобразить географию распространения языков и диалектов, а комбинирование с фреймовыми типами – представлять динамические языковые модели. Таким образом, для наиболее полного описания языковой семантики тюркских языков следует интегрировать в единую модель все описанные выше комбинации графов знаний.

Графы знаний также активно используются для представления лингвистической информации. Один из таких лингвистических графов описан в работе [Basile, 2022], модель данного лингвистического графа знаний представлена на рис. 2. По мнению авторов, данный граф позволяет моделировать: 1) отношения между понятиями и словами; 2) информацию о встречаемости слов; 3) диахроническая информация, как понятий, так и слов. Представленный в данной работе лингвистический граф включает такие вершины, как Концепты (Concept), Лексические концепты (Lexicon Concept), Лексемы (Lexicon Entry). Лексические концепты взаимосвязаны между собой семантическими отношениями типа гипонимии и гиперонимии. Лексемы взаимосвязаны с леммой (Lemma) и основой словоформы (Stem). Особенность данного графа знаний в том, что они не описывают ситуационную семантику, поэтому семантическое описание является достаточно ограниченным и отражает только таксономические отношения, аналогичные тем, что представлены в известном ресурсе WordNet. Для описания данной семантики подходят графы знаний фреймового типа, например такой достаточно известный ресурс, как FrameNet. Также в этом графе знаний не представлена грамматическая (морфологическая) структура словоформ. А для тюркских языков это очень важно, так как они обладают богатой морфологией агглютинативного типа. В отличие от языков флективного типа, куда относится русский язык, в тюркских языках четкое деление на структурные компоненты слова, которые называются морфемами. Такое четкое деление позволяет и морфологическую структуру словоформы представить в виде графа, вершинами которого являются морфемы.

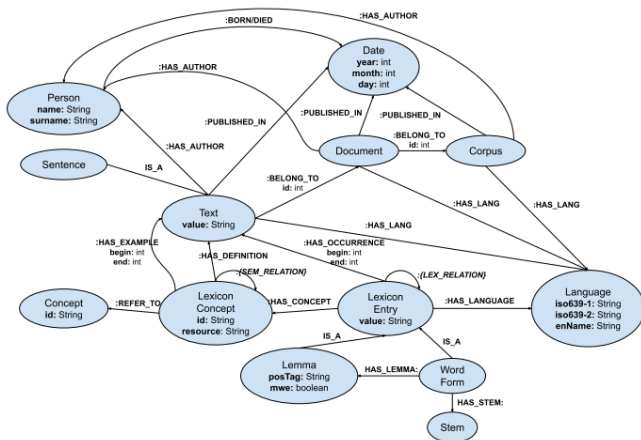


Рис. 2. Модель лингвистического графа знаний

2. Построение графов знаний

В технологиях представления знаний и работы с графами знаний главным вопросом является вопрос построения этих графов знаний. При их создании возможно извлекать данные как из структурированных, полуструктурированных, так и неструктурированных данных. Данный процесс представлен на рис. 3.

Наибольшее количество графов знаний создается для английского языка, а также для целого ряда других языков. В России они не получили широкого распространения и практически отсутствует литература на русском языке, можно лишь выделить ряд переводных работ. Например, перевод Боргестом статьи [Баклавски и др., 2020]. Возможно одна из причин в том, что в основе графов знаний находятся такие онтологические ресурсы, как WordNet и FrameNet, а для языков РФ подобных ресурсов мало (к таковым можно отнести проекты для татарского языка [Gatiatullin et al., 2020], [Kirillovich et al., 2021]) и они довольно разрозненны, что еще раз подтверждает малоресурсность данных языков. Среди тюркских языков наличием подобных ресурсов до настоящего времени выделялся только турецкий язык.

С целью решения этой проблемы для тюркских языков в Институте прикладной семиотики Академии наук Республики Татарстан ведется разработка портала “Тюркская морфема” [Gatiatullin et al., 2020]. Главной особенностью данного портала является то, что он прагматически ориентирован именно на структурно-функциональные особенности тюркских языков и направлен на решение целого класса задач.

Семиотическому исследованию особенностей лексико-грамматических характеристик тюркских языков и их описанию посвящены работы [Suleymanov, 2010] и [Сулейманов, 2021].

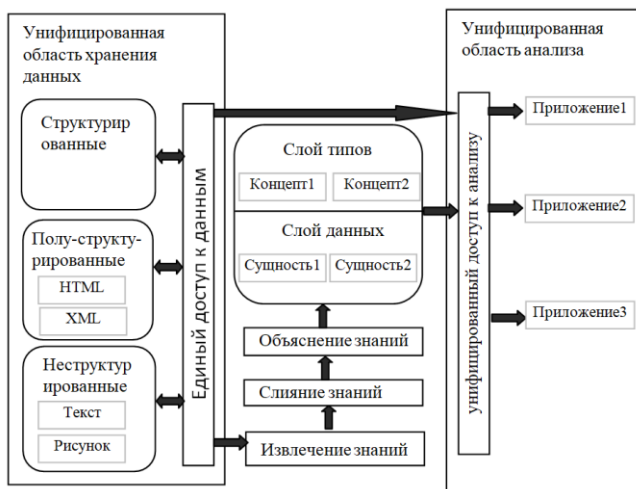


Рис. 3. Процесс построения графа знаний

Процесс построения графа знаний представлен на рис. 3. К структурированным данным относятся базы данных самого портала, к полуструктурированным – в основном размеченные корпусные данные, а к неструктурированным – произвольные данные на тюркских языках. В рамках портала реализуется единый интерфейс доступа к таким данным, в том числе электронные корпуса интегрируются с порталом в едином веб-интерфейсе.

Разрабатываемый инструментарий для построения графа знаний включает в себя инструменты извлечения и объяснения знаний (семантико-синтаксические анализаторы, использующие ресурсы портала), слияния знаний в онтологические структуры, которые в дальнейшем пополняют базы структурированных данных портала.

3. Портал «Тюркская морфема»

Интернет-портал «Тюркская морфема» представляет собой web-сайт (modmorph.turklang.net), который включает набор различных сервисов на базе лингвистических ресурсов с тюркскими языками и ориентирован на работу с тюркскими языками во всех аспектах: морфонологическом, морфологическом, синтаксическом, семантическом. Основной набор функций данного интернет-портала представлен на рис. 4.

Как указано на рис. 4, все функции портала подразделяются на базовые и прикладные. Базовый набор этих функций является более-менее устоявшимся, к нему относятся:

1. Информационно-справочная система по тюркским языкам, предоставляющая информацию о грамматике данных языков и о межязыковом тезаурусе.
2. Ресурсная онто-лингвистическая база для решения прикладных задач, таких как создание новых программных средств для обработки естественного языка, проведения лингвистических исследований.
3. Конвейер программных модулей для обработки естественного языка, уже создаваемых на базе портала и используемых для его развития.
4. Площадка для совместной работы экспертов типологов, лингвистов, диалектологов, разработчиков, так как задача создания ресурсной базы по множеству тюркских языков требует привлечения множества заинтересованных специалистов.

Набор прикладных функций включает:

1. Инструментарий для научных исследований, формируемый на ресурсной базе портала, к примеру инструменты сводных таблиц для проведения компаративных лингвистических исследований.
2. Инструментарий для создания обучающих систем, использующих базы данных портала для создания учебных материалов и проведения автоматизированного контроля знаний.
3. Инструментарий для создания терминологии и унификации обозначений для области электронной тюркологии, к примеру унификация аннотации тюркских электронных корпусов.

Данный набор будет расширяться по мере наполнения онто-лингвистической базы портала и расширения набора программных модулей, входящих в конвейер обработки естественного языка программного инструментария портала.



Рис. 4. Функции портала

4. Графы знаний портала «Тюркская морфема»

4.1. Архитектура графов знаний портала

Лингвистическая база данных портала представляет собой единый граф знаний, который подразделяется на несколько подграфов. Разделение на подграфы сделано в связи со структурными особенностями каждого из этих подграфов, а также с тем, что каждый из подграфов содержит наборы вершин одного типа. Вершины одного подграфа связаны между собой отношениями одного типа, а с вершинами из других подграфов отношениями другого типа. Схема разделения на подграфы представлена на рис. 5. Такое разделение связано и с задачами, для решения которых используются каждый из подграфов единого графа знаний портала.

Далее рассмотрим подграфы знаний портала, объединяемые в единый граф знаний.

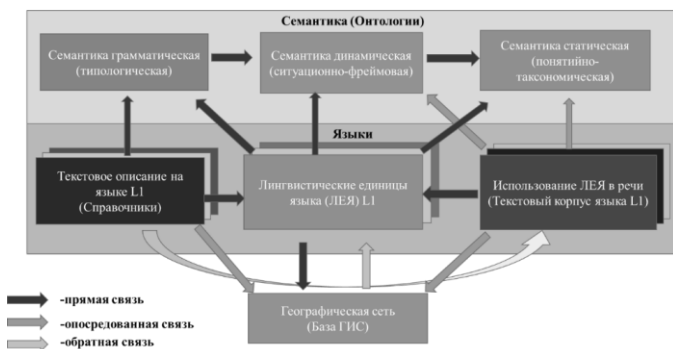


Рис. 5. Архитектура подграфов графа знаний портала

4.2. Понятийно-таксономический граф

Понятийно-таксономический граф – это граф знаний, образующий таксономию, аналогичную известному тезаурусу WordNet. Этот граф используется для описания значения лексических лингвистических единиц тюркских языков, описываемых в данном портале.

Таксономия строится с помощью триплетов, в которых субъектом и объектом триплета являются концепты, а ребро показывает отношение гиперонимии.

Каждый концепт имеет описание, состоящее из нескольких элементов:

1. Имя концепта на английском языке.
2. Синсет имен концепта на русском языке.
3. Описание концепта на английском языке.
4. Описание концепта на русском языке.

4.3. Графы знаний лингвистических единиц

Подграфы знаний с описанием лингвистических единиц содержат вершины, соответствующие лингвистическим единицам тюркских языков разного языкового уровня. Это и корневые и аффиксальные морфемы, и многословные выражения и т.д. Базовой лингвистической единицей в составе графов знаний с описанием лингвистических единиц являются морфемы. Рассмотрим один тип корневых морфем, которые несут в себе лексическое значение и могут соответствовать одному или нескольким концептам таксономического графа.

Рассмотрим схему взаимосвязей данных вершин графа. Пример фрагмента такого подграфа представлен на рис. 6. В данном примере КМ (ID 1905) и КМ (ID 1307) – это объекты – корневые морфемы татарского и турецкого языков с ID 1905 и 1307. Обе корневые морфемы в качестве значения имеют один и тот же концепт, но связь в графе осуществляется не напрямую, а через связующие вершины графа L(ID 1307x1583) и L(ID 1905x1583). ID этих элементов показывает, какие узлы графа знаний связывают эти элементы.

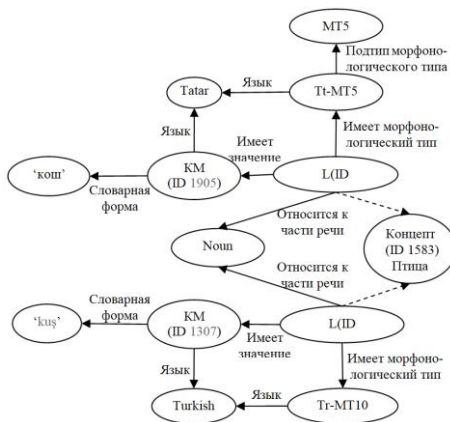


Рис. 6. Пример фрагмента таксономического графа знаний

Все связующие элементы привязаны к морфонологическим типам Tt-MT10 и Tt-MT5. Эти вершины графа, обозначающие морфонологические типы, сами являются связующими элементами между вершинами графа, соответствующих корневым и аффиксальным морфемами тюркских языков. Они содержат информацию о морфотактических правилах, показывающих какие аффиксальные морфемы могут следовать в словоформе после этих корневых морфем.

Таким образом, данные подграфы являются мультифункциональными лингвистическими ресурсами. Они, с одной стороны, используются для представления справочной информации, которая может быть выдана пользователю по запросам, а с другой стороны, описывают правила следования, используемые в морфологических анализаторах для тюркских языков.

5. Инструментарий для пополнения графов знаний портала «Тюркская Морфема»

Как было ранее сказано в данной статье, одна из основных задач при работе с графами знаний – это построение графов знаний из структурированных, полуструктурированных и неструктурированных данных. Источниками неструктурированных и полуструктурированных данных являются электронные корпуса текстов. Рассмотрим, как интегрируются данные источники с порталом «Тюркская морфема» (рис. 7).



Рис. 7. Взаимодействие графов знаний и электронных корпусов в рамках портала

Для интеграции корпусов тюркских языков в портал «Тюркская Морфема» в инструментарии портала разрабатываются две подсистемы:

1. Подсистема для создания новых корпусов.
2. Подсистема для интеграции существующих корпусов.

Подсистема для создания новых корпусов обеспечивает структуру базы данных, которая пригодна для большинства корпусов. Сущности этой базы данных связаны с языковыми элементами базы данных портала на основных уровнях: грамматическом, синтаксическом, семантическом.

Данная подсистема функционирует следующим образом: для каждого нового корпуса создается отдельное типовое веб-приложение, называемое «Модуль корпуса». Каждый модуль корпуса связывается с базой данных портала через предоставляемый единый программный интерфейс (API).

Внутри модуля корпуса реализуется база данных по схеме, описанной ранее. Кроме того, модули корпуса реализуют единый API, который обеспечивает доступ к данным при помощи поисковых запросов для корпус-менеджера и позволяет производить выгрузку данных для инструментов обработки естественного языка и графических интерфейсов для отображения сводных данных по нескольким языкам.

Для тюркских языков существует множество уже разработанных и разрабатываемых электронных корпусов, наполненных текстами, размеченных и имеющих собственные структуры данных, интерфейсы и сопутствующее программное обеспечение. Интеграция таких корпусов с порталом «Тюркская морфема» возможна при помощи соответствующей подсистемы. Основные задачи, которые должна решать подсистема:

- Независимость инструментария от структуры связываемого корпуса.
- Избежание дублирования данных.

Основной идеей для связывания существующего корпуса с порталом является создание отдельных программ-прослоек для каждого связываемого корпуса. Главной функцией подсистемы здесь является обеспечение единой архитектуры и требование наличия API для доступа к данным корпуса. Каждая такая программа-прослойка реализует функции для взаимодействия с API портала, API связываемого корпуса и собственную базу данных.

Заключение

В статье описаны перспективы использования лингвистических графов знаний в задачах обработки, накопления и изучения языковой информации. Представлены наработки и идеи развития портала «Тюркская морфема» в направлении построения графов знаний тюркских языков. В частности, рассмотрены инструменты для интеграции электронных корпусов с порталом, позволяющие, с одной стороны, пополнять базы данных портала за счет анализа корпусных данных, а с другой, использовать конвейер программных модулей для обработки естественного языка и унифицированную систему аннотации для разметки корпусов.

Таким образом, в настоящее время осуществляется формирование единого инструментария портала на основе использования разных типов интеграции графов знаний портала и электронных корпусов для тюркских языков.

Список литературы

- [Баклавски и др., 2020] Баклавски К., Беннет М., Берг-Кросс Г., Шнайдер Т., Шарма Р., Сингер Д. Онтологический Саммит 2020. Коммюнике: Графы знаний // Онтология проектирования. – 2020. – Т. 10, № 4 (38).
- [Сулейманов, 2021] Сулейманов Д.Ш. Инфокоммуникационные технологии и естественный язык: региональный опыт // Труды 19 Национальной конференции по искусственному интеллекту КИИ-2021. – Таганрог, 2021.

- [Adrian et al., 2020] Adrian W.T., Adrian M., Kluza K., Stachura-Terlecka B., Ligeza A. Extended knowledge graphs: A conceptual study // IC3K 2020 - Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 2020.
- [Ahmed et al., 2019] Ahmed A., Al-Masri N., Abu Sultan Y.S., Akkila A.N., Almasri A., Mahmoud A.Y., Zaqout I.S., AbuNaser S.S. Knowledge-based systems survey // International Journal of Academic Engineering Research (IJAER). – 2019. – Vol. 3 Issue 7.
- [Basile et al., 2022] Pierpaolo Basile, Pierluigi Cassotti, Stefano Ferilli1, and Barbara McGillivray New Time-sensitive Model of Linguistic Knowledge for Graph Databases // Proceedings of the 1st Workshop on Artificial Intelligence for Cultural Heritage co-located with the 21st International Conference of the Italian Association for Artificial Intelligence (AIXIA 2022), CEUR Workshop Proceedings. – 2022. – Vol. 3286.
- [Ehrlinger et al., 2016] Ehrlinger L., Wöß W. Towards a definition of knowledge graphs. SEMANTICS (Posters, Demos, SuCCESs) // Proc. 12th International Conference on Semantic Systems - SEMANTiCS2016, CEUR Workshop Proceedings. – 2016. – Vol. 1695.
- [Fensel et al., 2020] Fensel D., Şimşek U., Angele K., Huaman E., Kärle E., Panasiuk O., Toma I., Umbrich J., Wahler A. Knowledge Graphs. Methodology, Tools and Selected Use Cases. – Springer Nature, 2020.
- [Gangemi et al., 2016] Gangemi A., Alam M., Asprino L., Presutti V., Recupero D.R. Framester: A Wide Coverage Linguistic Linked Data Hub // Lecture Notes in Computer Science. – Springer Cham, 2016. – Vol. 10024.
- [Gatiatullin et al., 2020] Gatiatullin A., Suleymanov D., Prokopyev N., Khakimov B. About Turkic Morpheme Portal // Proc. Computational Models in Language and Speech Workshop 2020. CEUR Workshop Proceedings. – 2020. – Vol. 2780.
- [Hogan et al., 2020] Hogan A., Blomqvist E., Cochez M., d'Amato C., de Melo G., Gutierrez C., Gayo J.E.L., Kirrane S., Neumaier S., Polleres A. Knowledge graphs // arXiv preprint arXiv:2003.02320. – 2020.
- [Ji et al., 2020] Ji S., Pan S., Cambria E., Marttinen P., Yu P. S. A survey on knowledge graphs: Representation, acquisition and applications // arXiv preprint arXiv:2002.00388. – 2020.
- [Krauwert, 2003] Krauwert S. The basic language resource kit (BLARK) as the first milestone for the language resources roadmap // Proc. International workshop on speech and computer SPECOM-2003, Moscow. 2003.
- [Lawrynowicz, 2017] Lawrynowicz A. Semantic data mining: an ontology-based approach // Studies on Semantic Web. – 2017. – Vol. 29.
- [Pan et al., 2017] Pan J.Z., Vetere G., Gomez-Perez J.M., Wu H. Exploiting Linked Data and Knowledge Graphs in Large Organizations. – Springer Cham. 2017.
- [Suleymanov, 2010] Suleymanov D.Sh. Natural Cognitive Mechanisms in the Tatar language // In Proc. 20th European Meeting in Cybernetics and Systems Research, Austria. 2010.
- [Gatiatullin et al., 2020] Gatiatullin, A., Kirillovich, A., Nevzorova, O. On developing of the FrameNet-like resource for Tatar // CEUR Workshop Proceedings this link is disabled. – 2020. – 2790. – P. 344-349.
- [Kirillovich et al., 2021] Kirillovich A., Shaekhov M., Galieva A., Ilvovsky D., Loukachevitch N. TatWordNet: A linguistic linked open data-integrated WordNet resource for tatar // OpenAccess Series in Informatics. – 2021, – 93, 16.

УДК 004.82

ОНТОЛОГИЧЕСКИЕ ПАТТЕРНЫ ЗНАНИЙ ПО ДИАГНОСТИКЕ ПРОЦЕССОВ*

В.В. Грибова (*gribova@iacp.dvo.ru*)

Е.А. Шалфеева (*shalfe@iacp.dvo.ru*)

Институт автоматики и процессов управления ДВО РАН,
Дальневосточный федеральный университет, Владивосток

Онтологические паттерны являются инструментом снижения трудоемкости разработки интеллектуальных систем. В работе представлен паттерн для класса задач диагностики. Он содержит комплекс предметно-независимых семантических структур, универсальный диагностический решатель и средства адаптации предметно-независимой модели к специализированным, благодаря чему формирование экспертных знаний осуществляется в предметных терминах.

Ключевые слова: онтологический паттерн, знания о диагностике, предметно-ориентированная онтология, адаптация онтологии.

Введение

Создание интеллектуальных программных систем, решающих задачи диагностики процессов и состояний, является трудоемкой, поэтому необходимы эффективные технологии для снижения затрат на разработку и внедрение таких систем. Данная задача реализуется двумя основными методами: машинного обучения и инженерии знаний. В первом случае эффективность решения обеспечивается, прежде всего, наличием качественных (желательно больших по объему) датасетов. Второй метод применим, когда датасетов нет (либо они не содержат всю необходимую для диагностики информацию), когда требуется производить диагностику в соответствии с жесткими изменяющимися регламентами, когда исторические данные, составляющие основу датасетов, быстро устаревают (часто меняются методы диагностики), когда необходима «тонкая» настройка систем диагностики на конкретную ситуацию (объект, организацию и др.)

* Работа выполнена при поддержке Минобрнауки России: № FZNS-2023-0010 (модели информации и адаптации) и № 0202-2021-0004 (реализация решателя на IACaaS Платформе).

и набор методов диагностики, а также, когда требуется детализированное объяснение (обоснование) полученных результатов, а не их интерпретация (реализуемая в системах машинного обучения). Наиболее эффективным методом инженерии знаний является использование онтологического подхода.

Создание систем диагностики с использованием последнего можно реализовать двумя основными способами: реализацией предметно-зависимой и предметно-независимой диагностики. Очевидно, что предметно-независимая диагностика, являясь более универсальной, призвана снизить затраты на разработку диагностической системы для предметных областей, где еще не создано СППР, специализированной, ориентированной на конкретную задачу. Однако, у данного подхода существует ряд недостатков:

- использование унифицированных онтологий усложняет объяснение результата диагностики;
- использование терминов, не свойственных предметной области, в которой проводится диагностика, не позволяет корректно задавать знания в базе знаний и знания о входном диагностируемом объекте.

Целью является создание онтологического паттерна знаний (унифицированной онтологии) по диагностике процессов и метода его адаптации для конкретных предметных областей.

1. Постановка задачи

Паттерны онтологического проектирования [Загорулько и др., 2017; Gangemi et al, 2005] широко используются для создания тематических порталов знаний. Их разработка, как правило, ведется с учетом классификации паттернов, предложенной Ассоциацией ODPA (<https://github.com/odpa>) и включает построение специализированных онтологий предметных областей с использованием известных паттернов [Gangemi, 2009]. В работах [Боровикова и др., 2017; Gangemi, 2005] предлагается ряд таких решений, их основное преимущество – удобство и стандартизация формирования информации, которая фиксирует описание проверенных на практике решений проблем, и позволяет всесторонне описать различные аспекты предметной области и соответствующих задач. Однако, такие онтологические паттерны не являются реализациями, позволяющими автоматизировать процесс рассуждения наукоемких задач (прогноза, диагностики, планирования и т.п.), кроме отдельных решений, связанных с классификацией на основе значимых атрибутов [Боровикова и др., 2017].

С помощью понятий и отношений их базовых онтологий - онтологии задач и методов – могут быть описаны задачи, методы их решения и реализующие их веб-сервисы. Например, элементы описания паттерна мето-

да исследования представлены такими обязательными классами онтологии, как «Деятельность», «Научный Результат», «Задача», «Метод решения» и соответствующими отношениями "используется В", "реализуется В", "решает". Они делают удобным информацию для читателей (подобно резюме), которые ищут ответы на вопросы "В какой деятельности используется названный метод?", "В каких результатах метод реализован?" и т.п. Не предполагается существование алгоритма или процедуры построения программного решателя для класса задач на основе представляемого «метода решения». Во всяком случае, если речь идет о задаче диагностики, тем более для динамических систем.

Неформальную постановку актуальной задачи можно представить следующим образом. Необходимо разработать предметно-независимую программную оболочку для создания диагностических систем в различных предметных областях, содержащую онтологию знаний по диагностике процессов, онтологию входных данных, онтологию объяснения, универсальный решатель, основанный на данных онтологиях, и метод адаптации универсальных онтологий к предметно-ориентированным, т.е. специализированным. Адаптация универсальных онтологий к специализированным необходима для того, чтобы специалисты и эксперты конкретной предметной области могли формировать базы знаний и данных в понятных им терминах.

Универсальный решатель должен адаптироваться («понимать») специализированную онтологию без необходимости его модификации.

Онтологическая модель знаний должна учитывать внешние признаки и влияющие на диагностику факторы, признаки, присущие непосредственно процессу или объекту диагностики; необходимо учитывать также, что процесс может протекать во времени.

Задача проблемно-независимой диагностики процессов состоит в определении типа отклонения состояния объекта от нормы на основе знаний о его признаках, изменяемых и не изменяемых во времени.

Общая постановка задачи.

Дано:

- База знаний о диагностике отклонений в терминах конкретной предметной области, удовлетворяющая условию делимости классов-диагнозов.
- Признаковое описание объекта конкретной предметной области, содержащее информацию об объекте (признаки и их значения).
- Онтология объяснения диагностики процессов.
- Онтология предметно-независимой диагностики процессов.

Найти:

- Результат диагностики в виде набора гипотез.

- Объяснение диагностики процессов в терминах конкретной предметной области.

Признаковое описание входного объекта состоит из: множества *моментов* времени; для каждого момента определен комплекс *диагностических признаков*, как множество функций от момента времени; для каждого момента времени определено множество *событий*.

В требуемой программной системе адаптации для интеллектуальной системы на основе онтологий должен быть онтологический решатель классов задач (выбранным классом является диагностика), база терминов для формализации языка взаимодействия системы с пользователем и система определяющая связь абстрактных терминов решателя с терминами из соответствующей базы.

Решатель должен проводить дифференциальную диагностику и дальнейшую диагностику объекта с объяснением процесса идентификации дефекта, а также должна быть возможность рассмотрения параллельно нескольких гипотез о диагнозе.

2. Результат

В данном разделе представлен комплекс универсальных предметно-независимых онтологий, метод адаптации универсальных онтологий к специализированным на основе таблицы соответствий, а также структура универсального решателя, задача которого найти все подтвержденные, опровергнутые и возможные гипотезы по входным данным диагностируемого объекта. Результат работы алгоритма используется для объяснения процесса диагностики.

2.1. Универсальные предметно-независимые онтологии

Онтологическая модель знаний диагностики процесса или состояния динамической системы связывает класс отклонений или диагноз с динамикой внешних признаков (изменяющихся их значений), факторы, влияющие как на возможность зарождения аномального процесса, так и на варианты изменений признаков. Обобщенная модель выражена в предметно-независимых понятиях (рис. 1).

Онтология данных определяет формат заполнения наблюдаемых сведений в составе названия признака, его значения и даты-времени наблюдения (рис. 2). Предлагается также возможность представления составных наблюдений, которые включают имя, набор возможных характеристик некоторого наблюдаемого объекта, каждая из которых имеет свое множество возможных значений (например, наличие отклонения от нормы и характеристики отклонения: цвет, структура, смещение и пр.).

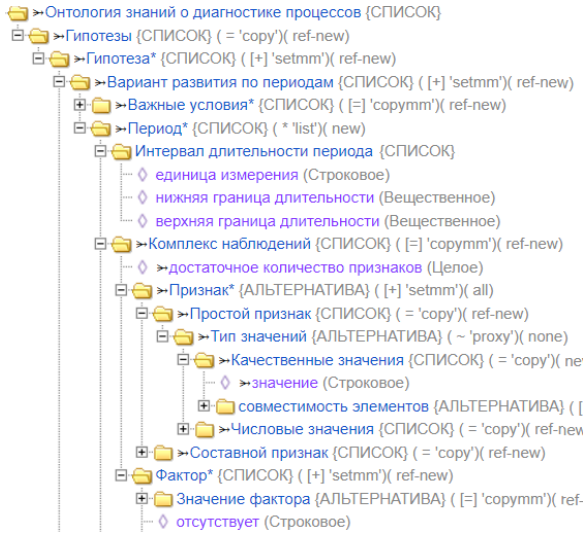


Рис. 1. Обобщенная онтология диагностики

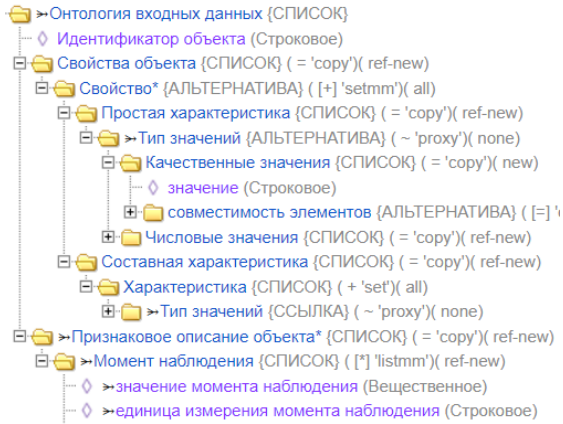


Рис. 2. Единая обобщенная онтология наблюдаемых фактов

Онтология объяснения (рис. 3) включает ссылку на признаковое описание объекта, статус диагностики, а также множество подтвержденных, возможных и опровергнутых гипотез (каждое из подмножеств может оказаться пустым).

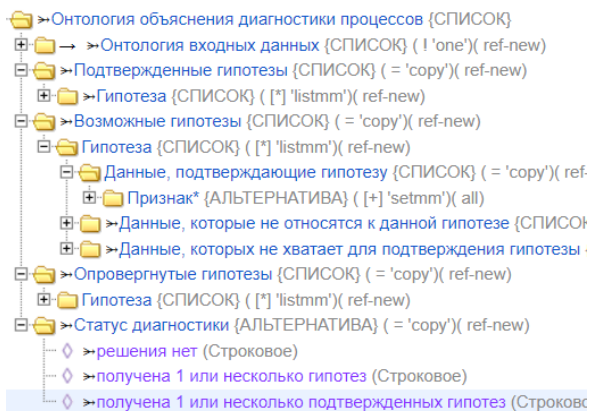


Рис. 3 Структура объяснения результатов диагностики

2.2. Метод адаптации программных обработчиков оболочки

Онтология диагностики имеет универсальную структуру и терминологию. Чтобы ее адаптировать на конкретную предметную область, предложена таблица соответствий, структура которой приведена на рис. 4. Она позволяет гибко настроить структуру понятий и отношений универсальной онтологии под специфику конкретной онтологии. Таблица соответствий сопоставляет абстрактные понятия из универсальной онтологии и конкретные понятия в целевой. Помимо задания терминологии, в целевых онтологиях возможно различное количество вершин на одном структурном уровне, т.е. одному абстрактному понятию могут соответствовать несколько конкретных. Таким образом, формируются специализированные онтологии диагностики, адаптированные к конкретным предметным областям. На основе данных онтологий формируются уже соответствующие базы знаний по диагностике процессов в конкретной предметной области. Таким образом, изменяя количество вершин и их имена, универсальная онтология адаптируется к конкретной, специализированной онтологии по диагностике.

2.3. Универсальный решатель

Решатель конструируется как описание агентов-обработчиков информации, созданной по предметной онтологии и онтологии таблицы соответствий (адаптации) ее терминов к терминам универсальной онтологии. Таким образом, совокупность онтологий и решателя, на них основанного, является программной оболочкой. Для того, чтобы адаптировать универсальный решатель к специализированному, необходимо указать имена

таблиц соответствий – входных данных, базы знаний, объяснения, а также, соответственно, ссылки на базу знаний, входные данных и объяснение (выступающие как фактические параметры решателя), рис. 5.

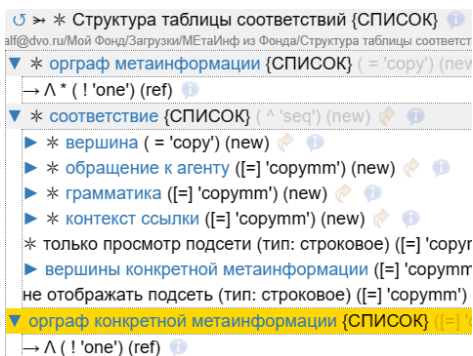


Рис. 4. Онтология таблицы соответствий

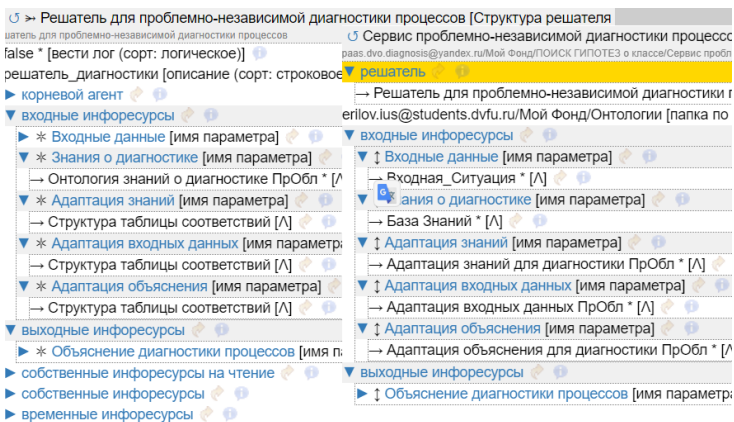


Рис. 5. Решатель и сервис

Программное средство адаптации унифицированной онтологии диагностики процессов для конкретной предметной области разрабатывалось на основе агентного подхода IASaaS с разработкой кода агентов на языке Java. Объём кода составляет 0,7 тыс. строк. Представление всех иерархических оргграфов обрабатывалось с помощью единого API платформы для доступа к единицам хранения всех типов, формируемых единым инструментальным структурным редактором платформы.

Заключение

В данной работе представлен подход к созданию интеллектуальной системы диагностики. В качестве основного метода решения предлагается подход на основе инженерии знаний. Система диагностики проектируется на основе онтологических паттернов, которые представляют собой комплекс предметно-независимых онтологий, универсальный решатель и систему адаптации предметно-независимых паттернов к специализированным. Основная задача – снизить трудоемкость разработки систем такого класса. В предложенном авторами решении интеллектуальная система, построенная по такому принципу, исключает сложный процесс создания онтологии знаний, онтологического решателя и системы генерации объяснений. Процесс разработки заключается в адаптации универсальных онтологий к специализированным через таблицу соответствий и построение базы знаний.

В настоящее время разработан прототип программной оболочки, проводится его опытная эксплуатация. Программная оболочка реализована на платформе IASPaas.

Список литературы

- [Загорулько и др., 2017] Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б. Применение паттернов онтологического проектирования при разработке онтологий научных предметных областей // Труды XIX Международной конференции «Аналитика и управление данными в областях с интенсивным использованием данных» (DAMDID/RCDL'2017). – М., 2017.
- [Gangemi et al, 2005] Gangemi A. Ontology Design Patterns for Semantic Web Content / In: Gil Y., Motta E., Benjamins V.R., Musen M.A. (eds) // The Semantic Web. – Springer: Berlin, Heidelberg. – ISWC 2005. – LNCS, 3729. – P. 262-276.
- [Gangemi, 2009] Gangemi A., Presutti V. Ontology Design Patterns // Handbook on Ontologies. – 2009. – P. 221-243. – DOI: 10.1007/978-3-540-92673-3_10.
- [Боровикова и др., 2017] Боровикова О.И., Загорулько Г.Б., Загорулько Ю.А., Шестаков В.К. Использование паттернов для разработки онтологии информационно-аналитического интернет-ресурса "поддержка принятия решений" // Информационные и математические технологии в науке и управлении. – 2017. – № 3 (7).

ЯЗЫК ЗАПРОСОВ К РАСШИРЕННЫМ СЕМАНТИЧЕСКИМ СЕТЯМ

Д.В. Демидов (*dvdemidov@mephi.ru*)

Национальный исследовательский ядерный университет МИФИ,
Москва

В работе описывается язык запросов к расширенным семантическим сетям, в которых отношения характеризуются в общем случае различной местностью. Дается ряд алгоритмов интерпретатора языка запросов.

Ключевые слова: расширенные семантические сети, язык запросов, интерпретатор запросов.

Введение

Семантические сети известны с работ Р.Куиллиана конца 1960-х как графы, в которых дуги моделируют отношения между понятиями или словами. С тех пор семантические сети применялись и в языках представления знаний, таких как SNePS, и расширялись логиками для повышения выразительных возможностей, в частности, дав развитие дескрипционным логикам и позднее концепции семантической паутины. Так, в [Segaran et al., 2009] описывается язык триплетов RDF, язык описания онтологий OWL, язык запросов SPARQL. В [Shapiro, 1979] описана система вывода и представления знаний SNePS, в основе которой лежат пропозициональные семантические сети, а в [Kandfer et al., 2008] SNePS сравнивается с машиной вывода Pellet для OWL. При обработке семантических сетей возникают задачи выборки подсетей, которые можно рассматривать как разновидность задачи программирования с ограничениями (constraint logic programming). В [Hentenryck, 1989], [Marriott et al., 1998] описаны подходы к программированию с ограничениями, реализованные в системе ECLiPSe. Интересно, что многоместные отношения моделируются в ECLiPSe и его прародителе языке Prolog в виде предикатов первого порядка. Многоместные отношения как раз находятся в фокусе внимания этой работы.

За время эволюции семантических сетей исследователями предлагались различные расширения формализма. Так, в работе [Хендрикс, 1975] семантические сети расширены подпространствами, а в [Попов, 1982] это

расширение модифицировано для более адекватного кодирования кванторных утверждений. В работах [Кузнецов, 1986], [Кузнецов, 1988], [Кузнецов, 2008] формализм расширен вершинами связи (с-вершинами), р-вершинами, выражающими истинность высказывания, множественными отношениями. Отметим, что триплеты RDF представляют собой бинарные отношения, которые удобно визуализировать в виде графа и представлять в реляционной таблице, а в формализме Кузнецова И.П. кортежи имеют различную длину, что делает их гораздо менее удобными для визуализации и реляционного представления. Эта статья посвящена построению языка запросов к расширенным семантическим сетям (РСС) в смысле Кузнецова.

В РСС вершины моделируют объекты, кортежи моделируют отношения между объектами. Вводятся вершины двух типов: *определённые вершины (о-вершины)* – моделируют объекты, которые формализуются на языке программирования в виде чисел, строк, булевых значений; и *неопределённые вершины (н-вершины)* – моделируют объекты реального мира, которые можно описать с помощью отношений. Н-вершины аналогичны переменным в языках программирования, поэтому их удобно представлять уникальными идентификаторами. Далее в примерах они представлены в виде строк, начинающихся с “_”.

Вершины в кортежах интерпретируются следующим образом:

1. Первый элемент кортежа – с-вершина. Интерпретируема как идентификатор кортежа, хотя в аппарате РСС ограничение на уникальность не вводится. Полезно, когда моделируемое кортежем отношение является аргументом другого отношения.
2. Второй элемент кортежа – р-вершина, выражающая истинность отношения между объектами. Позволяет отказаться от предположения о замкнутости мира и явно выражать значение истинности, причём не только истину и ложь, но и неопределённое значение истинности (с помощью н-вершины).
3. Третий элемент кортежа – вершина, выражающая отношение.

Следующие элементы кортежа выражают аргументы отношения.

Например, высказывание "Истинно то, что некий объект с идентификатором `text_s1_t6_h1` является предикатом" выразимо кортежем `<_, true, Является, _text_s1_t6_h1, ПРЕДИКАТ>`, где первый элемент выражен универсальным образцом `_`, что при вставке в сеть приводит к образованию уникальной н-вершины; о-вершина `true`, означает, что отношение истинно; о-вершина "Является" – имя отношения; именованная н-вершина `"_text_s1_t6_h1"` и о-вершина "ПРЕДИКАТ" – два аргумента отношения "Является".

1. Типы запросов

Базовый язык предоставляет возможности записи следующих типов запросов: выборка кортежей, выборка элементов, вставка, изменение, удаление кортежей, унификация вершин. Приведём несколько примеров.

1.1. Выборка подсети

Пример №1. Выбрать все кортежи длины 5, в которых третий аргумент равен "Ссылка на":

```
select * arity 5
where arg[3] == "Ссылка на"
```

Пример №2. Выбрать первые 4 кортежа по образцу, где третий аргумент равен "Ссылка на", упорядочив по 4-й колонке:

```
select * like [_, _, "Ссылка на", _, _]
order by 4
limit 4
```

Пример №3. Выбрать все кортежи с отношением экземпляр-класс, упорядочить по 4-му элементу, сгруппировать по 5-му элементу.

```
select * like [_, True, "Экземпляр", _, _]
group by 5
```

Результат в нотации JSON¹:

```
{
  'ДЕЛАТЬ': [
    ['_', True, 'Экземпляр', '_text_s1_t6_h1', 'ДЕЛАТЬ'],
    ['_', True, 'Экземпляр', '_text_s1_t15_h1', 'ДЕЛАТЬ']
  ],
  'РАБОЧИЙ': [
    ['_', True, 'Экземпляр', '__2263466468832', 'РАБОЧИЙ'],
    ['_', True, 'Экземпляр', '__2263466553936', 'РАБОЧИЙ']
  ],
  'СДЕЛАТЬ': [
    ['_', True, 'Экземпляр', '_text_s2_t5_h1', 'СДЕЛАТЬ']
  ]
}
```

1.2. Выборка вершин

Пример №4. Выбрать такие n-вершины из заданного множества, на которые есть ссылка "они" (отношение "Ссылка на").

```
select nodes ?x
where
```

¹ <https://www.json.org/json-en.html>.


```
[_, True, "Ссылка на", "они", ?x],  
?x in [1, "_text_s1_t15_h1", "__2263466553936"]
```

Результат:

```
'__2263466553936'
```

Пример №5. Выбрать такие x , которые являются первыми аргументами отношений Агент и Объект одновременно:

```
select nodes x  
where  
  [_, _, "Агент", x, _],  
  [_, _, "Объект", x, _]
```

Результат:

```
{'text_s1_t6_h1', 'text_s1_t15_h1', 'text_s2_t5_h1'}
```

Пример №6. Выбрать согласованные без учёта перестановок тройки x , y , z такие, что на один и тот же объект z (n -вершину) имеются различные ссылки x и y (o -вершины):

```
select nodes $x, $y, ?z  
where  
  ordered($x, $y),  
  [_, True, "Ссылка на", $x, ?z],  
  [_, True, "Ссылка на", $y, ?z]
```

Результат:

```
('1', 'первый', '__2263466468832')  
(2, 'второй', '__2263466471088')
```

Пример №7. Выбрать такие n -вершины x , которые являются агентами ситуаций и не являются экземплярами никаких объектов, выраженных o -вершинами:

```
select nodes ?x  
where  
  exists [_, True, "Агент", ?x, _],  
  not exists [_, True, "Экземпляр", ?x, $y]
```

Результатом будет пустое множество.

1.3. Вставка кортежей

Пример №8. Вставить кортежи вида $X \in Y$, где X , Y – n -вершины, X – агент Y , а "они" – ссылка на Y :

```
insert [_, True, "€", ?x, ?y]  
where  
  [_, True, "Агент", ?x, ?y],  
  [_, True, "Ссылка на", "они", ?y]
```

Добавляется один кортеж

```
[_, True, "€", _text_s2_t5_h1, __2263466553936]
```

1.4. Изменение кортежей

Пример №9. Изменить третий аргумент кортежей на "Ссылка", там, где он был равен "Ссылка на":

```
update set arg[2] = "Ссылка"  
where arg[2] == "Ссылка на"
```

1.5. Удаление кортежей

Пример №10. Удалить все кортежи, у которых третий аргумент равен "Ссылка на":

```
delete * arity 5  
where arg[2] == "Ссылка на"  
Альтернативный вариант запроса:  
delete * like [_, _, "Ссылка на", _, _]
```

1.6. Унификация вершин

Пример №11. Добавить отношение эквивалентности между каждой парой различных объектов, на которые ссылается одна и та же ссылка:

```
unify nodes ?x, ?y  
where  
  ?x != ?y,  
  [_, True, "Ссылка на", z, ?x],  
  [_, True, "Ссылка на", z, ?y]
```

Унификация вершин интерпретируется либо как вставка кортежа с отношением эквивалентности, либо как замена одной вершины на другую во всех кортежах с помощью запроса на изменение.

2. Описание ограничений

2.1. Ограничение в виде кортежа-образца для сопоставления

Ограничение может использоваться для выборки подсети, но не для выборки узлов. Может упоминаться 0 и более переменных.

Ограничение этого типа может встречаться в запросе на выборку подсети не более одного раза, а в запросе на выборку узлов – многократно. Может являться единственным в запросе, так как производит подсеть. Извлекает подсеть, кортежи которой сопоставлены с кортежем-образцом. Если в кортеже-образце упоминаются переменные-образцы, то соответствующие им значения связываются в каждом отобранном кортеже, образуя множество согласованных значений. Проекция подсети на упомянутые переменные сохраняет порядок следования переменных и производит множество укороченных кортежей. Если переменная упомянута в кортеже-образце более одного раза, то в проекции остаётся только первое вхождение.

2.2. Ограничение на существование подходящего кортежа

Ограничение может использоваться для отбора допустимых значений переменных, упомянутых в ограничении. Может упоминаться 0 и более переменных. Ограничение может встречаться в запросе многократно и может являться единственным в запросе, так как продуцирует подсеть.

Извлекает подсеть, кортежи которой сопоставлены с кортежем-образцом. Отсекает такие согласованные значения упомянутых переменных, для которых не нашлось ни одного кортежа. Если множество упомянутых переменных не пересекается с множеством целевых переменных, то играет роль только пуста или не пуста отобранная подсеть.

2.3. Ограничение на отсутствие подходящего кортежа

Ограничение может использоваться для отбора допустимых значений переменных, упомянутых в ограничении, но не для выборки подсети. Может упоминаться 0 и более переменных. Ограничение может встречаться в запросе многократно. Не продуцирует подсеть с допустимыми кортежами, поэтому не является достаточным для выборки.

Извлекает подсеть, кортежи которой сопоставлены с кортежем-образцом. Отсекает такие согласованные значения упомянутых переменных, для которых нашёлся хотя бы один кортеж. Должно накладываться после ограничения, продуцирующего допустимую подсеть, иначе не из чего будет вычитать исключаемые кортежи. Если множество упомянутых переменных не пересекается с множеством целевых переменных, то играет роль только пуста или не пуста отобранная подсеть.

2.4. Ограничение на тождество вершин, выраженных образцами

Ограничение может использоваться для отбора допустимых значений переменных, упомянутых в ограничении, но не для выборки подсети.

Должна упоминаться хотя бы 1 (в сочетании с константой) или более 1 переменной. Ограничение может встречаться в запросе многократно. Не продуцирует подсети с допустимыми значениями, поэтому не является достаточным для выборки. Может налагаться только на кортежи согласованных значений. В результате распространения ограничения остаются только такие сочетания элементов, в которых элементы равны.

2.5. Ограничение на различие вершин, выраженных образцами

Ограничение может использоваться для отбора допустимых значений переменных, упомянутых в ограничении, но не для выборки подсети.

Должна упоминаться хотя бы 1 (в сочетании с константой) или более 1 переменной. Ограничение может встречаться в запросе многократно. Не продуцирует подсети с допустимыми значениями, поэтому не является

достаточным для выборки. Может налагаться только на кортежи согласованных значений. В результате распространения ограничения остаются только такие сочетания, в которых элементы попарно различны.

2.6. Ограничение на принадлежность элемента списку

Ограничение может использоваться для отбора допустимых значений переменных, упомянутых в ограничении, но не для выборки подсети.

Должна упоминаться одна переменная-образец, которая проверяется на принадлежность списку. Ограничение может встречаться в запросе многократно. Может являться единственным в запросе на выборку узла, но не может быть единственным в запросе на выборку подсети. Область возможных значений переменной сужается до указанного в ограничении.

3. Задача наложения ограничений

Пусть необходимо построить множество троек $\{ \langle z, x, y \rangle \}$, например, тройки $\langle \text{фамилия, имя, отчество} \rangle$, имея следующие множества кортежей:

1. $\{ \langle x \rangle \}$ – множество кортежей с одним возможным значением переменной x (по сути, это вектор-столбец с именами);
2. $\{ \langle y \rangle \}$ – множество кортежей с одним возможным значением переменной y (вектор-столбец с отчествами);
3. $\{ \langle z \rangle \}$ – множество кортежей с одним возможным значением переменной z (вектор-столбец с фамилиями);
4. $\{ \langle z, x \rangle \}$ – множество двоек согласованных значений переменных z и x (фамилия, имя);
5. $\{ \langle x, y \rangle \}$ – множество двоек согласованных значений переменных x и y (имя, отчество).

Множеству кортежей будем ставить во взаимно-однозначное соответствие кортеж заголовков, показывающих порядок следования элементов кортежа, а также будем ставить в соответствие множество используемых уникальных заголовков. Например, множеству кортежей $A = \{ \langle x, y, z \rangle \}$ соответствует кортеж заголовков $\text{Head}(A) = \langle "x", "y", "z" \rangle$ и множество заголовков $\text{HeadSet}(A) = \{ "x", "y", "z" \}$. Множеству кортежей $B = \{ \langle x, x, y \rangle \}$ соответствует кортеж заголовков $\text{Head}(B) = \langle "x", "x", "y" \rangle$ и множество заголовков $\text{HeadSet}(B) = \{ "x", "y" \}$. Одно и то же множество заголовков может соответствовать разным множествам кортежей.

В примере множество выходных заголовков $\{ "y", "x", "z" \}$ имеет непустые пересечения с каждым из множеств заголовков для 1-5. В общем случае существует множество способов скомбинировать из имеющихся множеств кортежей множество кортежей с нужным заголовком, но не все способы будут порождать искомое множество кортежей.

Введём операцию проекции $\text{Project}(A, \text{Indices})$, которая возвращает множество укороченных кортежей, полученных из кортежей множества A путём удаления элементов с индексами, не указанными в Indices . Например, по определению $\text{Project}(\{\langle z, x \rangle\}, [1]) \rightarrow \{\langle z \rangle\}$, $\text{Project}(\{\langle z, x \rangle\}, [2]) \rightarrow \{\langle x \rangle\}$, $\text{Project}(\{\langle x, y, z \rangle\}, [2, 3]) \rightarrow \{\langle y, z \rangle\}$.

Следует отметить, что в общем случае $\text{Project}(\{\langle z, x \rangle\}, [2])$ может не совпадать с $\text{Project}(\{\langle x, y \rangle\}, [1])$, но множества $\{\langle z, x \rangle\}$ и $\{\langle x, y \rangle\}$ существуют так, что пары с необщими x отсутствуют, т.к. поддерживается общая система ограничений на значения связанных переменных.

3.1. Формирование множества кортежей $\{\langle x, y, z \rangle\}$ на основе имеющихся множеств кортежей без учёта требуемого порядка элементов

Рассмотрим варианты реализации.

Вариант 1. Декартово произведение $\{\langle x \rangle\} * \{\langle y \rangle\} * \{\langle z \rangle\}$. Это неверно, т.к. могут появиться сочетания x и z , отсутствующие в $\{\langle z, x \rangle\}$, равно как и сочетания x и y , отсутствующие в $\{\langle x, y \rangle\}$.

Вариант 2. Декартово произведение $\{\langle x, y \rangle\} * \{\langle z \rangle\}$. Неверно, т.к. могут появиться сочетания x и z , отсутствующие в $\{\langle z, x \rangle\}$. Симметричный вариант $\{\langle z, x \rangle\} * \{\langle y \rangle\}$ также неверен.

Вариант 3. $\text{Project}(\{\langle x, y \rangle\} * \{\langle z, x \rangle\}, [1, 2, 3])$. Неверно, т.к. разрушается согласованность между z и x . Тот же результат порождается и во втором варианте $\{\langle x, y \rangle\} * \{\langle z \rangle\}$.

Введём операцию эквисоединения $\text{NatJoin}(A, B)$, которая соединяет два множества кортежей A и B по общему подмножеству заголовков, порождая множество кортежей C такое, что:

- множество его заголовков равно объединению множеств заголовков соединяемых множеств: $\text{HeadSet}(C) = \text{HeadSet}(A) \cup \text{HeadSet}(B)$;
- заголовки упорядочены следующим образом: $\text{Head}(C) = \langle \text{Head}(A), \text{Head}(B) - \text{Head}(A) \rangle$, т.е.
 - сначала идут все заголовки из A ,
 - затем идут заголовки из B , отсутствующие в A .
- множество кортежей таково, что любые его проекции не шире, чем известные множества 1-5 с теми же заголовками.

Вариант 4. $\text{NatJoin}(\{\langle x, y \rangle\}, \{\langle z, x \rangle\})$.

Рассчитаем $\text{Head}(C) = \langle \text{Head}(\{\langle x, y \rangle\}), \text{Head}(\{\langle z, x \rangle\}) - \text{HeadSet}(\{\langle x, y \rangle\}) \rangle = \langle ["x", "y"], ["z", "x"] - ["x", "y"] \rangle = \langle ["x", "y"], ["z"] \rangle = \langle "x", "y", "z" \rangle$. Последний переход сделан с тем допущением, что оператор \diamond может конструировать кортеж не только из элементов, но и из списков/множеств/кортежей элементов, "распрямляя" структуры данных. В результате операции

эквисоединения по общему элементу x получаем множество C , которое включает только согласованные значения всех трёх переменных. Итак, вариант 4 верный.

Для упорядочивания элементов кортежа в нужном порядке введём операцию переупорядочивания $\text{Reorder}(A, \text{Indices})$, которая ставит элементы каждого кортежа множества A в порядке, указанном в списке Indices . Например, $\text{Reorder}(\{\langle y, x, z \rangle\}, [2, 1, 3]) \rightarrow \{\langle x, y, z \rangle\}$. Тогда для множества $C = \text{NatJoin}(\{\langle x, y \rangle\}, \{\langle z, x \rangle\})$ и $\text{Head}(C) = \langle "x", "y", "z" \rangle$ верно, что $\text{Reorder}(C, [3, 1, 2]) = \langle z, x, y \rangle$.

Обобщим задачу. Пусть требуется построить множество кортежей длины n , имея m множеств кортежей разной длины, причём в каждом из множеств используется свой набор переменных, которые могут как упоминаться в выходном множестве, так и не упоминаться.

3.2. Алгоритм построения оптимального порядка операций декартова произведения и эквисоединения

В основу алгоритма положим итеративную схему из двух шагов "выбор-соединение":

1. На этапе выбора будем оценивать меру полезности имеющихся множеств кортежей для наращивания результата.
2. На этапе соединения будем осуществлять комбинирование текущего результата с выбранным множеством кортежей с помощью операции эквисоединения или декартова произведения.

Критерий останова: множество заголовков результирующего множества кортежей включает все требуемые заголовки.

На этапе выбора каждое следующее множество кортежей должно наращивать целевой кортеж максимально длинными согласованными кортежами требуемых переменных, а с другой стороны, минимизировать появление лишних элементов. Положим меру полезности множества кортежей равной количеству использованных переменных, которые требуются, но ещё не включены в конечный результат.

Замечание 1. Пусть имеются $\{\langle x, y \rangle\}$, $\{\langle z, x \rangle\}$, $\{\langle z, x, m \rangle\}$, $\{\langle a, b, c, d \rangle\}$, а получить нужно $\{\langle x, y, z \rangle\}$. Множество $\{\langle a, b, c, d \rangle\}$ содержит самые длинные кортежи, однако не содержит ни одного из требуемых переменных, т.е. обладает нулевой полезностью. Множеств заголовков $\{"z", "x", "m"\}$ больше $\{"z", "x"\}$, но требуемых переменных в них поровну – 2. При этом первое множество, будучи выбранным, потребует дополнительной операции проекции, чтобы избавиться от лишней переменной m . Среди нескольких множеств заголовков, наращивающих имеющийся кортеж полезными заголовками одинаковым образом, стоит выбирать наименьшее множество.

Замечание 2. По мере построения результата полезность оставшихся множеств может снижаться, т.к. некоторые требуемые переменные уже будут включены в результат. Например, после первой итерации в результате отбора $\{<z, x>\}$ останется нарастить результирующие кортежи только переменной y , что снижает полезность множества $\{<x, y>\}$ с 2 до 1, а полезность $\{<z, x, m>\}$ – с 2 до 0.

Предлагаемый алгоритм:

1. Инициализировать остаточный список заголовков всеми требуемыми в запросе заголовками. В примере: $\text{RestHeadSet} \leftarrow \{ "x", "y", "z" \}$.
2. Инициализировать результирующее множество пустым множеством. $\text{Result} \leftarrow \emptyset$.
3. Пока остаточный список заголовков не опустеет ($\text{RestHeadSet} \neq \emptyset$):
 - a. Упорядочить множества кортежей по убыванию меры полезности. Для первой итерации это список множеств $\text{GivenHeadSets} \leftarrow [\{ "x", "y" \}, \{ "z", "x" \}, \{ "x" \}, \{ "y" \}, \{ "z" \}]$.
 - b. Извлечь первое множество заголовков $\text{HX} = \text{HeadSet}(X)$ из списка GivenHeadSets , вычеркнуть из RestHeadSet те заголовки, которые нашлись в HX .
 - c. Если множество заголовков результата пересекается с множеством заголовков HX , то выполнить эквисоединение: $\text{Result} \leftarrow \text{NatJoin}(\text{Result}, X)$, иначе выполнить декартово произведение: $\text{Result} \leftarrow \text{Result} * X$.

Такой алгоритм сделает два витка, выполнив операции $\text{NatJoin}(\emptyset * \{<x, y>\}, \{<z, x>\})$, и остановится. При этом кортеж заголовков будет выглядеть как $\langle "x", "y", "z" \rangle$.

3.3. Алгоритм эквисоединения двух множеств кортежей по общим столбцам

Предлагаемый алгоритм:

1. Вычислить результирующий список заголовков путём добавления в список заголовков первого множества заголовков второго множества, отсутствующих в первом: $\text{Head}(C) = \text{Head}(A) + (\text{Head}(B) - \text{Head}(A))$
2. Инициализировать список pairs – пар индексов соответствующих общих заголовков в списках заголовков.
3. Для каждого заголовка H из $\text{HeadSet}(A) \cap \text{HeadSet}(B)$:
 - a. $iA \leftarrow$ индекс H в списке $\text{Head}(A)$
 - b. $iB \leftarrow$ индекс H в списке $\text{Head}(B)$
 - c. Добавить в pairs пару $\langle iA, iB \rangle$
4. Вычислить список индексов элементов B , которые есть и в A : $\text{MutualB} \leftarrow \text{Project}[\text{pairs}, 2]$.

5. Инициализировать выходное множество кортежей пустым множеством: $C \leftarrow \emptyset$.
6. Для каждого кортежа a из A и каждого кортежа b из B :
 - a. Если все элементы в кортежах a и b с индексами из pairs , соответствующие общим заголовкам, равны, то добавить к выходному множеству кортежей C новый кортеж вида: $\langle a, \text{Project}[b, \text{MutualB}] \rangle$.
7. Вернуть множество кортежей со списком заголовков $\langle C, \text{Head}(C) \rangle$.

Осталось переупорядочить элементы кортежей требуемым образом. Для примера: $\langle "x", "y", "z" \rangle \rightarrow \langle "y", "x", "z" \rangle$. Этого можно добиться с помощью предписания вида: $[1, 2, 3] \rightarrow [2, 1, 3]$.

3.4. Алгоритм построения предписания для упорядочивания множества элементов в кортеж требуемой конфигурации

Пусть имеется множество кортежей $A = \{ \langle x, y, z \rangle \}$ и соответственно $\text{Head}(A) = \langle "x", "y", "z" \rangle$. Необходимо построить предписание для получения множества $B = \{ \langle y, x, z \rangle \}$ с $\text{Head}(B) = \langle "y", "x", "x", "z" \rangle$.

Возможный алгоритм, строящий подобные предписания порядка:

1. Инициализировать $\text{Indices} \leftarrow []$
2. Для каждого заголовка H из $\text{Head}B$:
 - a. $i \leftarrow$ индекс H в списке $\text{Head}A$
 - b. добавить в Indices найденный индекс i
3. Вернуть Indices

Например, $\text{Prescription}(\langle "y", "x", "z" \rangle, \langle "x", "y", "z" \rangle) = [2, 1, 3]$.

Эта же операция может быть использована, когда требуется несколько раз задействовать одну и ту же переменную. Например, $\text{Prescription}(\langle "x", "y", "z" \rangle, \langle "y", "x", "x", "z" \rangle) = [2, 1, 1, 3]$.

3.5. Синтез множества кортежей по имеющемуся множеству кортежей и предписанию порядка

Пусть имеется множество $A = \{ \langle x, y, z \rangle \}$ и предписание $\text{Indices} = [2, 1, 3]$. Необходимо построить множество $B = \{ \langle x, y, z \rangle \}$. Возможный алгоритм синтеза множества кортежей с требуемым порядком элементов:

1. Инициализировать множество кортежей $B \leftarrow \emptyset$
2. Для каждого кортежа a из A :
 - a. Инициализировать кортеж $b \leftarrow []$
 - b. Для каждого индекса i из Indices :
 - i. Нарастить справа кортеж b элементом $A[i]$
 - c. Добавить кортеж b во множество B
3. Вернуть B

Заключение

Описаны базовые типы запросов к РСС с некоторым набором ограничений, который может быть расширен при необходимости. Приведены алгоритмы работы интерпретатора. Пилотная реализация на языке Python с использованием парсера грамматик LARK² продемонстрировала применимость языка запросов к РСС для задачи глубинизации и трансформации синтаксических структур нескольких предложений в единую семантическую сеть. Подробное описание результатов применения разработанного языка запросов является предметом будущих работ.

Список литературы

- [Кузнецов, 1986] Кузнецов И.П. Семантические представления. – М.: Наука, 1986. – 296 с.
- [Кузнецов, 1988] Кузнецов И.П., Шарнин М.М., Пузанов В.В. Система обработки декларативных структур знаний "ДЕКЛАР-2". – М.: ИПИ РАН, 1988. – 81 с.
- [Кузнецов, 2007] Кузнецов И.П. Семантико-ориентированные системы на основе баз знаний. – М.: МТУСИ, 2007. – 173 с.
- [Попов, 1982] Попов Э.В. Общение с ЭВМ на естественном языке. – М.: Наука. Главная редакция физико-математической литературы, 1982. – 360 с.
- [Хендрикс, 1975] О расширении применимости семантических сетей введением разбиений // Труды IV Международной объединённой конференции по искусственному интеллекту. – 1975. – Т. 1. – С. 190-206.
- [Hentenryck, 1989] Pascal Van Hentenryck. Constraint Satisfaction in Logic Programming. Logic Programming Series. MIT Press, Cambridge, MA, 1989.
- [Kandefter et al., 2008] Michael Kandefter and Stuart C. Shapiro, Comparing SNePS with Topbraid/Pellet, SNeRG Technical Note 42, State University of New York at Buffalo, Buffalo, NY, July 18, 2008.
- [Marriott et al., 1998] Kim Marriott and Peter J. Stuckey. Programming with Constraints. MIT Press, 1998.
- [Segaran et al., 2009] Toby Segaran, Colin Evans, Jamie Taylor. Programming the Semantic Web. – O'Reilly Media, 2009. – 302 p.
- [Shapiro, 1979] Stuart C. Shapiro. The SNePS semantic network processing system. In N.V. Findler, editor // Associative Networks: The Representation and Use of Knowledge by Computers. – Academic Press, New York, 1979. – P. 179-203.

² https://lark-parser.readthedocs.io/en/latest/json_tutorial.html.

УДК 004.82: 620.9: 574.24

ИНЖЕНЕРИЯ ЗНАНИЙ В ИССЛЕДОВАНИЯХ УСТОЙЧИВОСТИ ЭНЕРГЕТИЧЕСКИХ И ЭКОЛОГИЧЕСКИХ СИСТЕМ*

Л.В. Массель (massel@isem.irk.ru)

А.Г. Массель (amassel@gmail.com)

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

В статье описано применение инженерии знаний в ходе исследований «упругой устойчивости» энергетических и экологических систем: онтологического и когнитивного инжиниринга. Показана система онтологий, интегрирующая онтологии объектов энергетики и онтологии влияния объектов энергетики на окружающую среду, приведены примеры онтологий и когнитивных карт. Описаны вычислительные эксперименты по оценке устойчивости энергетических и экологических систем, в том числе с применением методов машинного обучения.

Ключевые слова: инженерия знаний, онтологический и когнитивный инжиниринг, онтологии, когнитивные карты, вычислительный эксперимент.

Введение

Постановка задачи совместного исследования устойчивости (в смысле resilience) энергетических и социо-экологических систем была выполнена совместно с сотрудниками Международного института прикладного системного анализа (Австрия) [Массель и др., 2019], [Массель и др., 2020]. Это обусловлено тем, что исследования устойчивости энергетических и социо-экологических систем становятся все более актуальными, так как на первое место выходят не просто вопросы устойчивости в смысле стабильности этих систем (sustainability), а вопросы возможности и скорости возвращения этих систем в устойчивое состояние после возмущений, которым подвергаются эти системы (resilience) [Holling, 1996], [Davoudi, 2012], [Cen Nan et al., 2014], [Zhonglin Wang et al., 2017].

* Работа выполнена при частичной финансовой поддержке гранта РФФИ № 22-21-00841.

Направление, определяемое термином “Resilience”, который переводится на русский язык как “упругость”, или “упругая устойчивость”, в последнее время вызывает большой интерес. В России исследования в этой области ведутся в основном в области технической устойчивости, в то время как в Западной Европе рассматривают это направление шире и включают в рассмотрение также экологическую, психологическую, социальную и экономическую устойчивость. С другой стороны, факторы, определяющие социальную устойчивость в зарубежных работах, переключаются с факторами, используемыми при оценке качества жизни в российских исследованиях. Подходы к определению устойчивости (“Resilience”) подробно изложены в [Массель и др., 2019], анализ определений resilience выполнен в [Zhonglin Wang et al., 2017]. В представленном исследовании авторы придерживаются определения “Resilience”, приведенного Davoudi: «Устойчивость (Resilience) – это способность системы возвращаться к равновесию или устойчивому состоянию после возмущения, такого, как наводнения, землетрясения или другие стихийные бедствия, а также техногенные катастрофы, такие, как банковские кризисы, войны или революции» [Davoudi, 2012]. Согласно экологическому подходу, устойчивость – мера постоянства экосистем и их способности адаптироваться к изменениям и нарушениям и по-прежнему поддерживать одни и те же отношения между населением или государством [Holling, 1996].

Учитывая, что в исследованиях устойчивости (в смысле resilience) энергетических и социо-экологических систем большую роль играют неопределенность и плохоформализованные факторы, затруднительно использование в этих условиях строгих математических моделей. По этим причинам было предложено использовать на первом этапе исследований технологии искусственного интеллекта, в частности, семантическое моделирование, которое включает онтологическое, когнитивное, событийное и вероятностное (на основе Байесовских сетей доверия) моделирование, совместное применение которых требует инженерии знаний [Гаврилова и др., 2016], на втором этапе возможно подключение традиционного программного комплекса, в котором реализована экономико-математическая модель топливно-энергетического комплекса (ТЭК) России.

В статье рассмотрены результаты работы в этом направлении, выполненные при поддержке гранта РФФ №22-21-00841 «Адаптация и развитие методов и инструментальных средств ситуационного управления и искусственного интеллекта для исследований устойчивости энергетических и экологических систем».

1. Онтологический инжиниринг области энергетики и экологии

Под онтологическим инжинирингом понимается процесс проектирования и разработки онтологий, объединяющий две основные технологии проектирования сложных систем – объектно-ориентированный и структурный анализ. Он включает выявление основных классов сущностей в описании реальных взаимодействующих процессов, отношений между этими классами, а также совокупности свойств, которые определяют их изменение и поведение во взаимодействии [Гаврилова и др., 2016], [Гаврилова, эл. рес.].

Для целей проекта выполнена адаптация современной трактовки концепции ситуационного управления, предложенной в 1980-х гг. Д.А. Поспеловым [Поспелов, 1986]. Эта концепция использовалась авторами для исследований энергетической безопасности (ЭБ) и разработки интеллектуальной системы поддержки принятия стратегических решений в области энергетики.

В данном проекте предложено, по аналогии с исследованиями ЭБ, рассматривать и моделировать ситуации в шкале «норма-предкризис-кризис» (НПК), с учетом возмущений и предлагаемых мероприятий по их предотвращению (для критических – предкризисных ситуаций) или ликвидации последствий чрезвычайных – кризисных ситуаций [Пяткова и др., 2011]. В качестве критериев устойчивости энергетических систем используются индикаторы ЭБ, определяющие принадлежность ситуаций к интервалам шкалы НПК.

На первом этапе ситуации моделируются с использованием инструментальных средств семантического моделирования: онтологий, описывающих структуру ситуации (концепты и отношения между ними) и когнитивных моделей, описывающих причинно-следственные связи между концептами. Использование семантического моделирования позволяет получить качественную оценку ситуаций и возможностей возвращения системы в устойчивое состояние.

Причем совместное использование онтологического и когнитивного моделирования убирает дополнительную нагрузку с экспертов за счет того, что у экспертов появляется формализованная база знаний о предметной области в виде онтологий, используя которую эксперты могут строить когнитивные модели. Использование одного понятийного аппарата снимает часть субъективного восприятия при описании различных ситуаций.

На втором этапе для получения количественных оценок формируются сценарии возмущений, которые передаются в программный комплекс для прогнозирования развития топливно-энергетического комплекса (ТЭК) – ПК ИНТЭК-А, реинжиниринг которого выполнен в коллективе, возглав-

ляемом авторами [Массель и др., 2021]. Для интеграции математического и когнитивного моделирования в состав ПК включен агент семантического (когнитивного) моделирования

Определены критерии устойчивости, угрозы нарушений устойчивости энергетических и экологических систем на основе анализа угроз энергетической безопасности (ЭБ), в частности, в Байкальском регионе; сформирован перечень возможных чрезвычайных ситуаций в энергетике, вызываемых выделенными угрозами и управляющих воздействий, направленных на их предотвращение или устранение последствий ЧС; выполнен анализ видов ущербов от реализации природных и техногенных угроз (наступления критических и/или чрезвычайных ситуаций).

В результате проведённого онтологического инжиниринга была разработана система онтологий, интегрирующая онтологии объектов энергетики и онтологии влияния объектов энергетики на окружающую среду, представленная на рис. 1. На рис. 2 приведена онтология нарушений устойчивости энергетической и экологической систем Байкальского региона и управляющих воздействий по их предупреждению и/или ликвидации.

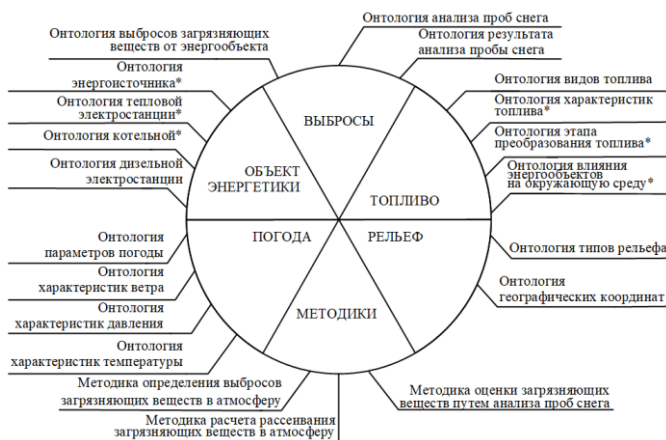


Рис. 1. Система онтологий, интегрирующая онтологии объектов энергетики и онтологии влияния объектов энергетики на окружающую среду

2. Когнитивный инжиниринг предметной области

По аналогии с онтологическим инжинирингом, под когнитивным инжинирингом понимается процесс проектирования и разработки когнитивных моделей (в графическом представлении – когнитивных карт), ото-

бражающих основные концепты и причинно-следственные отношения (связи) между ними [Трахтенгерц, 1998]. Связи могут дополняться весами (в простейшем случае +1 или -1). Разработана и применена методика интеграции онтологического, когнитивного и математического моделирования топливно-энергетического комплекса (ТЭК) в исследованиях устойчивости энергетических систем (ЭС), включающая следующие этапы:

- I. Формирование технико-экономической модели (в ПК ИНТЭК-А).
- II. Формирование системы онтологий для технико-экономической модели:
 - 2.1. Определение концептов, соответствующих элементам технико-экономической модели и элементам балансовых таблиц (автоматически).
 - 2.2. Формирование концептов факторов угроз.
 - 2.3. Установление связей между концептами.
- III. Формирование когнитивной карты угроз ЭБ района:
 - 3.1. Формирование концептов и связей при помощи онтологий (автоматически).
 - 3.2. Установление значений связей.
- IV. Расчет при помощи когнитивной карты:
 - 4.1. Расчет базового варианта ТЭК (автоматически).
 - 4.2. Корректировка технико-экономической модели при помощи когнитивной карты.
- V. Расчет варианта развития ТЭК (автоматически).
- VI. Интерпретация результата (автоматически).

Предложена усовершенствованная технология анализа когнитивных карт путем перевода когнитивной карты (КК) в продукционные правила экспертной системы (ЭС) и последующего анализа КК с использованием ЭС. Для поддержки методики и технологии выполнено развитие средства когнитивного моделирования CogMap: разработана подсистема трансформации когнитивных моделей в продукционные правила экспертной системы (ЭС) TCogMap, для автоматизации интерпретации когнитивных карт с помощью ЭС.

Перевод когнитивных карт в продукционные правила в свою очередь позволяет из когнитивной карты получить определенный сценарий, и облегчает в дальнейшем обработку этих сценариев.

На рис. 3 показана когнитивная карта, построенная для проведения одного из вычислительных экспериментов по оценке устойчивости энергетических систем.

Для оценки экологической устойчивости используется web-ориентированная информационно-вычислительная система WICS для оценки влияния объектов энергетики на окружающую среду. Она реализует методику, показанные в нижней части системы онтологий на рис. 1.

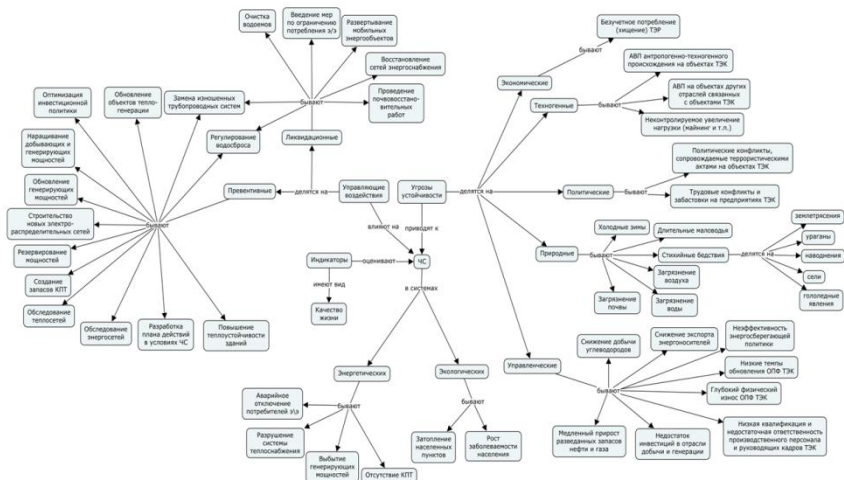


Рис. 2. Онтология нарушений устойчивости энергетической и экологической систем Байкальского региона и управляющих воздействий по их предупреждению и/или ликвидации

Для оценки экологической устойчивости используется web-ориентированная информационно-вычислительная система WICS для оценки влияния объектов энергетики на окружающую среду. Она реализует методики, показанные в нижней части системы онтологий на рис. 1.

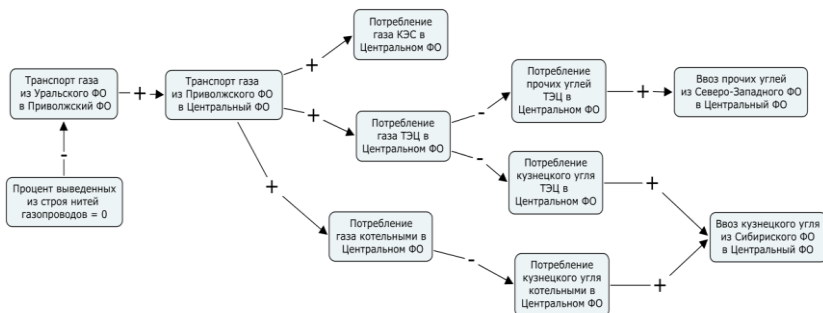


Рис. 3. Когнитивная карта снижения транспорта природного газа из Уральского федерального округа в Приволжский федеральный округ

3. Вычислительные эксперименты для оценки устойчивости энергетических и экологических систем

3.1. Моделирование критической ситуации в системе газоснабжения

Одной из наиболее опасных (по последствиям) является возможность повреждения трансконтинентальных газопроводов, идущих из Западной Сибири через Урал в европейскую часть России, поскольку природный газ в Европейской части России, на Урале и в Поволжье стал практически монопольным ресурсом в производстве электроэнергии и тепла.

В качестве расчетных условий критической ситуации в системе газоснабжения была рассмотрена возможность выхода из строя участка магистральных газопроводов, идущих через Урал из Западно-Сибирского газодобывающего района (северная нитка) (см. рис. 2).

Рассматривалось снижение транспорта природного газа на 28% из Уральского федерального округа в Приволжский федеральный округ. Значительные изменения в потреблении топливных ресурсов должны возникнуть в Центральном федеральном округе. Большинство ТЭЦ в Центральном федеральном округе используют в качестве топливного ресурса газ, большая часть газа добывается в Уральском федеральном округе, из Уральского ФО через Приволжский ФО газ транспортируется в Центральный ФО. На рис. 4 приведена когнитивная карта, соответствующая рис. 3, с указанием измененных значений концептов модели (по результатам ВЭ) в условных единицах

3.2. Оценка устойчивости электроэнергетической системы (ЭЭС) Сибирского федерального округа (СФО) на примере угрозы маловодья с применением методов машинного обучения

В последнее время широкое распространение получили методы машинного обучения, в том числе в энергетике [Mosavi, 2019]. В коллективе под руководством авторов было выполнено экспериментальное исследование с применением методов машинного обучения, со следующим примером.

Опасным природным явлением, способным нанести существенный ущерб производству электроэнергии в Сибирском ФО, является маловодье на Ангаро-Енисейском каскаде ГЭС. Выработка электроэнергии на этих станциях составляет в среднем 88 тыс МВт·ч в год или 46% от общей выработки электроэнергии в энергосистеме Сибири. В примере рассматривалась угроза маловодья, последствием которой является снижение генерации электроэнергии ГЭС на 0,976 млрд кВт·ч в год.

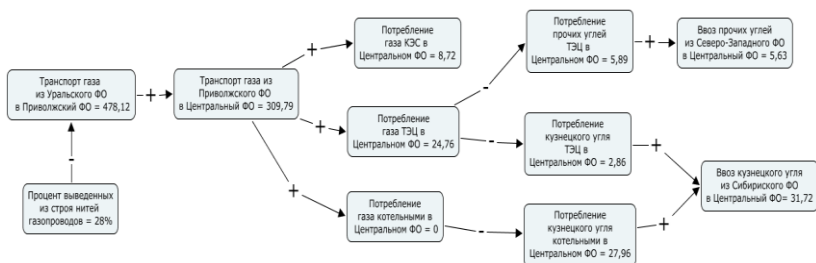


Рис. 4. Когнитивная карта, соответствующая рис. 3, с указанием измененных значений концептов модели (по результатам ВЭ) в условных единицах

Для обучающей выборки были взяты открытые данные Системного оператора ЕЭС, которые характеризуют потребление за период с 2017 год по 2021 год по Сибирскому федеральному округу. Обучающая выборка содержит значения суточного потребления электроэнергии в МВт·ч за период с 13.08.2017 по 31.12.2020, что соответствует 1207 объектам для обучения нейронной сети. Тестовая выборка составила 576 объектов. Выборки нормализованы в диапазоне от 0 до 1.

В примере использовались LSTM-сети. Программная реализация содержит 2 слоя LSTM и полносвязный слой. Первый и второй слои LSTM состоят из 128 нейронов каждый, среднеквадратическая ошибка используется в качестве функции потерь, а для уменьшения потерь используется оптимизатор “Adam”. Средства разработки программной реализации: язык Python.

Полученное прогнозируемое значение потребления электроэнергии Сибирским ФО равно 214,491 млрд кВт·ч. Реальное значение потребления за выбранный период равняется 217,546 млрд кВт·ч., при этом среднеквадратичное отклонение (RMSE) составляет 8887.09.

3.3. Оценка влияния объектов энергетики на окружающую среду в центральной экологической зоне (ЦЭЗ) Байкальской природной территории (БПТ) с использованием информационно-вычислительной системы WICS

Вычислительный эксперимент в рамках выполненных исследований был проведен с использованием сведений о 48 объектах энергетики с различными характеристиками, расположенных в ЦЭЗ БПТ и использующих в качестве топлива уголь (на каждом объекте используются различные виды). Метеосведения взяты из базы данных сервиса Gismeteo, частота обновления погодных данных – 12 часов, сведения о рельефе – из базы

данных сервиса Topocoding. Результаты анализа проб снега взяты из публикаций ЛИН СО РАН, ИГ СО РАН, БИП СО РАН, ФГБУ «Иркутское УГМС» и Государственных докладов.

Пример геовизуального представления результатов расчетов показан на рис. 5. Синими маркерами обозначены объекты энергетики. Чем краснее область на карте, тем выше выброс от рассматриваемых объектов.

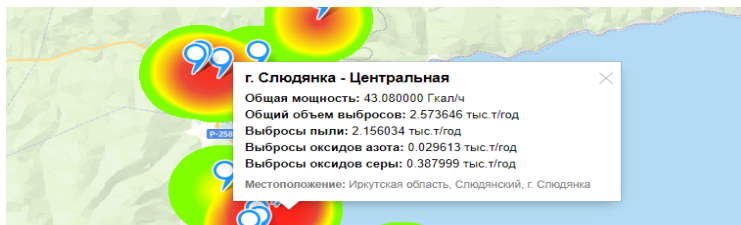


Рис. 4. Геовизуальное представление результатов расчетов выбросов в г. Слюдянка

На основе результатов вычислительного эксперимента сформулированы рекомендации для снижения негативного воздействия от объектов энергетики на окружающую среду, полученные на основании экспертных оценок, исходя из результатов исследования.

Заключение

В статье рассмотрено применение инженерии знаний в исследованиях упругой устойчивости (resilience) энергетических и экологических систем. Введены основные определения, выделены процессы онтологического и когнитивного моделирования, показаны примеры онтологий и когнитивных карт. Приведено описание вычислительных экспериментов с использованием описанной методики интеграции онтологического, когнитивного и математического моделирования ТЭК в исследованиях устойчивости энергетических систем: моделирование критической ситуации в системе газоснабжения; оценка устойчивости ЭЭС СФО) на примере угрозы маловодья с применением методов машинного обучения; оценка влияния объектов энергетики на окружающую среду в ЦЭЗ БПТ с использованием геовизуального представления результатов расчетов.

Список литературы

- [Гаврилова и др., 2016] Гаврилова Т.А., Кудрявцев Д.В., Муромцев Д.И. Инженерия знаний. Модели и методы: учебник. – СПб.: Изд-во «Лань», 2016. – 324 с.
- [Гаврилова, эл. рес.] Гаврилова Т.А. Онтологический инжиниринг. Электронный ресурс. Технологии менеджмента знаний. – Режим доступа: http://www.kmtec.ru/publications/library/authors/ontolog_engeneering.shtml (дата обращения: 20.05.2020).

- [Массель А.Г. и др., 2021] Массель А.Г., Мамедов Т.Г. Адаптация методики реинжиниринга унаследованных программных систем // Информационные и математические технологии в науке и управлении. – 2021. – № 4 (24). – С. 88-99.
- [Массель и др., 2019] Массель Л.В., Комендантова Н.П. Оценка рисков природных и техногенных угроз устойчивости энергетических, экологических и социальных систем на основе интеллектуальных информационных технологий // Информационные и математические технологии в науке и управлении. – 2019. – № 4 (16). – С. 31-45.
- [Массель и др., 2020] Массель Л.В., Массель А.Г., Комендантова Н.П. Подход к исследованиям устойчивости энергетических и экологических систем на основе интеллектуальных информационных технологий // Труды Международной научной конференции «Устойчивое развитие энергетики республики Беларусь: состояние и перспективы». – Минск: Беларус. навука, 2020. – С. 33-43. – ISBN 978-985-08-2654-1.
- [Массель и др., 2022] Массель Л.В., Массель А.Г., Мамедов Т.Г., Гаськова Д.А., Цыбиков А.Р., Шукин Н.И. Оценка устойчивости энергетических систем с применением методов машинного обучения // Информационные и математические технологии в науке и управлении. – 2022. – № 4 (28). – С. 248-260.
- [Поспелов, 1986] Поспелов Д.А. Ситуационное управление. Теория и практика. М.: Наука, 1986. – 284 с.
- [Пяткова и др., 2011] Пяткова Н.И., Рабчук В.И., Сендеров С.М., Чельцов М.Б. Энергетическая безопасность России: проблемы и пути решения / отв. ред. Воропай Н.И. – Новосибирск: Изд-во СО РАН, 2011. – 211 с.
- [Трахтенгерц, 1998] Трахтенгерц Э.А. Компьютерная поддержка принятия решений. – М.: СИНТЕГ, 1998. – 376 с.
- [Cen Nan et al., 2014] Cen Nan, Sansavini G., Kroeger W. Building an integrated metric for quantifying the resilience of interdependent infrastructure systems // 9th Intern. Conf. on Critical Information Infrastructure Security. Limassol. Cyprus. October 13-15. 2014. – 12 p.
- [Davoudi, 2012] Davoudi S. Resilience: A Bridging Concept or a Dead End // Planning Theory and Practice. – 2012. – 13 (2). – P. 299-307.
- [Holling, 1996] Holling C. Engineering Resilience Versus Ecological Resilience, Engineering Within Ecological Constraints, ed.: Peter Schultz, National Academy Press, Washington D.C. – 1996. – P. 31-43.
- [Mosavi, 2019] Mosavi A., Salimi M., Faizollahzadeh Ardabili S., Rabczuk T., Shamshirband S., & Varkonyi-Koczy A. State of the Art of Machine Learning Models in Energy Systems, a Systematic Review // Energies. – 2019. – 12 (7). – 1301.
- [Zhonglin Wang et al., 2017] Zhonglin Wang, Nistor M.S., Pickl S.W. Analysis of the definitions of resilience // 20th IFAC World Congress. Toulouse. France. July 9-14. 2017. – P. 11136-11144.

УДК 004.822

ОБЪЕКТНО-ОРИЕНТИРОВАННЫЙ ОНТОЛОГИЧЕСКИЙ ИНЖИНИРИНГ НА ОСНОВЕ МЕТА-АССОЦИАТИВНЫХ ГРАФОВ*

А.Е. Мисник (*anton@misnik.by*)

Белорусско-Российский университет, Могилев

В.В. Борисов (*vbor67@mail.ru*)

Филиал Национального исследовательского университета «МЭИ»
в г. Смоленске, Смоленск

Объектно-ориентированный подход является одним из наиболее распространенных подходов к разработке сложных кибер-физических систем. Он основан на представлении системы в виде набора взаимодействующих объектов, которые работают вместе для достижения целей системы. В контексте онтологического инжиниринга этот подход может быть особенно полезен, так как он упрощает описание и моделирование сложных кибер-физических систем, включающих множество взаимосвязанных объектов и понятий. Также объектно-ориентированный подход упрощает разработку и поддержку онтологий, которые могут использоваться для описания сложных кибер-физических систем. Онтология кибер-физической системы должна быть удобной и гибкой основой для моделирования процессов и обеспечения функционирования информационно-аналитических и системных процессов. Для построения онтологий системы и извлечения онтологий процессов может использоваться метаграф. Предлагается модификация метаграфа(мета-ассоциативный граф) для включения в онтологию событий и системных методов. Такой подход к построению метаграфа позволяет включать процессную составляющую в онтологическую модель системы как неотъемлемую ее часть. Это позволяет гибко и с меньшими временными затратами проектировать, актуализировать и поддерживать процессные модели на основе сегментов общей онтологической модели.

Ключевые слова: объектно-ориентированный подход, онтологический инжиниринг, метаграфы, мета-ассоциативные графы.

* Работа выполнена в рамках государственного задания, проект № FSWF-2023-0012.

Введение

Онтология – это формализованное описание понятий и их отношений в определенной области знаний, которое играет ключевую роль в обеспечении эффективной работы сложных кибер-физических систем. Онтология кибер-физической системы должна быть удобной и гибкой основой для моделирования процессов и обеспечения функционирования информационно-аналитических и системных процессов, что позволяет системе эффективно функционировать и адаптироваться к изменяющимся условиям.

Одной из основных проблем, возникающих при онтологическом инжиниринге и дальнейшей поддержке сложных кибер-физических систем, является необходимость быстрой реструктуризации и рекомбинации онтологии в зависимости от меняющихся условий. Это связано с тем, что сложные кибер-физические системы работают в реальном или квазиреальном времени, и задержки в перестройке онтологии могут привести к негативным последствиям для системы в целом.

Реструктуризация и рекомбинация онтологии необходимы при изменениях в области знаний, которую она описывает. Причиной могут быть появление новых сущностей или существенные изменения существующих, а также необходимость адаптации к новым условиям, не предусмотренным в исходной онтологии. Возможность таких изменений позволяет онтологиям более точно описывать систему и среду, обеспечивая надежную основу для функционирования процессов и принятия решений [Argpirez et al., 1998].

Использование объектно-ориентированного подхода для онтологического инжиниринга сложных кибер-физических систем является эффективным и гибким способом описания и управления данными и знаниями в системе. Он позволяет создавать модульную и легко повторно используемую онтологию, обеспечивает простую интеграцию с другими системами и, при необходимости, легко добавлять новые сущности, свойства и методы в существующую онтологию. Это особенно полезно, если онтология системы развивается и изменяется во времени.

Таким образом, использование объектно-ориентированного подхода при создании онтологии кибер-физической системы позволяет достичь гибкости и эффективности в описании и управлении данными и знаниями. Это также обеспечивает надежную основу для адаптации к изменяющимся условиям и эффективного функционирования процессов и принятия решений в реальном или квази-реальном времени.

1. Онтология кибер-физической системы как иерархическая структура

При работе со сложными кибер-физическими системами наблюдается тенденция к объединению разномодальных подсистем в единую систему. Для интеграции разномодальных систем необходимо использовать единый концептуальный интерфейс, который может быть представлен в виде онтологии.

Онтология кибер-физической системы представляет собой упорядоченные знания о её сущностях, их свойствах и отношениях между ними. Онтология может быть представлена в виде иерархической структуры, где каждый узел представляет сущность, а связи между узлами представляют отношения между сущностями. В такой иерархии все узлы имеют свою семантику и некоторые свойства, а также могут иметь отношения с другими узлами.

Иерархическая структура онтологии может быть древовидной или графовой. Однако, в древовидной структуре каждый узел имеет только одного родителя, в то время как в графовой структуре узлы могут иметь несколько родителей. Графовая структура более гибкая и позволяет представлять сложные отношения между сущностями, что предпочтительнее для формирования онтологий кибер-физических систем.

Иерархическую структуру можно расширять, добавляя новые узлы и связи между ними. Это позволяет уточнять понимание системы и углублять знания о ней. Однако также возможен пересмотр иерархической структуры, если новые данные или открытия противоречат ее текущей структуре. Все эти действия требуют наличия соответствующего модельного, алгоритмического и программного обеспечения.

Граф, отображающий онтологию сложной кибер-физической системы, может быть огромным и содержать множество узлов и связей. Поиск информации в такой онтологии может быть сложным и затратным по времени. Одной из проблем, связанных с поиском по онтологии сложной кибер-физической системы, представленной в виде графа, является необходимость понимания ее структуры и взаимосвязей между сущностями. Если структура онтологии неизвестна, то поиск информации и формирование представлений для просмотра может занимать весьма продолжительное время. Кроме того, онтология может содержать множество отношений между сущностями, что может затруднить понимание того, какие сущности связаны и как именно они связаны.

Таким образом, использование онтологии в качестве концептуального интерфейса для интеграции разномодальных систем в единую кибер-физическую систему может быть полезным. Однако, для эффективного использования онтологии необходимо понимание ее структуры и взаимо-

связей между сущностями. Расширение иерархической структуры и постоянное уточнение знаний о системе могут помочь углубить понимание системы и улучшить ее работу.

2. Объектно-ориентированный подход к онтологическому инжинирингу

Онтологический инжиниринг является процессом проектирования и разработки онтологий для объединения двух основных технологий проектирования сложных кибер-физических систем - объектно-ориентированного и структурного анализа. Он включает в себя выделение основных классов сущностей, описывающих реальные взаимодействующие процессы, отношения между этими классами и их свойства.

Целью онтологического инжиниринга является повышение уровня интеграции информации для принятия управленческих решений, эффективности информационного поиска и возможности совместной обработки знаний на основе единого семантического описания пространства знаний [Gomez-Perez et al., 2004].

Онтологический инжиниринг сложной кибер-физической системы как объектно-ориентированный подход следует разделить на три основные части: организация структуры данных и знаний в онтологическую структуру системы (классы), Онтологическая структуризация данных системы на основе онтологической структуры (объекты), и организация данных пользовательских интерфейсов, пригодных для реализации бизнес-процессов сложных кибер-физических систем, на основе как онтологической структуры системы, так и ее данных (представления) (рис. 1).

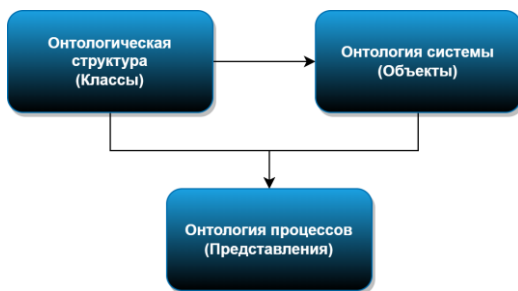


Рис. 1. Компоненты объектно-ориентированного подхода к онтологическому инжинирингу кибер-физической системы

Пользовательские интерфейсы бизнес-процессов сложной кибер-физической системы, обычно, также являются сложными, но каждый низкоуровневый узел такого интерфейса представляет собой достаточно про-

простое "плоское" представление данных – форму или таблицу. Такая простая структура может состоять из нескольких объектов разных классов, поэтому проблема надежного функционирования пользовательских интерфейсов является нетривиальной [Bobryakov et al., 2020], [Bobryakov et al., 2021], [Borisov et al., 2021].

3. Метаграфы как инструмент онтологического инжиниринга

Метаграфы являются обобщением представлений о структурах графов, используемых в различных областях. Они содержат элементы как орграфов, так и гиперграфов, и строятся на основе иерархического графа.

Метаграф представляет собой набор направленных отображений множества в множество, и его многообещающими приложениями являются моделирование отношений данных, где каждое отображение рассматривается как отображение набора ключевых элементов на набор элементов контента, а также моделирование задач рабочего процесса, где каждое отображение можно рассматривать как отображение набора элементов входных документов на набор выходных документов [Basu et al., 2007].

Метаграф определяется тремя множествами:

$$S = \{V, MV, E\},$$

где V – множество вершин метаграфа, MV – множество мета-вершин, E – множество ребер. Мета-вершина метаграфа представляет собой вершину, которая включает в себя множество вершин v , где $v \in V$ (рис. 2).

Каждая вершина метаграфа может быть атрибутивной, то есть с ней может быть связано произвольное количество атрибутов. Обозначим вершину метаграфа как v_i , и она будет содержать набор атрибутов A :

$$v_i = \{A\},$$

где $v_i \in V$ (рис. 2).

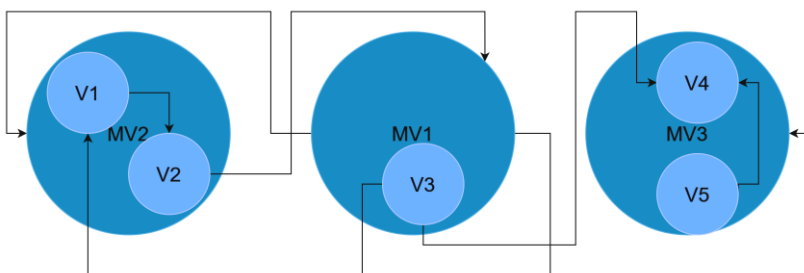


Рис. 2. Пример метаграфа

Таким образом, метаграфы представляют собой удобный инструмент для моделирования отношений данных и задач рабочего процесса. Они могут содержать как оргграфы, так и гиперграфы, и строятся на основе иерархического графа. Каждая вершина метаграфа может быть атрибутивной и содержать произвольное количество атрибутов [Chernenkiy et al., 2018], [Garanuyuk, 2019].

При описании онтологии сложной системы, метаграф с атрибутивными вершинами не всегда является идеальным инструментом. Требования к динамичности изменения структуры и параметров системы, с точки зрения онтологического инжиниринга процессов в сложной системе, могут сделать отличия между атрибутами и вершинами внутри мета-вершины достаточно неясными. В некоторых случаях вершина метаграфа может, фактически, являться вырожденным случаем рассмотрения мета-вершины без включенных в нее вершин.

Кроме того, классический метаграф не имеет естественных механизмов для идентификации событий и реакции на них, что может затруднять моделирование динамических процессов в сложных системах.

Предлагается обобщить понятия вершины и метавершины до понятия узла метаграфа (N) и дополнить определение узла именем, наборами событий и методов, принадлежащих узлу:

$$N = \{I, AS, EV, M\},$$

где I – имя узла, EV – набор событий, связанных с узлом метаграфа, M – набор методов, связанных с узлом метаграфа. AS , трансформируется из множества атрибутов A в множество ассоциативных атрибутов, являющихся либо обычными атрибутами, либо ссылками на узлы метаграфа:

$$AS = \{A, N^*\}.$$

Такая разновидность метаграфа позволяет увеличить гибкость онтологического инжиниринга, особенно при моделировании процессов в сложных системах. Ссылки на узлы метаграфа, на практике, мало чем отличаются от ребер метаграфа, что обеспечивает ассоциативные связи между узлами. Любая вершина может стать мета-вершиной, поскольку ссылка на другой узел метаграфа предоставляет все возможности для взаимодействия с ним [Misnik, 2022].

С другой стороны, если не требуется установка ассоциативной связи, ссылка может быть представлена в виде обычного атрибута с именем узла в качестве значения. Такой подход позволяет более гибко описывать онтологию и использовать метаграфы для моделирования сложных киберфизических систем (рис. 3).

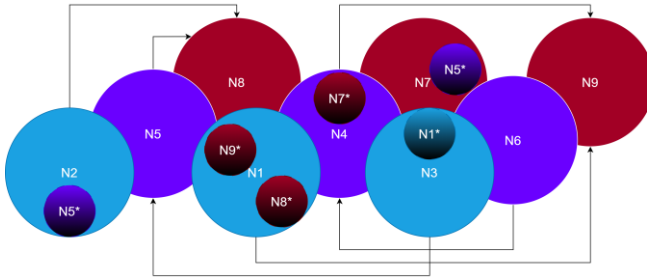


Рис. 3. Мета-ассоциативный граф

Такой подход к построению метаграфа дает возможность включать процессную составляющую в онтологическую модель системы как её неотъемлемую часть, что позволяет гибко и экономично формировать процессные модели на основе сегментов общей онтологической модели.

4. Специализированная программно-инструментальная среда для объектно-ориентированного онтологического инжиниринга

Перспективный подход к проектированию сложных кибер-физических систем направлен на уменьшение роли разработчиков программного обеспечения. Вместо программной реализации бизнес-процессов разработчики создают программно-инструментальную среду, которая позволяет экспертам проектировать онтологию и бизнес-процессы предметной области [Bobryakov et al., 2019], [Bobryakov et al., 2022], [Rasheed et al., 2019] (рис. 4).

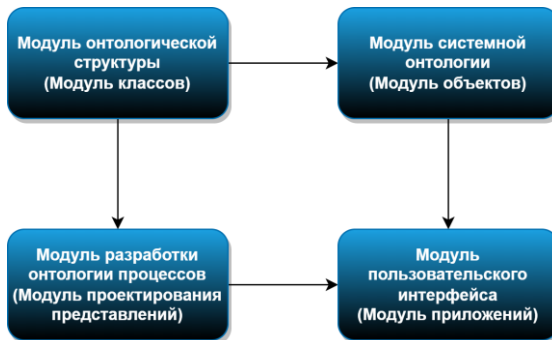


Рис. 4. Структура программно-инструментальной среды

Такой подход позволяет преодолеть семантический разрыв между экспертами, архитекторами и разработчиками программного обеспечения. Разработчики могут привлекаться только в случаях, когда требуется создание новых или доработка существующих инструментов среды.

Специалистам предметной области достаточно базовых навыков автоматизированного проектирования для создания онтологии предметной области или реализации бизнес-процессов сложной кибер-физической системы.

Для онтологического инжиниринга предметной области кибер-физической системы необходим модуль онтологической структуры, который позволяет спроектировать онтологию в виде мета-ассоциативного графа, включая процессную составляющую в виде методов (рис. 5).

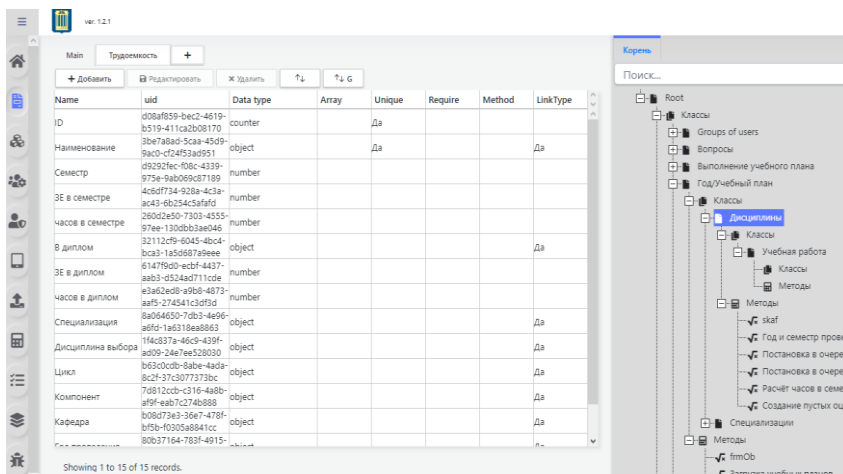


Рис. 5. Модуль онтологической структуры

Для реализации методов разработан язык программирования, который интегрируется с онтологией и позволяет использовать как классы онтологии так и её объекты (рис. 6).

Модуль системной онтологии служит контейнером данных и административным инструментом.

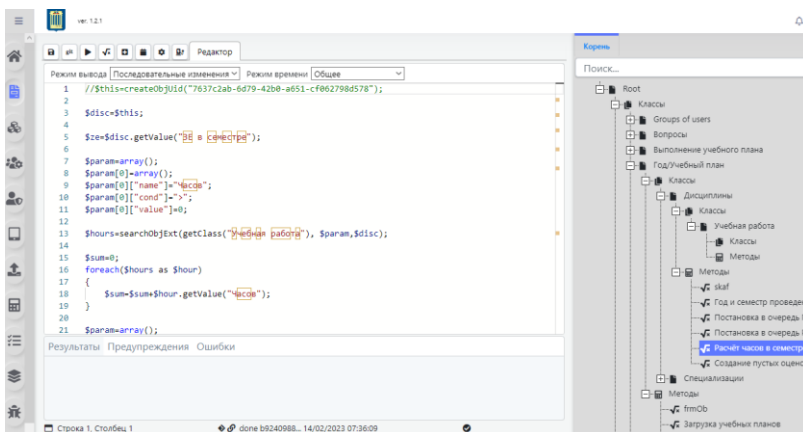


Рис. 6. Проектирование методов онтологии на основе разработанного языка программирования

Для разработки и реализации бизнес-логики сложной кибер-физической системы необходим модуль разработки пользовательского интерфейса, в который интегрированы возможности вызова методов онтологии. Одни и те же онтологические сущности могут быть организованы разными способами, что обеспечивает гибкость построения бизнес-процессов кибер-физической системы (рис. 7).

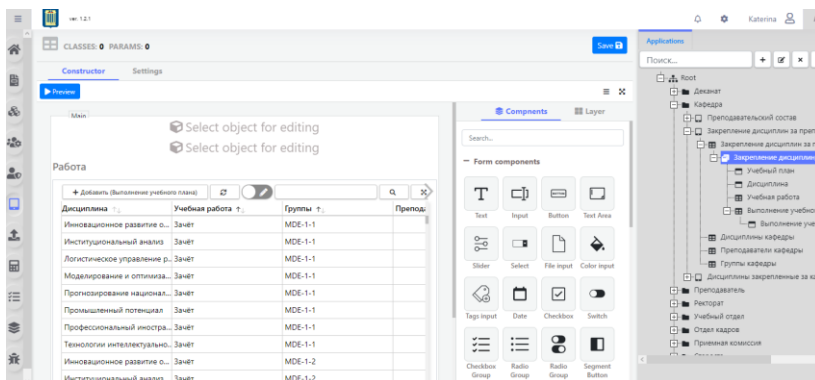


Рис. 7. Модуль разработки пользовательского интерфейса

Модуль пользовательского интерфейса необходим для взаимодействия пользователей с бизнес-процессами кибер-физической системы (рис. 8).

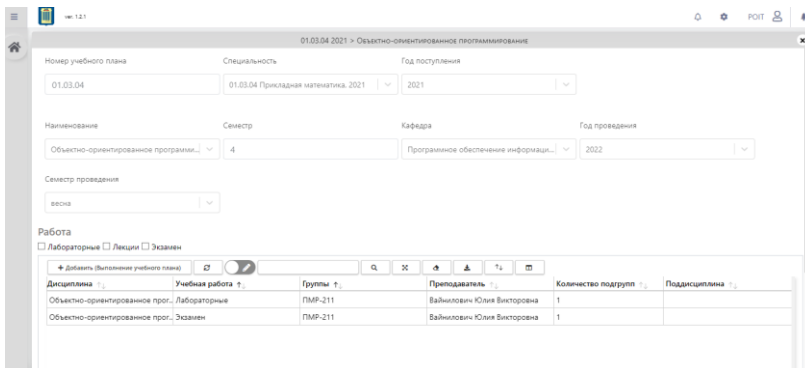


Рис. 8. Модуль пользовательского интерфейса

Заключение

Онтологический инжиниринг сложных кибер-физических систем является сложным и многогранным процессом, который требует использования различных методов и инструментов.

Рассмотрена комбинация объектно-ориентированного подхода и предложенной разновидности метаграфов – мета-ассоциативных графов, важной обладающих универсальной структурой и отличающихся возможностями интеграции процессной составляющей.

Использование мета-ассоциативных графов позволяет сократить время и усилия, затрачиваемые на создание и модификацию онтологий, а также обеспечить более наглядное и понятное представление сложных онтологических структур.

Для создания и разработки бизнес-процессов в сложных кибер-физических системах необходимы специальные программные инструменты. Описана разработанная программно-инструментальная среда для онтологического проектирования и разработки бизнес-процессов в сложных кибер-физических системах.

В настоящее время эффективность данного подхода и программно-инструментальной среды подтверждена более чем десятью успешными внедрениями в различных областях от медицины и образования до применений в промышленности.

Список литературы

[Arpirez et al., 1998] Arpirez J., Gomez- Perez A., Lozano A., Pinto S. (ONTO) 2Agent: An ontology- based WWW broker to select Ontologies // Workshop on Applications of ontologies and Problem Solving Methods. ECAI'98.

[Basu et al., 2007] Basu A., Blanning R. Metagraphs and their applications. – Springer, 2007. – 174 p.

- [**Bobryakov et al., 2020**] Bobryakov A.V., Yanukovich S.P., Zakharchenkov K.V., Borisov V.V. A Method for Managing Engineers Training Processes using Swarm Intelligence Algorithms // 2020 5th International Conference on Information Technologies in Engineering Education, Inforino 2020 – Moscow, Russia. – 2020. – DOI: 10.1109/Inforino48376.2020.9111788.
- [**Bobryakov et al., 2021**] Bobryakov A., Misnik A. and Prakapenka S. Management of industrial and technological processes of complex systems based on modified neuro-fuzzy Petri nets // CEUR Workshop Proceedings: Russian Advances in Fuzzy Systems and Soft Computing: Selected Contributions to the 10th International Conference on «Integrated Models and Soft Computing in Artificial Intelligence», IMSC 2021. – 2021. – Vol. 2965. – P. 276-283.
- [**Bobryakov et al., 2022**] Bobryakov A.V., Borisov V.V., Misnik A.E. and Prokopenko S.A. Design and Implementation of Information-Analytical and Industrial and Technological Processes in Production Based on Neuro-Fuzzy Petri Nets // 2022 VI International Conference on Information Technologies in Engineering Education (Inforino). – 2022. – P. 1-6. – DOI: 10.1109/Inforino53888.2022.9782997.
- [**Bobryakov et al., 2019**] Bobryakov A.V., Borisov V.V., Gavrilov A.I., Tikhonova E.A. Compositional Fuzzy Modeling of Energy- and Resource Saving in Socio-Technical Systems // EAI Endorsed Transactions on Energy Web and Information Technologies. – DOI: 10.4108/eai.12-9-2018.155863.
- [**Borisov et al., 2021**] Borisov V.V., Zakharchenkov K.V., Kutuzov V.V., Misnik A.E. and Prokopenko S.A. Modeling educational processes based on neuro-fuzzy temporal Petri nets // Applied Informatics. – 2021. – Vol. 16, No. 4. – P. 35-47. – DOI: 10.37791 / 2687-0649-2021-16-4-35-47.
- [**Chernenkiy et al., 2018**] Chernenkiy V.M., Gapanyuk Yu.E., Revunkov G.I., Andreev A.M., Kaganov Yu.T., Dunin I.V. The Principles and the Conceptual Architecture of the Metagraph Storage System / In: Manolopoulos, Ya., Stupnikov, S. (eds.) // 20th International Conference, DAMDID/RCDL 2018, Revised Selected Papers, CCIS. – Vol. 1003. – P. 73-87.
- [**Gapanyuk, 2019**] Gapanyuk Yu.E. Metagraph Approach to the Information-Analytical Systems Development // In: Proceedings of the 6th International Conference Actual Problems of System and Software Engineering, Moscow, Russia, 2019. – P. 428-439.
- [**Gomez-Perez et al., 2004**] Gomez-Perez A., Fernández-López M., Corcho O. Ontological Engineering: With Examples from the Areas of Knowledge Management, E-Commerce and the Semantic Web. – 2004.
- [**Soshnikov, 2000**] Soshnikov D. Software Toolkit for Building Embedded and Distributed Knowledge-Based Systems // In: Proceedings of the 2nd International Workshop on Computer Science and Information Technologies CSIT' 2000. – Ufa, 2000. – P. 103-111.
- [**Misnik, 2022**] Misnik A.E. Ontological Engineering on Metagraphs Basis // 2022 VI International Conference on Information Technologies in Engineering Education (Inforino). – 2022. – P. 1-6. – DOI: 10.1109/Inforino53888.2022.9782909.
- [**Rasheed et al., 2019**] Rasheed B., Popov A.Yu. Network graph datastore using DiSc processor // In: Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2019. – 2019. – P. 1582-1587.

УДК 004.8+004.65

РЕАЛИЗАЦИЯ ЗАПРОСОВ К ИЕРАРХИЧЕСКИМ ГРАФОВЫМ БАЗАМ ЗНАНИЙ И ДАННЫХ НА ПЛАТФОРМЕ IACRAAS*

Ф.М. Москаленко (*philipmm@iacp.dvo.ru*)

Институт автоматики и процессов управления ДВО РАН,
Владивосток

Дальневосточный федеральный университет, Владивосток

В работе представлены механизмы поддержки формирования и выполнения онтологических запросов для орграфовой связанной двухуровневой модели информационных единиц (баз знаний и данных). Применение предложенных подходов в процессе обработки баз знаний должно позволить задавать такие запросы без посредников (в лице программистов). Также описывается реализация предложенных решений на платформе IACRaaS, предназначенной для создания и использования облачных систем искусственного интеллекта.

Ключевые слова: онтология, графовая база данных, база знаний, редактор запросов, язык запросов, IACRaaS.

Введение

В современном мире всё большую важность получает деятельность по обработке накапливающихся массивов знаний и данных. Появились новые направления – data science, data mining. Специалисты в этой области должны снабжаться адекватными их задачам механизмами по работе с информацией. Одним из начальных шагов обработки данных является их извлечение из цифрового хранилища. Для этого могут применяться, например, конвертеры данных из баз данных (БД) в человеко-машинный формат. Для более продвинутых специалистов бывает достаточным наличие средства составления (в виде текста или, в более продвинутых случаях, в графическом виде, как, например, в MS Access) и исполнения так

* Работа выполнена в рамках госзаданий 0202-2021-0004 (разработка языка спецификации запросов к онтологическим базам знаний, апробация результатов) и FZNS-2023-0010 (разработка методов реализации исполнителя запросов к базам знаний).

называемых запросов к БД. Эти же механизмы обычно используются и программистами при разработке программных продуктов, использующих базы знаний (БЗ) и данных.

Средства последнего типа создаются по-возможности как можно более стандартизованными. Например, для реляционных БД создан язык SQL, задающий основные типы и формы запросов на получение и обработку информации. Далее каждая реализация может дополнять язык своими конструкциями (создавая свой диалект SQL), позволяющими использовать индивидуальные особенности реализации (например, язык PL-SQL для БД Oracle [Feuerstein et al., 2014]; язык Transact-SQL для БД Microsoft SQL Server [Dye et al., 2015]). Синтаксис таких языков зачастую прост и доступен пользователям, не являющимся разработчиками программных продуктов, однако, перед использованием язык необходимо изучить с целью ознакомления со всеми его выразительными средствами и возможностями. Также обстоит дело и в сфере БД, формат которых не является реляционным. Сюда относятся key-value БД (Redis; Riak; Amazon DynamoDB – SQL-совместимый язык PartiQL), базы документов (Apache CouchDB – запросы через HTTP, с применением JSON), графовые базы (neo4j – язык Cypher [Scifo, 2023]) и др. Здесь для каждого решения вводится свой язык. Можно отметить, что разработчики порой стараются приблизить его к синтаксису SQL.

С другой стороны с течением времени цифровизация затрагивает всё большее число отраслей деятельности человека. Квалификация экспертов в тех или иных предметных областях растёт – им нужны удобные средства накопления и сохранения знаний, а также формирования программных решателей задач, обрабатывающих такие данные. Развивающимся направлением разработки программного обеспечения является «программирование без программирования». То есть создание методов и инструментов, посредством которых эксперт, не являющийся разработчиком, мог бы решить свои задачи без написания кода на языках программирования. Языки и редакторы запросов, которыми дополняются некоторые движки баз данных, отчасти, могут служить таким средством.

Решением таких задач занимается коллектив разработчиков, которые в работах [Грибова и др., 2015a], [Грибова и др., 2015b] предложили особый подход к формированию информационных ресурсов (БД и БЗ), а также методы реализации средств накопления и обработки знаний, которые апробируются на платформе IASaaS [Грибова и др., 2022]. Применяемый формализм для представления информации – корневые размеченные иерархические однородные бинарные орграфы. Графовые БЗ и БД в такой форме предлагается формировать на основе двухуровневого подхода с использованием единого редактора. Реализация предлагаемых решений

выполнена на платформе IACPaaS (<https://iacpaas.dvo.ru>), которая предназначена для создания и использования СИИ и за более чем 10 летний срок применения и постоянной доработки продемонстрировала высокую эффективность и удобство использования. Двухуровневость подхода заключается в том, что вначале, с использованием специализированного метаязыка, формируется оргграф, описывающий метапонятия и их связь (модель онтологии предметной области), который представляет собой язык для описания оргграфов целевой информации (БД или БЗ).

Предложенный подход подразумевает не только онтолого-ориентированное конструирование взаимосвязанных сложно-структурированных информационных единиц, но и обеспечивает отслеживание выполнения соглашений, заложенных в онтологию инженером знаний, в процессе формирования знаний и информации экспертом [Москаленко и др., 2022]. Очередным шагом в расширении функциональных возможностей подхода и платформы является организация поддержки удобного пользователям-экспертам и инженерам знаний языка запросов и средства их задания и выполнения. Причём это должны быть такие средства, которые позволяли бы, в том числе, обрабатывать информацию без создания решателей задач, которые являются агрегациями программных агентов, использующих специальные программные интерфейсы для доступа к содержимому графовых баз знаний (API).

Для поддержки задания и выполнения запросов необходимо разработать:

1. Язык (онтология) спецификации запросов. Он граф, который содержит набор вершин, позволяющих описать тип запроса, его тело (в виде описания применяемых для вычисления *термов*, логических *формул* и пр.), параметры для различных выполнений запроса и результаты этих выполнений.

2. Метод исполнения запросов к графовым базам знаний и данных (сформированным согласно упомянутому подходу), составленных на предложенном языке, выполняющий сохранение результатов запроса в информационный ресурс, где содержится запрос, а также модификацию обрабатываемого информационного ресурса (в случае запроса модифицирующего типа).

3. Модель и методы реализации программного средства *Конструктор запросов*, поддерживающего задание, исполнение и отображение результатов запросов к графовым базам знаний и данных, сформированным согласно 2-уровневому подходу и используемому формализму.

Предложенные модели и методы реализуются на платформе IACPaaS.

1. Язык спецификации запросов

Модель языка запросов представляет собой модель онтологии, которая используется для формирования запросов (и результатов их исполнения) в виде иерархических графовых баз данных. Каждая база с запросами может содержать множество разноименных запросов, каждый из которых имеет *имя* (и может иметь описание в виде *комментария*), а его описание это тройка – *тело, формальные параметры, выполнения*.

Формальные параметры содержат именованные ссылки на корневые вершины ресурсов, описывающих модели онтологии ресурсов, над которыми может быть выполнен запрос, а также описание имён и типов значений параметров. Типом значения параметра может быть строковое, целое, вещественное, логическое, дата и время, бинарные данные, ссылка на вершину, множество однотипных значений (включая множество множеств).

Тело запроса описывается на языке логико-математических, теоретико-множественных и пр. формул, которые задаются над параметрами и той графовой грамматикой, по которой порождаются исследуемые оргграфы информации (графовые БЗ и БД). *Тело* запроса есть альтернатива: *создание, чтение, модификация, удаление*. В каждую из этих четырёх вершин в том или ином виде помещается структура *терм*, описание которой дано ранее в работе [Москаленко и др., 2022]. В структуру *терма* входят *переменная* и (логическая) *формула*.

Для запросов на *чтение* вершина, с которой начинается описание тела запроса, есть сам *терм*.

Для запроса типа *создание* под такой вершиной помещаются следующие две: *предки* и *создаваемое*. Под вершиной *предки* помещается ссылка введённую ранее вершину *переменная* – с её помощью задаётся множество вершин, для которых необходимо создавать новое содержимое. Под вершиной *создаваемое* указывается тип операции создания вершин(ы): *прямой предок, путь потомков* (более общий случай для *прямой предок* – здесь задаются шаги, состоящие из *прямых потомков*), *копия подсети, дуга*. Для первого типа создания и каждого шага второго типа указывается *мета-вершина потомка* (ссылка в оргграф модели онтологии) и *метка* новой вершины (вычисляемый *терм*). Для создания описания *копии подсети* в языке сделана специальная ссылка, которая позволит в самом запросе сослаться на заранее приготовленную исходную подсеть (её необходимо подготовить в некотором отдельном информационном ресурсе с требуемой метаинформацией). Для описания создания *дуги* язык потребует указать её *конец* – в виде вычисляемого *терма*, значением которого быть вершина.

Запрос типа *изменение* описывается двумя понятиями: *модифицируемое* и *изменение*. *Модифицируемое* есть *переменная* – здесь потребуются задать *терм-множество* [Москаленко и др., 2022], результатом вычисле-

ния которого должно стать множество вершин, для которых затем выполняется *изменение*. В качестве последнего предложены операции *установить метку* и *установить комментарий*. В качестве средства для задания этого используется опять же *терм*, значением которого должно стать примитивное данное (строка, число и т.п.).

Запрос типа *удаление*, выполняющий удаление определённых вершин (возможно, с потомками), дуг (возможно, с концом и потомками) или всё содержимое графовой базы данных, в качестве тела имеет соответственно три эти альтернативы. В случае необходимости удаления *вершин* – их множество указывается с помощью *терм-множество*. Таким же образом указывается множество вершин, являющихся начальными – при удалении *дуг*. При этом задаётся и один из трёх вариантов удаления:

- одну дугу, соединяющую начальную вершину с определённым её прямым потомком (вершиной, вычисляемой с помощью *терма*);
- все дуги, соединяющие начальную вершину с вершинами, являющейся её прямыми потомками и имеющими некоторую мета-вершину (задаётся ссылкой на мета-вершину);
- все дуги, исходящие из начальной вершины.

Результат выполнения запроса любого типа есть либо *терм* того или иного типа, либо некоторая *строковая константа* (сообщение о результатах выполнения). Каждому запросу может быть сопоставлено несколько *выполнений*, определяемых значениями *фактических параметров*. Каждому *выполнению* может быть сопоставлено несколько *результатов* – каждый *результат* получается отдельным запуском запроса на исполнение. В случае изменения БЗ или БД между запусками запроса на исполнение – разные *результаты* могут содержать разные полученные данные (при одних и тех же значениях параметров).

2. Метод исполнения запросов

Метод исполнения запросов, сформированных на предложенном языке, основывается на интерпретации структуры запроса (заданного в информационном ресурсе) и выполнении необходимых действий в зависимости от его типа и *тела*. При этом вначале либо в запросе должно быть сформировано *выполнение* со ссылками на обрабатываемые информационные ресурсы и заданными значениями параметров, либо такие данные поступают на вход методу и такое *выполнение* формируется самим этим методом. Затем определяется тип запроса и выполняется операция, соответствующая этому типу. Результат операции записывается под новой вершиной с метапонятием *результат* в существующее (в первом случае – см. выше) или новое (во втором случае) *выполнение*.

В случае запроса на *чтение* выполняются интерпретация *терма*: в зависимости от его типа (вычисляется в условии условного оператора, обрабатывающего вершину с метапонятием *терм*) выполняется вычислительное действие на основе указанных данных. Результат, сохраняемый под вершинами *выполнение/результат*, есть само вычисленное значение *терма* (строка, число, ссылка на вершину и т.д.).

В случае запроса на *создание* вначале вычисляется множество значений описанной *переменной*, если они есть множество вершин, то организуется цикл, перебирающий это множество, в теле которого выбранная вершина используется как предок, для которого формируется содержимое (*прямой потомок* по указанной *мета-вершине* с вычисляемой по *терму* меткой, *путь потомков* или же делается *копия подсети*, причём в качестве мета-вершины, по которой формируется корень подсети, используется мета-вершина исходного корня. Результат работы метода – созданная вершина – в случае прямого потомка, последняя созданная вершина – в случае пути потомков и вершина, являющаяся корнем скопированной подсети, – в случае копирования подсети.

В случае запроса на *изменение* вначале вычисляется множество значений описанной (в *модифицируемые вершины*) *переменной*, если они есть множество вершин, то организуется цикл, перебирающий это множество, в теле которого выбранной вершине устанавливается вычисляемая *метка* или *комментарий*. Результат работы метода – найденные вершины с сохранёнными изменениями.

В случае запроса на *удаление* вначале вычисляется множество значений *терма-множество* и, если они есть множество вершин, то организуется цикл, перебирающий это множество, в теле которого текущая вершина удаляется. При удалении некоторой вершины работает удаление всего поддерева, с неё начинающегося (согласно применяемой модели представления информации), поэтому может быть удалена и некоторая другая вершина из полученного множества – такая ситуация является нормальной и не приводит к ошибке. Результат выполнения метода – число найденных (и удалённых) вершин.

Описанные шаги выполнения запросов объединяются в метод `executeQuery` с двумя сигнатурами (см. начало раздела).

3. Конструктор запросов

Для визуального формирования запросов, их выполнения и отображения полученных результатов предлагается программный инструмент *Конструктор запросов*. Он должен позволять формировать и отображать запросы к иерархическим графовым БЗ и БД (с использованием онтологических терминов предметной области), инициировать исполнение этих

запросов и просматривать результаты их выполнения. Согласно модели онтологии запросов – в отдельной базе может быть сформировано несколько запросов, различающихся по именам. Для каждого запроса может быть сформировано множество вершин *выполнение*. Если запрос содержит *формальные параметры* – их фактические значения указываются под вершиной *выполнение* с помощью данного средства перед выполнением запроса. Для каждого такого набора значений параметров (отдельное *выполнение*) может быть инициировано множество выполнений запроса, причём каждый *результат* выполнения сохраняется, хранится и может быть просмотрен в любое время.

В качестве технологии реализации такого инструмента предлагается «Технология разработки сервисов по управляющей процессом решения задачи онтологии» [Грибова и др., 2017]. Специальный информационный ресурс *Таблица соответствия для конструктора запросов* (формируемая по модели онтологии *Структура таблицы соответствий*) создается и подключается к решателю *Конструктор запросов* в качестве *собственного*. Для реализации выполнения запроса разрабатывается агент *Исполнитель запросов* – он имеет блок продукций, отвечающий за создание *результата* запроса, где выполняется вызов метода выполнения запроса. В упомянутой выше таблице соответствий создано соответствие с мета-вершиной *результат*, которой сопоставлен этот агент.

4. Реализация и пример использования на платформе IASaaS

Формирование модели описанного языка для спецификации запросов выполнено на платформе IASaaS с использованием системного редактора моделей онтологий. Результирующая модель онтологии содержится в информационном ресурсе *Язык описания запросов*, который расположен в папке *Ядро платформы* в разделе *Платформа IASaaS* общего Фонда. Верхний уровень графа, представляющего собой этот язык, выглядит следующим образом (используется нотация текстового представления информационных ресурсов, используемая на платформе):

```
Язык описания запросов {
  ~setmm ~new запрос {
    ~new формальные параметры {
      ~new инфоресурсы {
        ~setmm ~new-clone имя параметра {
          ~one ~ref -> Λ$; #r ссылка на корень онтологии
        } }
      ~new данные {
        ~setmm ~new-clone имя параметра
        ~ALT{
          ~new ["строковое"]
          ~new ["целое"]
        }
      }
    }
  }
}
```

```

...
~new множество
~ALT {...}

} } }
~new тело запроса
~ALT {
  ~new чтение {
    терм
    ~ALT {...}
  }
  ~new создание {
    ~new предки {
      ~one ~new -> .../переменная;
    }
    ~new создаваемое
    ~ALT{
      ~new прямой потомок {...}
      ~new путь потомков {...}
      ~new копия подсети {...}
      ~new дуга {...}
    }
  }
  ~new изменение
  {
    ~new модифицируемые вершины {...}
    ~new изменяемое {...}
  }
  ~new удаление
  ~ALT{
    ~new вершины {...}
    ~new дуги {...}
    ~new инфоресурс {...}
  }
} }
~сорупм ~new выполнения {
  ~listmm ~new выполнение {
    ~new фактические параметры {...}
    ~new результаты {
      ~seqmm результат {...}
    }
  }
} } } } }

```

Реализация метода выполнения запросов выполнена на платформе IASaaS на языке Java в рамках *процессора информационных ресурсов*, а функции его вызова включены в предоставляемое платформой API разработчика (часть класса Agent):

1) Исполнить запрос (вычислить его уже готовое "выполнение" – сформированными параметрами):

```

IConcept ru.dvo.iacp.is.iacpaas.mas.Agent.executeQuery (
    IConcept queryExecutionHead,
    Object caller)

```

2) Исполнить запрос, сформировав новое выполнение с параметрами:

```
IConcept ru.dvo.iacp.is.iacpaas.mas.Agent.executeQuery (
    IConcept      queryHead,
    Map< String, IInforesource > inforesourceParams,
    Map< String, Object >      miscParams,
    Object caller )
```

При вызове в параметрах передается: (1) либо ссылка на *выполнение* заранее сформированного запроса (в некоторой графовой базе запросов), (2) либо ссылка на корневую вершину заранее сформированного запроса и два множества отображений, описывающих связь имён инфоресурсов и параметров с обрабатываемыми инфоресурсами и значениями параметров соответственно. Также передаётся ссылка на вызывающего агента – для проверки наличия доступа к обрабатываемым данным.

Результат выполнения реализованного метода есть ссылка на созданную вершину *результат*, формируемую при выполнении запроса. Описанные результаты выполнения операций помещаются в виде вершин, являющихся прямыми потомками этой вершины. При реализации использованы методы объектно-ориентированного программирования.

Описанный *Конструктор запросов* реализован на платформе IACPaas, имеющей инструментальную поддержку упомянутой в разделе 3 технологии разработки. Для этого был разработан специальный агент «Агент исполнитель запросов», выполняющий вызов метода вычисления запроса при формировании вершины с мета-вершиной «результат». Данная связь зафиксирована в разработанном информационном ресурсе «Таблица соответствия для конструктора запросов», подключенного к разработанному решателю задач «Конструктор запросов» в качестве «собственного на чтение».

Для демонстрации работы *Конструктора* рассмотрим пример из предметной области «медицинская диагностика». Здесь для решения интеллектуальных задач обычно вначале формируется «база знаний о наблюдениях», выступающая общим терминологическим словарём для баз знаний о заболеваниях (их диагностики и лечении), баз историй болезни и т.п. В качестве демонстрационного запроса к «базе знаний о наблюдениях» выступает запрос на чтение всех признаков из группы *Жалобы* информационного ресурса «База медицинской терминологии и наблюдений ДЕМО» (созданного по модели онтологии «Онтология базы медицинской терминологии и наблюдений»). Интерфейс *Конструктора*, отображающий ресурс «Тест запросов» с сформированным пользователем запросом и вычисленными результатами его выполнения (множество из 205 вершин с мета-вершиной *Признак*), представлен на рис. 1.

Вербально тело данного запроса может звучать следующим образом:
найти в информационном ресурсе «База Наб» все вершины с мета-вершиной «признак», которые являются непрямыми потомками всех вершин с мета-вершиной «группа признаков» и меткой «Жалобы».

В формальных параметрах указывается «тип» параметра «База Наб» – «Онтология базы терминологии и наблюдений», а при описании выполнения «1» задано его фактическое «значение» – «База медицинской терминологии и наблюдений ДЕМО».

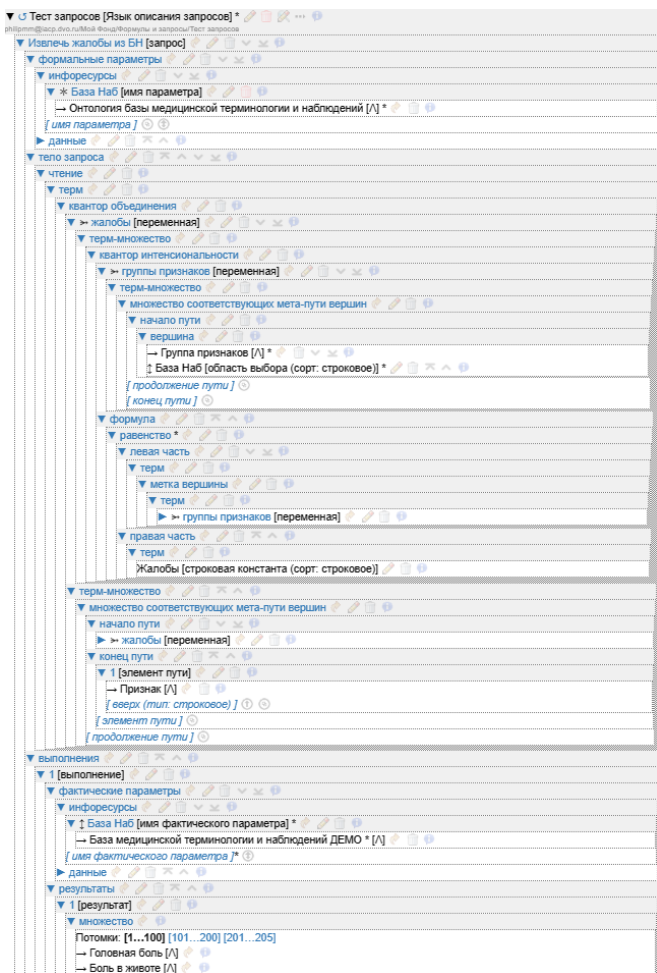


Рис. 1. Интерфейс Конструктора запросов на платформе IACPaas

Заключение

В работе даётся описание языка спецификации запросов для орграфовой связанной двухуровневой модели информационных единиц, метода выполнения запросов и реализации средства их конструирования и исполнения. Предложенные решения апробируются на облачной платформе IASaaS, предназначенной для разработки и использования облачных интеллектуальных сервисов. С использованием разработанного языка здесь можно описать CRUD (create-read-update-delete) запросы для графовых БЗ и БД. Инженер знаний, эксперт, дата-саентист при помощи разработанного решателя *Конструктор запросов* может в информационном ресурсе сформировать необходимые запросы к своим онтологическим информационным ресурсам в той или иной предметной области – в виде орграфа. При этом используется тот же процесс формирования орграфа, что и при создании самих БЗ и БД, что существенно снижает «порог входа». Модель онтологии *Язык описания запросов* управляет работой *Конструктора запросов* и позволяет формировать только корректные (с точки зрения модели онтологии) запросы в структурном виде. Это устраняет необходимость изучать язык запросов с целью формирования таковых в текстовом виде, как это делается в множестве существующих решений. Тем самым снижается возможность допустить ошибку в процессе составления *тела* запроса.

Вместе с тем, есть несколько путей дальнейшего развития *Конструктора* и применяемого подхода. Например, в текущей реализации структура *термножество* применяется слишком общо в модифицирующих запросах. Результатом вычисления такого *терма* могут быть не только вершины, но и примитивные данные (строки, числа и т.п.), которые, по сути, не могут быть модифицированы. Во избежание создания запроса с таким некорректным *телом* необходимо ввести вспомогательный термин *множество вершин* в модели онтологии. Помимо этого некоторые конструкции языка (заимствованные из математического диалекта), возможно, стоит переименовать, используя общепринятые термины, позаимствовав их из терминологии SQL.

Список литературы

- [Dye et al., 2015] Dye et al. SQL Server T-SQL Recipes. – Apress, 2015. – 972 p.
- [Feuerstein et al., 2014] Steven Feuerstein, Bill Pribyl, Oracle PL/SQL Programming: Covers Versions Through Oracle Database 12c. O'Reilly Media, Inc., 2014. – 1382 p.
- [Scifo, 2023] Estelle Scifo, Graph Data Science with Neo4j: Learn how to use Neo4j 5 with Graph Data Science library 2.0 and its Python driver for your project. – Packt Publishing, 2023. – 288 p.

- [Грибова и др., 2015a] Грибова В.В., Клещев А.С., Москаленко Ф.М., Тимченко В.А. Двухуровневая модель сложноструктурированных информационных единиц, соответствующая метафоре анкетирования // Научно-техническая информация. Сер. 2. – 2015. – № 10. – С. 1-10.
- [Грибова и др., 2015b] Грибова В.В., Клещев А.С., Москаленко Ф.М., Тимченко В.А. Модель порождения орграфов информации по орграфу метайнформации для двухуровневой модели сложноструктурированных информационных единиц // Научно-техническая информация. Сер. 2. – 2015. – № 12. – С. 26-38.
- [Грибова и др., 2017] Грибова В.В., Клещев А.С., Москаленко Ф.М., Тимченко В.А., Федорищев Л.А., Шалфеева Е.А., Управляемая графовыми грамматиками разработка оболочек интеллектуальных сервисов на облачной платформе IASaaS // Программная инженерия. – 2017. – № 10. – С. 435-447.
- [Грибова и др., 2022] Грибова В.В., Москаленко Ф.М., Тимченко В.А., Шалфеева Е.А. Платформа IASaaS для разработки систем на основе онтологий: десятилетие использования // Искусственный интеллект и принятие решений. – 2022. – № 4. – С. 55-65.
- [Москаленко и др., 2022] Москаленко Ф.М., Тимченко В.А. Формирование баз знаний с учётом онтологических соглашений на платформе IASaaS // Двадцатая национальная конференция по искусственному интеллекту с международным участием, КИИ-2022, (Москва, 21-23 декабря 2022г.): Труды конференции. В 2 т. Т. 1. – М.: Изд-во МЭИ, 2022. – С. 157-176.

ГЕНЕРАЦИЯ СОБЫТИЙНЫХ ДАННЫХ НА ОСНОВЕ ОНТОЛОГИЧЕСКОГО ПОДХОДА*

А.М. Наместников (*am.namestnikov@gmail.com*)

Ульяновский государственный технический университет,
Ульяновск

В настоящее время активно развивается научное направление process mining (интеллектуальный анализ процессов), который позволяет решать задачи машинного обучения, применительно к событийным данным. В статье приводится формальная модель прикладной онтологии, которая применяется для формирования журнала событий по результатам имитационных экспериментов с дискретно-событийной моделью. Дается описание архитектуры разработанной онтологической системы, структура основных понятий с соответствующими семантическими отношениями и наборами экземпляров.

Ключевые слова: онтология, событийные данные, процесс, имитационная модель.

Введение

Интеллектуальный анализ процессов (ИАП) (англ. – process mining) является областью исследований, которая находится на пересечении вычислительного интеллекта, интеллектуального анализа данных и моделирования и анализа процессов [Aalst, 2016], [Pegoraro, 2021]. В последнее время рост интереса исследователей к задачам process mining привел к созданию рабочей группы Task Force for Process Mining [Aalst, 2016].

Имитационные модели сложных систем позволяют сохранять в журналах событий важную информацию для последующего анализа их поведения. В свою очередь, методы ИАП могут иметь доступ к событиям, записанным в указанных журналах. Каждое событие относится к определенному действию и является частью отдельного экземпляра процесса

* Работа выполнена при финансовой поддержке государственного задания № 075-00233-20-05 по проекту «Исследование интеллектуального предиктивного мультимодального анализа больших данных и извлечения знаний из разных источников».

(кейс, англ. – case). Для включения дополнительной информации в журналах событий могут содержаться дополнительные атрибуты: данные о ресурсах, временные метки и другие.

Имитационное моделирование является очевидным подходом к генерации событийных данных. Основная сложность заключается в необходимости трансформации событийных данных из контекста имитационного моделирования (где применяются такие понятия, как агент, сообщение, состояние и т.д.) в контекст process mining (включающего понятия экземпляра процесса, действия, события, ресурса и т.д.).

Прикладная онтология, как средство представления знаний, способна выполнять функцию семантического медиатора, определяя семантические связи между понятиями имитационного моделирования и понятиями ИАП.

1. Онтологическая модель событийных данных

Онтология событийных данных, генерируемых имитационной моделью, включает в себя следующие компоненты:

$$O_D = \langle C_D, R_D, F_D, I_D \rangle, \quad (1)$$

где C_D – множество понятий предметной области, R_D – множество семантических отношений онтологии, F_D – множество функций интерпретации, позволяющие формировать новые факты базы знаний, I_D – множество экземпляров (индивидов), относящиеся к понятиям онтологии.

Множество понятий (концептов) C_D включает четыре компонента знаний о терминологии предметной области моделирования и анализа иерархических систем:

$$C_D = \{C_D^{St}, C_D^{Ms}, C_D^{Im}, C_D^{Pm}\}, \quad (2)$$

где C_D^{St} – множество понятий диаграмм состояний имитационной модели; C_D^{Ms} – множество понятий, включаемых в процессы обмена сообщениями между агентами модели; C_D^{Im} – множество понятий, описывающих структурные элементы имитационной модели; C_D^{Pm} – множество понятий, представляющих структурные особенности журнала событий в контексте ИАП.

Аналогичным образом определяются множества семантических отношений $R_D = \{R_D^{St}, R_D^{Ms}, R_D^{Im}, R_D^{Pm}\}$ и множество функций интерпретации $F_D = \{F_D^{St}, F_D^{Ms}, F_D^{Im}, F_D^{Pm}\}$ [Загоруйко, 2004].

Место онтологии событийных данных представлено на рис. 1. Журнал событий так, как он понимается в контексте ИАП, формируется на основе трех видов журналов имитационного моделирования:

- журнал обмена сообщениями;
- журнал состояний агентов;
- журнал прохождения агентов по имитационной модели (на рисунке – ИМ).



Рис. 1. Место прикладной онтологии в процессе формирования журнала событий

Состав журнала событий включает информацию об экземпляре процесса. В данной работе под процессом будем понимать конкретную реализацию процесса обработки задачи в иерархической системе. Идентификатор события определяет уникальный номер строки журнала событий. Метка времени является обязательной для процедуры упорядочивания событий. Деятельность представляет функцию преобразования данных в процессе обработки задачи. Дополнительно журнал событий может включать атрибуты, которые представляют информацию о субъектах деятельности, стоимости или других параметрах, которые относятся к конкретному событию.

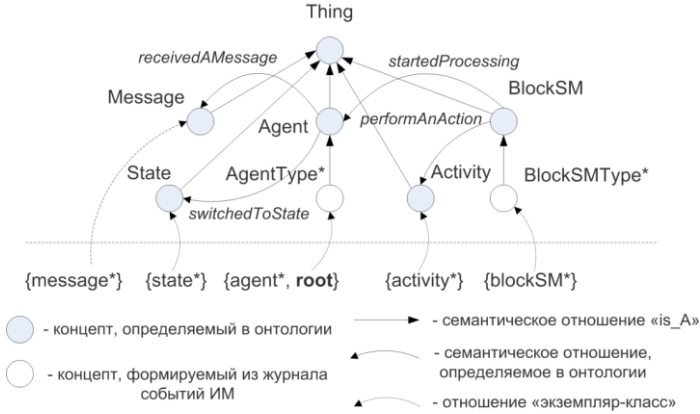


Рис. 2. Структура онтологии событийных данных

На рис. 2 представлена структура онтологии событийных данных, используемая для формирования журнала событий. Множество понятий включает в себя два подмножества:

$$C_D = C_D^{fix} \cup C_D^{gen},$$

где C_D^{fix} – множество понятий предметной области, которое формируется экспертным способом (вручную) и является инвариантным относительно конкретной имитационной модели; C_D^{gen} – генерируемое множество понятий предметной области на основе содержимого событийных журналов конкретной имитационной модели.

Обозначение класса с символом (*) поясняет тот факт, что мы имеем дело с несколькими классами, количество которых заранее неизвестно. На рис. 2 такими понятиями являются $AgentType^*$ и $BlockSMType^*$. Индивиды понятий онтологии предметной области на рисунке представлены с помощью множеств $\{message^*\}$, $\{state^*\}$, $\{agent^*, root\}$, $\{activity^*\}$ и $\{blockSM^*\}$. Здесь индивид *root* выделен среди остальных индивидов как предопределенный (другими словами, он присутствует в каждой имитационной модели в любом случае).

Концепт «Агент» запишем в следующем виде:

$$Agent \equiv T \prod \exists receivedMessage.Message \prod \exists switchedToState.State,$$

где *receivedMessage* и *switchedToState* – наименование ролей «получено сообщение» и «перешел в состояние», соответственно; *Message* – домен типа «Сообщение», *State* – домен типа «Состояние».

Концепт «Блок имитационной модели» запишем в виде логического выражения онтологии:

$$BlockSM \equiv T \sqcap \exists startedProcessing.Agent \sqcap \exists performAnAction.Activity,$$

где *startedProcessing* и *performAnAction* – наименование ролей «начата обработка» и «выполнение действия», соответственно; *Agent* – домен типа «Агент», *Activity* – домен типа «Действие».

Логический вывод по онтологии выполняется для решения следующих задач:

- Все экземпляры процессов, которые подвергаются анализу, должны быть полностью завершены. Для этого определим отдельный класс онтологии *CompleteAgent*, который является подклассом понятия *Agent* и содержит только те экземпляры (конкретные агенты, появляющиеся в имитационной модели), которые получили как начальное сообщение, так и конечное сообщение:

$$\begin{aligned} CompleteAgent \equiv \\ Agent \sqcap \exists receivedMessage.StartMessage \sqcap \\ \exists receivedMessage.FinalMessage, \end{aligned}$$

где *StartMessage* – подкласс понятия *Message*, который содержит сообщения, инициирующие создание агента, *FinalMessage* – подкласс понятия *Message*, который содержит сообщения, относящиеся к моменту удаления агента из имитационной модели.

Рассматривая агентов с полным жизненным циклом мы имеем возможность учитывать только те фрагменты журнала обмена сообщениями, которые описывают полные экземпляры процессов.

- Не все блоки имитационной модели (экземпляры класса *BlockSM*) имеют отношение к моделированию пользовательских действий по обработке заданий. Некоторые из блоков носят исключительно системный характер и соответствующие действия модели не должны быть включены в процедуру анализа поведения системы. Для исключений действий системных блоков имитационной модели создан класс онтологии *SystemActivity*, который является подклассом понятия *Activity* и содержит только те действия, которые выполняют системную функцию в имитационной модели (которые в последующем анализе не будут включены в поведенческую модель процесса):

$$SystemActivity \equiv Activity \sqcap \forall performedBy.SystemBlockSM,$$

где свойство *performedBy* является обратным свойством для *performAnAction*, концепт *SystemBlockSM* – подкласс концепта *BlockSM*. В данном случае используется универсальное ограничение на класс онтологии.

Процесс онтологического формирования журнала событий модели иерархической системы включает в себя следующие шаги:

Шаг 1. Доопределение множества понятий онтологии (1) C_D^{fix} понятиями C_D^{gen} . Для этого необходимая информация извлекается из журнала обмена сообщениями, журнала состояний агентов и журнала прохождения агентов по имитационной модели.

Шаг 2. Доопределение онтологии фактической информацией. На данном шаге каждое событие из журналов, анализируемых на предыдущем шаге, преобразуется в факты вида:

instance : *Class* И $\langle instance1, instance2 \rangle$: *Relation*.

Таким образом, множество понятий $C_D = C_D^{fix} \cup C_D^{gen}$ с соответствующими семантическими отношениями и функциями интерпретации, а также содержимое базы фактов, сформированное на данном шаге, определяют то содержимое онтологии, которое необходимо для выполнения логического вывода.

Шаг 3. Выполнение логического вывода по базе знаний с целью формирования класса событий, которые являются инициаторами действий, которые, в свою очередь, входят в состав событий, как экземпляров соответствующего журнала.

Шаг 4. Восстановление $\langle instance, event, timestamp, activity \rangle$ как кортежа журнала событий для последующего анализа.

2. Реализация онтологической системы генерации событийных данных

Архитектура системы генерации событийных данных приведена на рис. 3 и включает в себя следующие подсистемы:

- Подсистема онтологического моделирования.
- Подсистема генерации журнала событий.

В основе подсистемы онтологического моделирования находятся: библиотека OWLReady 2, RDF-хранилище, реализованное с помощью SQL-базы данных SQLite 3, механизм логического вывода HermiT. Среда выполнения OWLReady 2 позволяет оперировать классами (понятиями) онтологии, экземплярами (индивидами) и свойствами [Гуськов, 2021]. Указанные структуры сохраняются в RDF-хранилище [Beckett, 2001], [Bertails, 2012], [Brickey, 2000]. Реализация данной подсистемы позволяет выполнять сохранение онтологии в формате RDF/XML [Hayes, 2001] с последующей загрузкой ее в редактор Protege. Соответственно, выполне-

ние логического вывода возможно как непосредственно в среде OWLReady, так и средствами онтологического редактора.

Основной задачей *подсистемы онтологического моделирования* является формирование семантической модели предметной области, которая определяет ограничения для генерации журнала событий.

Подсистема генерации журнала событий состоит из двух основных модулей (рис. 3):

- Модуль анализа ограничений.
- Модуль формирования событийной структуры.

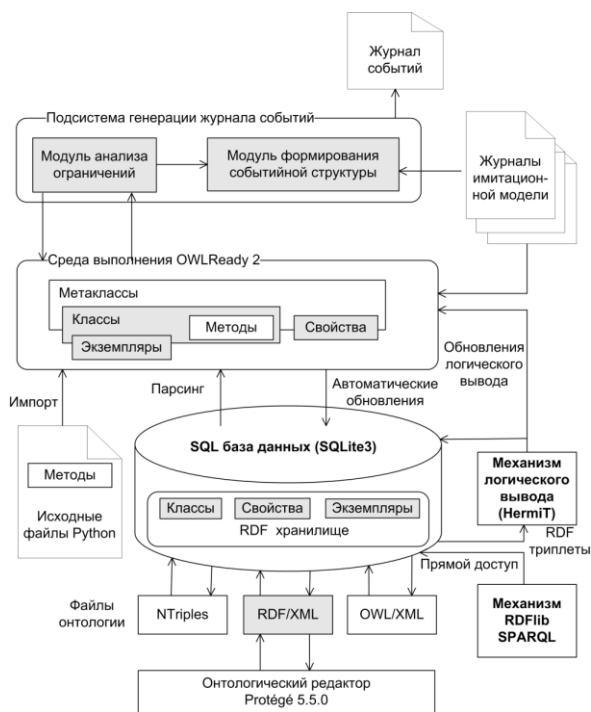


Рис. 3. Архитектура системы генерации событийных данных

3. Вычислительный эксперимент формирования журнала событий

В качестве примера системы, поведение которой представляется в виде журнала событий так, как он понимается в теории process mining, в статье рассматривается иерархическая система обработки задач. Для каждой

задачи определяется ее сложность, которая влияет на то, на каком уровне иерархии системы она будет решаться. Чем сложнее задача, тем выше по иерархии располагается уровень ее решения.

Построенная имитационная модель учитывает, что узлы иерархической системы могут находиться в нерабочем состоянии (в модели для этих целей используются стохастические параметры выхода из строя узлов системы).

На рис. 4 представлен пример сформированной онтологии после ее автоматического доопределения из журналов имитации иерархической системы решения задачи.

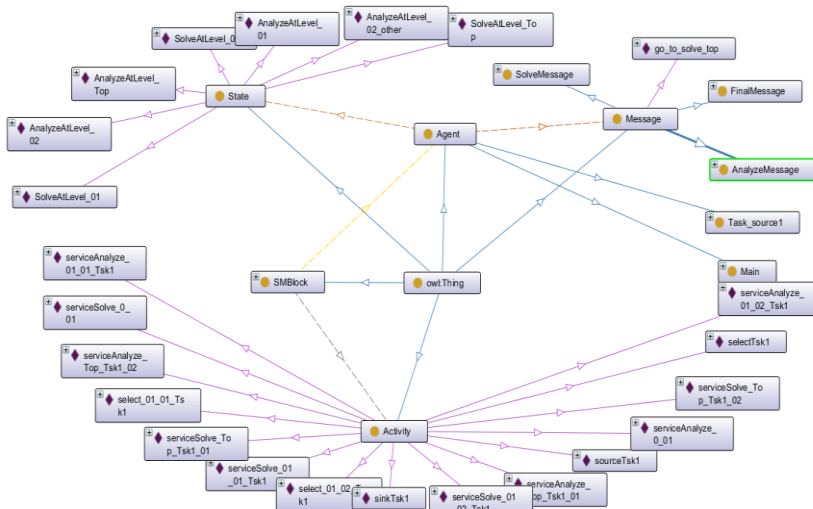


Рис. 4. Состояние онтологии после ее автоматического доопределения

Рассмотрим пример формирования журнала событий на основе журналов имитационной модели иерархической системы обработки задач. Были выполнены прогоны имитационной модели в среде AnyLogic PLE. Табл. 1 содержит подмножество строк журнала обмена сообщениями, соответствующее нескольким вариантам решения задачи. Каждая задача в имитационной модели представляется в виде агента.

В табл. 1 содержатся данные о получателе сообщений (recipient), который относится к определенному классу (recipient_type). Известно сообщение (message) и время, когда оно было получено агентом (date).

Таблица 1

Фрагмент структуры обмена сообщениями

message	Recipient	recipient_type	Date
back_to_level_01	<population>[7] : 3477	Task_source1	08.02.2023 0:05:37
finish_processing	<population>[7] : 3477	Task_source1	08.02.2023 0:07:30
next_level_02	<population>[7] : 3478	Task_source1	08.02.2023 0:14:29
back to level 02	<population>[7] : 3478	Task_source1	08.02.2023 0:21:26
next level 02	<population>[8] : 3479	Task_source1	08.02.2023 0:21:56
back to level 01	<population>[9] : 3480	Task_source1	08.02.2023 0:24:30
back to level 01	<population>[7] : 3478	Task_source1	08.02.2023 0:26:02
finish_processing	<population>[9] : 3480	Task_source1	08.02.2023 0:26:12

Таблица 2

Фрагмент журнала состояний агентов в моделируемой системе

agent type	Agent	state	entry_date
Task_source1	<population>[10] : 3481	AnalyzeAtLevel_01	08.02.2023 0:25
Task_source1	<population>[10] : 3481	AnalyzeAtLevel_02	08.02.2023 0:29
Task_source1	<population>[10] : 3487	AnalyzeAtLevel_01	08.02.2023 1:07
Task_source1	<population>[10] : 3487	AnalyzeAtLevel_02	08.02.2023 1:10
Task_source1	<population>[10] : 3487	AnalyzeAtLevel_Top	08.02.2023 1:19
Task_source1	<population>[10] : 3487	SolveAtLevel_Top	08.02.2023 1:24

Таблица 3

Фрагмент журнала прохождения агентов по имитационной модели

agent type	Agent	block type	Block
Task_source1	<population>[7] : 3477	Source	sourceTsk1
Task_source1	<population>[7] : 3477	Service	serviceAnalyze_0_01
Task_source1	<population>[7] : 3477	SelectOutput5	selectTsk1
Task_source1	<population>[7] : 3477	Service	serviceSolve_0_01
Task_source1	<population>[7] : 3477	Sink	sinkTsk1

Табл. 2 и 3 содержат фрагмент журнала состояний агентов и фрагмент журнала прохождения агентов по имитационной модели, соответственно.

Фрагмент созданного журнала событий, сформированного из выше представленных журналов с учетом ограничений, которые зафиксированы в онтологии, представлен в табл. 4.

Таблица 4

Пример фрагмента сгенерированного журнала событий

case_id	Activity	timestamp	<i>difficulty_level</i>
<population>[10] : 3487	AnalyzeAtLevel_01	08.02.2023 1:07	0.85
	AnalyzeAtLevel_02	08.02.2023 1:10	0.85
	AnalyzeAtLevel_Top	08.02.2023 1:19	0.85
	SolveAtLevel_Top	08.02.2023 1:24	0.85

Здесь атрибуты case_id, activity, timestamp – обязательные компоненты журнала событий. Кроме того, могут присутствовать дополнительные атрибуты (в таблице выше – атрибут difficulty_level), характеризующие экземпляр в целом или отдельное действие.

Заключение

В статье предложен онтологический подход к генерации событийных данных на основе имитационного моделирования. Новизна полученных результатов заключается в применении онтологического анализа в переходе от контекста имитационных моделей к контексту интеллектуального анализа процессов.

Рассмотрен способ применения данного подхода в моделировании иерархической системы решения задач. Показана архитектура программной системы генерации событийных данных, содержимое онтологического ресурса и генерируемый фрагмент журнала событийных данных, применительно к решаемой задаче.

Сформированная онтология содержит 22 класса (понятия), из них 10 – сгенерированы автоматически на основе журналов имитационной модели. Количество индивидов в онтологии составляет более 100.

Практическая ценность предложенного подхода состоит в сокращении времени формирования журналов событий для последующего интеллектуального анализа за счет исключения необходимости внесения изменений в программный код обработчиков журналов имитационного моделирования.

Список литературы

- [Aalst, 2016] W. van der Aalst. Process Mining – Data Science in Action. Second Edition. – Springer, 2016.
- [Pegoraro, 2021] Pegoraro M., Bakullari B., Uysal M.S., and W.M.P. van der Aalst. Probability estimation of uncertain process trace realizations // in International Workshop on Event Data and Behavioral Analytics (EdbA). – Springer, 2021.
- [Загоруйко, 2004] Загоруйко Н.Г., Налетов А.М., Соколова А.А., Чурикова В.А. Формирование базы лексических функций и других отношений для онтологии предметной области // Труды международной конференции Диалог-2004. – М.: Наука, 2004. – С. 202-204.
- [Гуськов, 2021] Гуськов, Г.Ю., Наместников А.М., Романов А.А., Филиппов А.А. Формирование базы знаний для поддержки процесса архитектурного проектирования программных систем // Онтология проектирования. – 2021. – Т. 11, № 2 (40). – С. 154-169. – DOI: 10.18287/2223-9537-2021-11-2-154-169.
- [Beckett, 2001] David Beckett. The Design and Implementation of the Redland RDF Application Framework. In Proceedings of Semantic Web Workshop of the 10th International World Wide Web Conference, Hong-Kong, China, May 2001.
- [Bertails, 2012] Bertails A., Arenas M., Prud'hommeaux E., Sequeda J., Editors. A Direct Mapping of Relational Data to RDF. – <http://www.w3.org/TR/rdb-direct-mapping/>.
- [Brickley, 2000] Brickley D. and Guha R.V. Resource Description Framework (RDF) Schema Specification 1.0. Candidate recommendation, World Wide Web Consortium, March 2000. – See <http://www.w3.org/TR/2000/CR-rdf-schema-20000327>.
- [Hayes, 2001] Patrick Hayes. RDF Model Theory. Working draft, World Wide Web Consortium, September 2001. – See <http://www.w3.org/TR/rdf-mt/>.

УДК 004.056.5, 004.8

ОРГАНИЗАЦИЯ БЕЗОПАСНОЙ МАРШРУТИЗАЦИИ ДАННЫХ В ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ КОМПЛЕКСАХ С ИСПОЛЬЗОВАНИЕМ ОНТОЛОГИЧЕСКИХ И ГЛУБОКИХ МОДЕЛЕЙ*

А.С. Федулов (*director@sbmpei.ru*)

А.И. Лазарев (*anonymous.project@gmail.com*)

Филиал Национального исследовательского университета «МЭИ»
в г. Смоленске, Смоленск

Индустриальное развитие структур электроэнергетических, электромеханических, теплотехнологических комплексов оказывает непосредственное влияние на развитие интеллектуального распределения данных по протоколу TCP/IP. Существующие системы распределения и учета трафика направлены на достоверную доставку пакетов без потери данных с учетом возможностей интеллектуального выставления приоритетов. Одной из проблем реализаций указанных топологий является наличие множественных конфигураций магистральных и управляемых коммутаторов, включая mesh покрытия. Побочными проблемами данных средств могут являться возможные утечки данных, вызванных неавторизованным доступом к сети как через кроссовер соединение, так и через беспроводные wlan-интерфейсы. Решением указанных проблем является реализованный подход к гибкому управлению сегментом электроэнергетических установок. Основным результатом предлагаемого решения является возможность централизованного анализа изменений с учетом адаптации к сетевым нагрузкам на основе выделенных онтологических переменных. Дополнительной возможностью при реализации данного подхода является совместимость с существующими аппаратными сетевыми устройствами за счет уникальной архитектуры построенной топологии.

Ключевые слова: электроэнергетические комплексы, онтологические модели, классификация данных, принятие решений, нечеткая логика.

* Работа выполнена в рамках государственного задания, проект № FSWF-2023-0012.

Введение

Синергетический эффект, представляемый совокупностью множественных управляемых электросетевых, энергосбытовых компаний, включая энергосистемы и подразделения поддержки коммерческой инфраструктуры Российского рынка, представляет собой комплексную иерархию электроэнергетического комплекса, осуществляющего бесперебойное снабжение потребителей необходимыми средствами [Склюев и др., 2022]. Наряду с развитием данной области, идентичные структуры присущи электромеханическим и теплотехнологическим системам, сопровождаемые активным внедрением IT-технологий в процессы автоматизации деятельности, отслеживания неполадок и проведения критически-важных испытаний.

На текущий момент, развитие подразделений в указанных инфраструктурах поддерживается внедрением сетевых структур на базе протоколов TCP/IP, включая функциональные возможности резервирования важных данных, поддержки отказоустойчивости, оперативного изменения состава сегментов предприятия. Большая часть аппаратных решений, таких как cisco, microtik, так и программных решений pfSense, NethServer, ClearOS поддерживает интеллектуальное управление трафиком клиентов с учетом корректной настройки маршрутизации до конечного клиента [Muthukumar et al., 2019]. Нарушение корректного функционирования таких решений может привести к эксплуатации CVE-уязвимостей, не требующих привилегированного доступа к системе.

С другой стороны, немаловажной проблемой, проиллюстрированной на рис. 1, можно выделить правильность в настройке сетевого оборудования, включая возможное взаимодействие IoT-оборудования с аппаратными техническими комплексами. Развитие данной сферы подтверждается исследованиями многих ученых. Например, авторами публикации [Yi et al., 2016] описывается процесс разработки протокола туннелирования сети с использованием IPv6 адресации и перезаписи полей заголовков для идентификации клиента – данная работа представляет перспективные направления в развитии методов безопасной передачи данных, однако предлагаемое решение обладает новизной в области прогнозирования изменений в сетевом трафике. В тоже время возможности использования обученных моделей не накладывают дополнительных затрат на усовершенствование аппаратной составляющей электроэнергетических комплексов.

Как можно увидеть из рисунка, представленная топология взаимодействия затрагивает два магистральных узла, а также ряд устройств, осуществляющих прямое взаимодействие с управляющими электроэнергетическими установками. Нарушение корректной конфигурации в одном из узлов представленной топологии может повлечь за собой ряд проблем,

таких как перегрузка трафика на одном из узлов сети, нарушение конфиденциальности ввиду использования устаревших версий протоколов и программного обеспечения (включая полнофункциональный доступ к узлу сети), что подтверждает актуальность исследуемой проблемы.

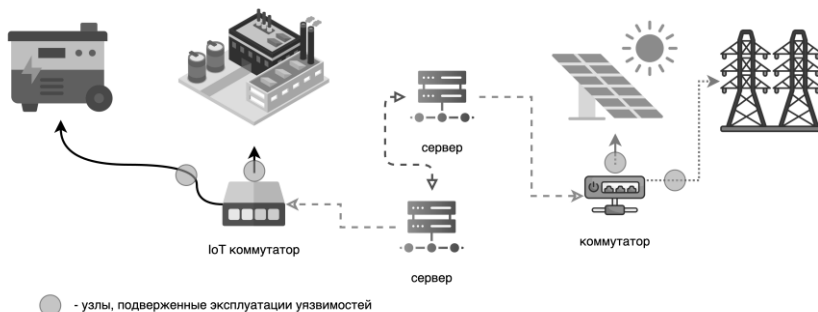


Рис. 1. Топология взаимодействия устройств электроэнергетических структур

В качестве основного решения указанных проблем предлагается реализованная система управления электроэнергетическими ТСП/ІР структурами с использованием данных онтологической модели для автоматизации процессов принятия решений оптимизации работы сегментов сети.

1. Исследование основных параметров эксплуатации уязвимостей на базе онтологических моделей

Рассматривая большую часть уязвимостей и полезных нагрузок на ресурсах offsec, предполагающих прямое или косвенное воздействие на программную платформу информационных структур, следует обратить внимание на таргетированные атаки по протоколам доставки данных, определяемых конкретным портом приложения – на текущий момент насчитывается более 300 портов приложений и служб, для которых существуют уязвимости различных уровней [Offsec, 2023].

Для решения потенциальных проблем с обеспечением безопасности различные компании – поставщики сервисов, на постоянной основе предлагают клиентам возможности своевременного обновления интегрируемого ПО, однако часть специализированных программных средств может напрямую зависеть от аппаратной составляющей, изменение которой в большинстве случаев нерентабельно. В качестве основного подхода к реализации системы оптимального управления данными в сети электроэнергетических структур предлагается использование онтологической

модели, продемонстрированной на рис. 2 – указанная модель представляет совокупный набор большинства полиформатных данных, обрабатываемых на электроэнергетических комплексах.

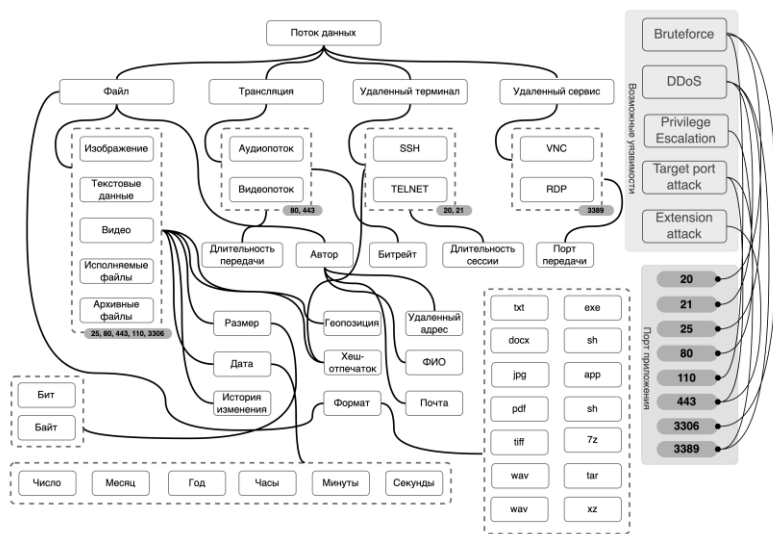


Рис. 2. Онтологическая модель вариативности данных и уязвимостей в электроэнергетическом комплексе

Как можно увидеть из представленной модели, на текущий момент основными сервисами для обработки данных являются приложения, работающие на портах 20, 21, 25, 80, 443, 3306, 3389, 110 [Бобряков и др., 2022]. Указанные порты в большинстве случаев являются стандартными для доступа к интерфейсу управления, что также нарушает политику информационной безопасности.

Условное подразделение, представленное на рис. 2, указывает на множественные точки уязвимостей, например – внедрение вредоносного исполняемого кода в передаваемый файл, изменение и перехват аудио / видеопотока, использование слабых ключей шифрования к удаленному терминалу, а также обеспечение удаленного управления через устаревшие версии протоколов RDP / VNC. Использование представленной онтологической модели в процессе разработки программного алгоритма также позволяет выявить основные изменяемые треггами лицами показатели для маскировки изменений в оригинальной сигнатуре, такие как хеш-отпечаток файла, информация о авторе, геолокация. В качестве основных

примеров реализации уязвимостей можно привести CVE-2019-0708, CVE-2019-1935, CVE-2020-0688, некоторые из которых не требуют дополнительного вмешательства со стороны атакуемого лица. Программную реализацию представленной онтологической модели возможно провести с использованием средств захвата трафика на основе двоичных файлов pcap (Packet Capture Data), таких как tcpdump, Wireshark.

С другой стороны, рассматривая структуру обмена данными между пользователями, возможно выделить основные факторы отличия передаваемых данных, как представлено в рис. 3 – выделяемые классы трафика позволяют в последующем оптимизировать работу отдельных подсетей для обеспечения минимальной задержки с внешним сервером.

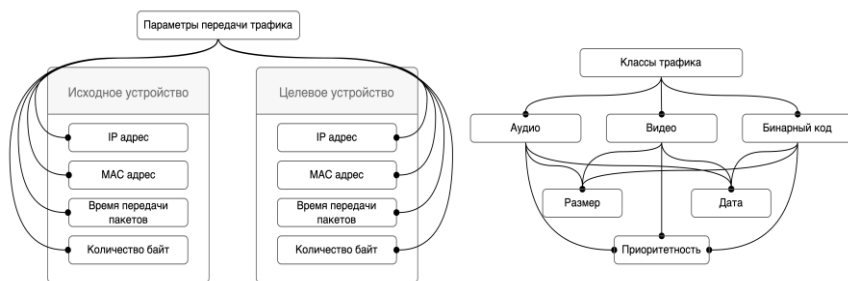


Рис. 3. Онтологические модели набора параметров в пакетах трафика / упрощенная модель разделения трафика

Результат рассмотрения представленных моделей, в частности, упрощенной модели классификации трафика, позволяет в последующей реализации подхода к оптимизации сети электроэнергетических образований использовать сетевые параметры исходных и целевых устройств как основные прогнозируемые показатели в системе принятия решений по управлению сегментированными участками.

2. Разработка многомодульной системы управления электроэнергетическими структурами

Немаловажной составляющей при разработке интеллектуального подхода к управлению сетевыми структурами в электроэнергетических комплексах является схема определения ключевых устройств в сегменте сети, а также выявления возможных мест эксплуатации уязвимостей в соответствии с представленной онтологической моделью классификации данных. В предлагаемой схеме интеллектуального управления, изображенном на рис. 4, основным является процесс первичного анализа доступных уст-

роЙств в сегменте сети с последующим выявлением доступных портов (глубокое сканирование) для последующей классификации трафика и обучения глубокой модели.



Рис.4. Схема адаптации топологии электроэнергетических комплексов

Как можно увидеть из представленной схемы, процесс оптимизации нацелен на внесение изменений в реальном времени с учетом выявленных отклонений в соответствии с первоначальной структурой TCP/IP – этапы работы в режиме изменений обособлены наличием изменчивости в структуре электроэнергетических систем, то есть добавлением / удалением различных сегментов.

В качестве основной программной составляющей предполагается наличие базового DHCP сервера разделения сегментации (*dhcpcd*), виртуальных компьютерных сетей (VLAN), а также нескольких управляемых точек виртуальных внешней сетей (VPN), инициализируемых сервером WireGuard / OVPN [Yi et al., 2023]. Выделяя возможные подразделения сегментов сети, также стоит отметить, что соединения между выделяемыми подразделениями осуществляются через виртуальный сегмент с использованием частных виртуальных сетей.

Организация процесса классификации трафика в данном решении основана на использовании дополнительных средств анализа выходных данных фреймворка NFStream [Aouini et al., 2021]. Используемый пакет задействует глубокий анализ пакетов за счет выделения множества параметров начальной и конечной точек IP адресации, задержки отправки и получения данных, количества выходной и входной информации (в байтах), а также размеры пакетов и нагрузки сети. В данной модели прогнозирования используется процесс захвата потока из адаптера *en0* с последующим сохранением в csv файл – в качестве основных параметров для

прогнозирования используются *bidirectional_packets*, *bidirectional_bytes* составляющие аккумулятор потоков и количества байтов для двунаправленных пакетов.

В качестве дополнительных средств для прогнозирования изменений предлагается использование глубокой модели на основе выходных данных rсар файла, как входных данных для обучения сети [Пучков и др., 2022]. Основу предлагаемой нейронной сети составляет Long-Short Term Memory (LSTM) модель, представляемая возможностью долговременного хранения зависимостей [Мешалкин и др., 2021].

Программная реализация LSTM модели основана на библиотеке Keras, включающая слои *Dense*, *LSTM*, *SimpleRNN* – основным количеством нейронов для задаваемых слоев являются значения 1, 4, 3 соответственно [Дли и др., 2022]. Основной функцией оптимизатора является *adaptive moment estimation* (*adam*) с целевым параметром потерь – расчет среднеквадратичной ошибки [Дли и др., 2020].

В результате, при использовании комбинации фреймворка NFStream и стороной LSTM модели, возможно получать данные как по классификации трафика, так и прогнозируемые значения по расходуемому трафику в различных периодах за счет представления.

3. Разработка модуля обеспечения безопасного функционирования электроэнергетической ТСП/Р топологии

Немаловажной составляющей в процессе ТСП/Р взаимодействия являются процессы шифрования данных с использованием как протоколов защиты трафика, так и средств шифрования данных на клиенте. Существующие средства туннелирования трафика в большинстве случаев основаны на первичной установке клиент-серверных корневых сертификатов, обмена двухсторонними ключами (подключение как пира), а также с применением пары логин-пароль. Альтернативная поддержка безопасности также возможна сертификацией с использованием SSL шифрования, однако актуальным остается вопрос проведения таргетированных атак на отдельные службы и сервисы – как можно увидеть из представленной онтологии рис. 2, непреднамеренное использование хотя бы одного из уязвимых сервисов может привести к утечке данных, следуемых за несанкционированным доступом.

В качестве решения указанной проблемы предлагается использование средств хэширования на основе алгоритма BLAKE3 с последующим комбинированием ТОТР алгоритма обновления ключей шифрования, реализующих совокупный алгоритм непрерывной аутентификации субъектов. Алгоритм шифрования BLAKE3 является хорошей альтернативой существующим хэш-функциям, представляя возможности быстрого получения

зашифрованного сообщения наряду с надежностью, идентичной SHA-3. В качестве основного тестирования производительности была проведена проверка вычисления хеш-суммы двоичного файла на процессоре MAC M1 (рис. 5). Согласно результатам последовательного тестирования указанных функций можно сказать, что для расчета хеш-функции алгоритму SHA-224 требуется 3,708 сек., в то время как для алгоритма BLAKE3 – 0,18 сек.

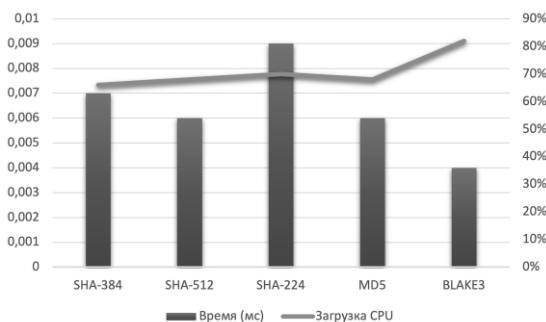


Рис. 5. Сравнение алгоритмов получения хеш-функций на процессоре M1

С учетом того, что при выполнении TOTP шифрования важным фактором является процесс обновления функции за определенный фиксированный интервал, применение алгоритма шифрования BLAKE3 является наиболее целесообразным. Алгоритмическая составляющая BLAKE3 образована за счет комбинирования средств предыдущего хеш-шифра (BLAKE2) и Bao, что в совокупности снимает ограниченное ветвление, как продемонстрировано на рис. 6. Архитектура BLAKE2 обособлена наличием дополнительного блока *genBlock*, который осуществляет склейку входного массива данных и ключа шифрования блоком фиксированного размера с последующей передачей количества блоков в следующий блок. *blake2bDigest* в свою очередь осуществляет получение хэш-значений для входных блоков.

Из приведенной иллюстрации можно сказать, что архитектура BLAKE3 ограничивается блоками (c_i) по 1024 байта, при которых возможно переполнение с последующим образованием 2х родительских узлов (d_i), причем переполнение возможно начиная с предоставления 1 байта для следующего фрагмента. В качестве зависимой переменной для TOTP-шифрования используется входная переменная (id), рассматриваемая как идентификатор аппаратной составляющей для получения временной аутентифицируемой последовательности (*TOTP GEN*).

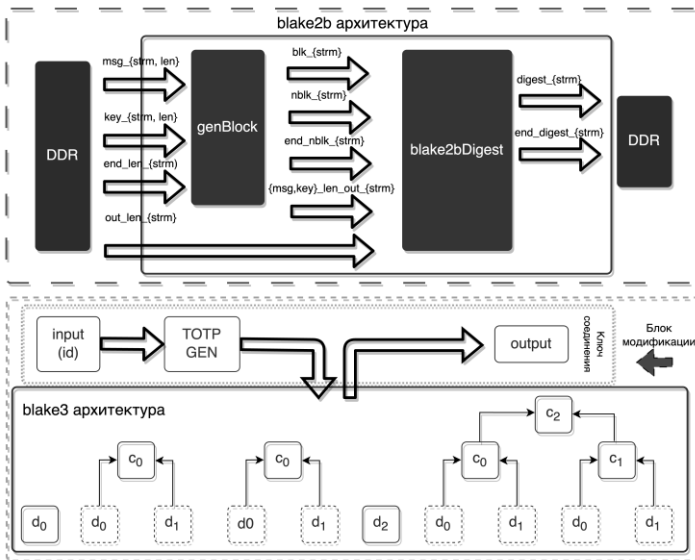


Рис. 6. Архитектуры BLAKE2B и BLAKE3 с динамическим шифрованием

Для организации процесса динамического обновления ключа предлагается использовать TOTP генерацию зависимого ключа, образуемого по формуле 3.1.

$$BLAKE3(VALUE) = CLIENT_{id},$$

$$VALUE(ID) = \left[\frac{T_1 - T_0}{T} \right] \cdot DELAY, \quad (3.1)$$

где $VALUE$ – цифровой временный отпечаток; $BLAKE3$ – общий идентификатор для пары устройств, образуемый от $VALUE$ и хэш-функции; T_1 – текущее системное время; T_0 – статичный параметр времени; T – дискретное значение времени; $DELAY$ – время действия отпечатка; $CLIENT_{id}$ – идентификатор пары.

Программная реализация указанного метода осуществляется с учетом передаваемого потока информации – предлагаемое решение за счет своей гибкости позволяет гибко вычислять цифровые отпечатки для определенных данных, действительный в течении n -го количества времени.

4. Разработка модуля принятия решений по изменению структуры распределения трафика ТСП/Р

На основании того, что большая часть прогнозных данных в предлагаемом подходе представлена числами с плавающей точкой, предлагается использовать нечеткое сопоставление данных, которое нацелено на идентификацию похожих, но не идентичных данных. На текущий момент, для сопоставления данных может использоваться множество алгоритмов, включая расстояние Хэмминга, Дамерау-Левенштейна, Левенштейна. Реализация последнего из приведенных алгоритмов сводится к измерению расстояния между двумя целевыми последовательностями, цель которого – вычисление количества односимвольных правок согласно формуле 4.1.

$$\text{lev}(a,b) = \begin{cases} |a| & \text{if } |b| = 0, \\ |b| & \text{if } |a| = 0, \\ \text{lev}(\text{tail}(a), \text{tail}(b)) & \text{if } a[0] = b[0], \\ 1 + \min \begin{cases} \text{lev}(\text{tail}(a), b) \\ \text{lev}(a, \text{tail}(b)) \\ \text{lev}(\text{tail}(a), \text{tail}(b)) \end{cases} & \text{otherwise} \end{cases} \quad (4.1)$$

где $\text{tail}(x)$ – символьная строка, исключаяющая первый символ последовательности x ; $x[n]$ – n -ый символ строки x , начиная с нулевой позиции.

С учетом указанной формулы, возможно вычисление процентного совпадения по прогнозируемым изменениям – например, процент совпадения по «1.023» и «1.029» Mb/s (расход трафика) позволяет выполнить перераспределение приоритетов классификации для снижения нагрузки на сеть. Программная реализация указанного метода произведена с использованием библиотеки *Fuzzy* на языке *rpython*, где реализованы функции как частичного, так и посимвольного сравнения [Дли и др., 2021].

Заключение

Рассмотрение инфраструктуры электросетевых компаний как единой экосистемы взаимодействия устройств в рамках ТСП/Р адресации позволило выявить существующие проблемы в электроэнергетических компаниях, такие как недостаточная эффективность при наличии множественных удаленных управляющих узлов, использование устаревших средств прикладного программного обеспечения и протоколов передачи данных. Для решения указанных проблем в статье предложены онтологические модели классификации трафика, которые позволили разработать функ-

циональные модули классификации и прогнозирования трафика как между локальными, так и центральными узлами сети. В качестве основного решения от несанкционированного доступа и предотвращения возможных утечек данных в статье предложен алгоритм комбинированного использования средств шифрования на основе алгоритма BLAKE3 и TOTP авторизации. Предложенный алгоритм динамического шифрования на основе своей уникальности не уступает существующим алгоритмам шифрования по скорости вычисления хэш-функции, а также обладает поддержкой уникальности при обработке полиформатного трафика.

Существенным программным модулем в предлагаемом решении является алгоритм автоматизированного принятия решений по управлению изменениями в структуре электроэнергетических систем на основе расстояния Левенштейна – представленный результат посимвольного сравнения позволил вычислять отклонения между исходными и прогнозируемыми значениями для изменения топологий трафика. Результаты последующих исследований в данной сфере также помогут внести особый вклад в развитие теории взаимодействий TCP/IP маршрутизации и искусственного интеллекта.

Список литературы

- [Бобряков и др., 2022] Бобряков А.В., Борисов В.В., Мисник А.Е., Прокопенко С.А. Моделирование и проектирование информационно-аналитических производственных процессов на основе нейронечетких темпоральных сетей Петри // Прикладная информатика. – 2022. – Т. 17, № 2. – С. 65-78. – DOI: 10.37791/2687-0649-2022-17-2-65-78.
- [Дли и др., 2020] Дли М.И., Булыгина О.В., Соколов А.М. Рубрицирование текстовой информации на основе голосования интеллектуальных классификаторов // Прикладная информатика. – 2020. – Т. 15, № 5. – С. 29-36. – DOI: 10.37791/2687-0649-2020-15-5-29-36.
- [Дли и др., 2021] Дли М.И., Власова Е.А., Соколов А.М., Моргунова Э.В. Создание цифрового двойника химико-технологической системы с использованием языка Python // Прикладная информатика. – 2021. – Т. 16, № 1. – С. 22-31. – DOI: 10.37791/2687-0649-2021-16-1-22-31.
- [Дли и др., 2022] Дли М.И., Синявский Ю.В., Рысина Е.И., Василькова М.А. Метод классификации перемещающихся устройств с использованием глубоких нейронных сетей с расширенным рецептивным полем // Прикладная информатика. – 2022. – Т. 17, № 5. – С. 51-61. – DOI: 10.37791/2687-0649-2022-17-5-51-61.
- [Мешалкин и др., 2021] Мешалкин В.П., Дли М.И., Пучков А.Ю., Лобанева Е.И. Предварительная оценка прагматической ценности информации в задаче классификации на основе глубоких нейронных сетей // Прикладная информатика. – 2021. – Т. 16, № 3. – С. 9-20. – DOI: 10.37791/2687-0649-2021-16-3-9-20.

- [Пучков и др., 2022] Пучков А.Ю., Дли М.И., Прокимнов Н.Н., Шутова Д.Ю. Многоуровневые алгоритмы оценки и принятия решений по оптимальному управлению комплексной системой переработки мелкодисперсного рудного сырья // Прикладная информатика. – 2022. – Т. 17, № 6. – С. 102-121. – DOI: 10.37791/2687-0649-2022-17-6-102-121.
- [Скляев и др., 2022] Скляев А.М., Хабаров В.И., Мусатова И.В., Попова О.В. Организационно управленческие инновации в электронной промышленности России: современные тренды // Современная конкуренция. – 2022. – Т. 16, № 6. – DOI: 10.37791/2687-0657-2022-16-6-103-116.
- [Aouini et al., 2021] Aouini Z., Pekar A. NFStream: A flexible network data analysis framework // Computer Networks. – 2021. – Vol. 204. – DOI: 10.1016/j.comnet.2021.108719.
- [Muthukumar et al., 2019] Muthukumar M., Senthilkumar P., Jawahar M. Firewall Scheduling and Routing Using pfSense // Intelligent Computing and Applications. – 2019. – Vol. 1172. – DOI: 10.1007/978-981-15-5566-4_67.
- [Offsec, 2023] Exploit Database - Exploits for Penetration Testers, Researchers, and Ethical Hackers [Электронный ресурс]. Exploit Database. – URL: <https://www.exploit-db.com> (дата обращения: 10.05.2023).
- [Yi et al., 2016] Yi B., Congxiao B., and Xing L. FlowLAN: A non-tunneling distributed virtual network based on IPv6 // Proc. IEEE ITNEC. – 2016. – P. 229-234. – DOI: 10.1109/ITNEC.2016.7560355.
- [Yi et al., 2023] Yi T., Chen X., Zhu Y., Ge W., Han G. Review on the application of deep learning in network attack detection // Journal of Network and Computer Applications. – 2023. – Vol. 212. – DOI: 10.1016/j.jnca.2022.103580.

Секция 2 | ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ

УДК 004.89

О ФОРМИРОВАНИИ ЭМПИРИЧЕСКИХ ТЕОРИЙ В ПЛОХО ФОРМАЛИЗОВАННЫХ ОБЛАСТЯХ

С.М. Гусакова (*svem45@yandex.ru*)

Федеральный исследовательский центр
«Информатика и управление» РАН, Москва

М.А. Михеенкова (*m.mikheyenkova@yandex.ru*)

Федеральный исследовательский центр
«Информатика и управление» РАН, Москва

Российский государственный гуманитарный университет, Москва

В работе представлены проблемы формирования нового знания средствами искусственного интеллекта в открытых эмпирических областях, где в значительной степени отсутствует формальный аппарат и процедуры формирования теорий носят эвристический характер. Рассматриваются вопросы представления данных и знаний, формирования баз фактов, выбора процедурных решений. Охарактеризованы подходы к заданию алгебраической операции сходства – ключевого элемента правдоподобных рассуждений средствами ДСМ-метода автоматизированной поддержки исследований – для различных структур. Подчёркивается, что эффективные решения указанных проблем могут быть предложены лишь в условиях междисциплинарного взаимодействия.

Ключевые слова: представление данных и знаний, база фактов, ДСМ-метод, структура данных, операция сходства.

Введение

Порождение нового знания и формирование теорий в плохо формализованных областях – от наук о жизни до наук о человеке и обществе – как правило, носит эвристический характер. Живучее представление о принципиальном несовпадении процесса познания в «науках о природе» и «науках о культуре» [Риккерт, 1998] неизбежно сталкивается в последних с проблемой обоснованности порождаемых конструктов. И если для умозрительных построений разнообразие мнений может рассматриваться как вполне допустимое, для эмпирических исследований указанная проблема является критичной. Отсюда – растущее внимание к возможностям формализации познавательного процесса в таких областях, в последнее время инициировавшееся повсеместным распространением методов искусственного интеллекта (ИИ), в особенности – методов нейронных сетей и глубокого обучения.

Как известно, ключевыми проблемами таких подходов являются проблемы формирования эффективной выборки и недостаточной интерпретируемости результатов (см., например, [Gao et al., 2023]), связанные с доминирующим вниманием к вычислительным аспектам в противовес к собственно проблемам представления декларативных знаний о предметной области [Chao et al., 2023]. И это при том, что фундаментальное значение представления знаний и формализации рассуждений для искусственного интеллекта считается самоочевидным для специалистов в области ИИ и имеет развитые традиции формирования различных подходов [Brachman et al., 2004]. Несмотря на это, по-прежнему открытыми остаются многие вопросы оценки полноты и содержательности эмпирической базы фактов и языка представления этих фактов и знаний предметной области, а также процедурных решений для их анализа и построения на его основании эмпирических теорий [Забежайло, 2022].

Рефлексия по этому поводу содержательным образом отражена в понятии квазиаксиоматической (открытой) теории КАТ [Финн, 2023, с. 310–380] $T = \langle \Sigma, \Sigma', R \rangle$, являющимся базовой формальной конструкцией для формирования эмпирических теорий. Здесь Σ – априорное знание, которое может быть представлено множеством принципов (аксиом), заведомо неполно характеризующих предметную область, Σ' – открытое множество элементарных высказываний о фактах и гипотезах этой области. Множество правил вывода R включает как правила достоверного вывода R_d , так и правила правдоподобного вывода R_p , $R = R_d \cup R_p$. Конечная цель эмпирического исследования состоит в пополнении знаний Σ – формировании теорий – из анализа имеющихся данных Σ' , средствами рассуждений с использованием правил вывода R . Трудности этого процесса в плохо

формализованных областях сопряжены с неотъемлемыми их особенностями: неполнотой (открытостью) эмпирического материала (базы фактов, БФ), неполнотой имеющихся знаний (базы знаний, БЗ) и – как следствие – трудностями формирования языка представления этих знаний и самих данных и, наконец, открытостью множества процедур порождения нового знания – правдоподобных рассуждений.

1. Проблемы подготовки эмпирического материала

Описанный подход конструктивно реализован в ДСМ-методе автоматизированной поддержки исследований (ДСМ-метод АПИ) [Автоматическое порождение..., 2020], являющемся методологией создания интеллектуальных систем и средством формализации, имитации и усиления интеллектуальных процессов. Метод воспроизводит исследовательские эвристики плохо формализованных областей в виде синтеза неэлементарных познавательных процедур: эмпирической индукции (анализ), структурной аналогии (предсказание), абдуктивного объяснительного принятия гипотез. Взаимодействие различных таких процедур в процессе анализа данных из пополняемых БФ порождает новое знание, на основании которого формируется семейство открытых (квазиаксиоматических) эмпирических теорий.

Многолетний опыт применения ДСМ-метода в различных науках о жизни и социальном поведении (см. [Автоматическое порождение..., 2020] и др.) подтверждает необходимость предварительной содержательной работы эксперта предметной области для подготовки эмпирического материала с последующим взаимодействием с ИИ-специалистом для тщательного выбора языка представления данных и знаний и адекватных процедур их анализа, в частности – R_p . Во многих случаях охарактеризованные выше проблемы формирования эмпирических теорий усугубляются невозможностью управления экспериментом (медицина, социология) и, соответственно, коррекции БФ, т.е. множества Σ' . Одновременно изменение БФ (к примеру, её расширение) может приводить как к появлению новых зависимостей [Климова и др., 2020], так и исчезновению старых [Gross et al., 2020]. Для нивелирования такого негативного эффекта в ДСМ-методе развит аппарат порождения эмпирических закономерностей (законов и тенденций) [Финн, 2023, с. 310–380], представляющих собой инварианты выявленных зависимостей и доопределяемых ими фрагментов БФ, свойства которых заранее неизвестны.

Другим объективным источником проблем при формировании эмпирических теорий служит отсутствие формального языка представления данных. Очевидно, что этот язык зависит от уже имеющегося знания, представленного аксиомами предметной области (элементами Σ), и традиционной для области формы сбора эмпирического материала – резуль-

татов опросов в социальных науках, психологических тестов, клинических данных в сочетании с субъективными жалобами и наблюдениями на приеме в медицине и т.п. Необходимым условием успешного исследования и построения каузальной эмпирической теории в таких случаях является проработанность теоретической (содержательной) модели и включение значимых факторов влияния в описание БФ – в противоположность статистической модели, ориентирующейся на широкую представленность множества параметров. Лишь после этого может начинаться работа по разработке адекватного языка представления данных и включению во множество Σ соответствующих аксиом структуры данных.

Однако даже при соблюдении всех указанных условий исследование может показать необходимость уточнения языка и/или развитие инструментов анализа. Так, в [Wang et al., 2022] при решении задачи прогнозирования свойств молекул и открытия лекарств исходное традиционное топологическое представление молекул в виде графа дополняется представлением об эквивалентности молекул по отношению к химическим реакциям. Такой подход существенно улучшает прогнозирование продукта реакции и предсказание свойств молекулы. В другой задаче – прогнозировании энерго-эффективности химических соединений – расширение графово-молекулярного представления включением релевантных проблемно-ориентированных квантово-химических параметров существенно меняет результаты [Митин и др., 2020].

Аналогичный опыт имеется и при решении задачи формирования социальной структуры средствами ДСМ-метода [Гусакова и др., 2016]. Для типологизации социума и формирования социальных групп использовался обратный ДСМ-метод: анализ «от следствия – к причине». Социальная группа представлена индивидуумами, имеющими общие характеристики V и выражающими сходные мнения W . В исходных данных БФ⁺ в ДСМ-методе представлена (+)-фактами наличия исследуемых свойств (в данном случае – мнений) $J_{(1,0)}(X \Rightarrow_1 Y)$ – «объект X обладает множеством свойств Y » и (-)-фактами их отсутствия $J_{(-1,0)}(X \Rightarrow_1 Y)$ – «объект X не обладает множеством свойств Y ».

Предикат простого обратного положительного сходства $\tilde{M}_{a,n}^+(V,W,k)$, использующийся при формулировке индуктивных правил правдоподобного вывода, распознает локальное сходство¹ $(\cap_{i=1}^k C_i = V) \& (V \neq \emptyset) \& (\cap_{i=1}^k Q_i = W) \& (W \neq \emptyset)$ на множестве (+)-примеров $J_{(1,n)}(C_i \Rightarrow_1 Q_i)$, $i = 1, \dots, k$, ($k \geq 2$), которое является основанием для правдоподобного вывода о причинах рассматриваемого явления. Предикат описывает эмпирическую зависимость $\forall X \forall Y (J_{(1,n)}(X \Rightarrow_1 Y) \& \quad \forall U (J_{(1,n)}(X \Rightarrow_1 U) \rightarrow U \subseteq Y) \& W \subseteq Y) \rightarrow (V \subseteq X \& V \neq \emptyset)$,

¹ См. раздел 2.

содержательно интерпретируемую как «сходство мнений W субъектов X в (+)-примерах есть следствие сходства V самих субъектов X ». Соответственно, объединённые общим мнением W и общими характеристиками V субъекты образуют некоторую социальную группу. Таким образом, личностные свойства, формирующие класс групповых признаков (общие характеристики V индивидуумов), становятся идеально-типическими свойствами. Парадокс состоит в том, что процедура допускает одновременное выполнение предикатов $\check{M}_{a,n}^+(V,W)$ и $\check{M}_{a,n}^+(V,W)$ таких, что $W \cap W' = \emptyset$, где непараметрический предикат $\check{M}_{a,n}^+(V,W) \Leftrightarrow \exists k \check{M}_{a,n}^+(V,W,k)$. На уровне интерпретации это означает, что один и тот же класс становится основой различных типологизаций, связанных с различными комплексными мнениями. Такая ситуация может быть разрешена, в частности, изменением/расширением языка представления данных – например, вследствие пересмотра исходной модели предметной области с соответствующим выбором нового набора факторов влияния – параметров представления.

Одним из вариантов расширения языка представления данных в ДСМ-методе является учёт контекста – внутреннего, распознаваемого тернарным каузальным отношением [Финн и др., 1993; Финн и др., 2016], или внешнего [Автоматическое порождение..., 2020, с. 428–445]. Такое изменение автоматически влечёт уточнение и развитие исследовательских процедур.

2. Выбор процедурных решений

Если в вопросах формирования БФ и выборе языка представления доминирующую роль играют первичные содержательные соображения экспертов предметной области, процедурные решения являются прерогативой специалистов в области ИИ при одновременном выполнении требования интерпретируемости результатов, а при возможности – и самого процесса исследований. Это означает создание и поддержку широкого спектра конструктивных эвристик [McCarthy, 2014]. В ДСМ-методе предусмотрена возможность выбора эвристических подходов – процедур рассуждения, которые соответствуют различным типам каузального вынуждения, представленным в БФ, и опираются на результаты метатеоретического исследования предметных областей [Финн, 2023, с. 381–436]. Выявляемые в этом случае причинные зависимости представляют собой обособленную, конструктивно порождаемую и верифицированную обусловленность исследуемого эффекта.

Источником детерминаций – каузальных эмпирических зависимостей – в ДСМ-методе является поиск структурного (неметрического) схождения объектов. Операция схождения определяется не просто как алгебраическая операция, удовлетворяющая свойствам идемпотентности, коммутативности, ассоциативности и наличию нуля, но и с учетом аксиом предметной

области и структуры данных, выраженной в языке представления данных. Таким образом, определение операции сходства и связанного с ней отношения вложения является неотъемлемой составляющей процедурных решений, которые на одной и той же БФ могут приводить к результатам различного прагматического уровня.

ДСМ-метод АПИ позволяет рассматривать объекты различной структуры, определяемой языком представления данных. С математической точки зрения это могут быть множества, кортежи, графы, алгебраические системы. Свойства могут быть, также как объекты, представлены математическими структурами, а могут быть просто помечены наличием или отсутствием у данного объекта и тогда могут быть описаны в каждом положительном или отрицательном примере как множество имен. Для структурного представления объектов выделяется множество признаков $At = \{A_1, \dots, A_n\}$ и множества значений каждого признака $V(A_i)$ ($i = 1, \dots, n$), с помощью которых можно описать объекты. Каждое из этих множеств содержит пустой элемент. Соответственно, сходство определяется для каждого заданного представления данных.

Для объектов, представленных в виде подмножеств множества At без уточнения значений признака, операция сходства может быть задана как операция пересечения этих подмножеств. Это имеет место в таких задачах, как «структура – активность», где химическое соединение описано в виде подмножества дескрипторов языка ФКСП [Блинова и др., 2016].

Объект может быть описан как подмножество имен признаков из множества At , где каждому признаку приписано одно или несколько значений, проявившихся в данном объекте. Тогда объекты могут быть представлены как подмножества объединения декартовых произведений $\bigcup_{i=1}^n A_i \times V(A_i)$. Такое описание объектов встречается, например, в задачах почерковедческой направленности, когда $\{A_1, \dots, A_n\}$ – это все буквы алфавита и быть может безбуквенные элементы, а $V(A_i)$ – все значения частных признаков почерка для данной буквы [Гусакова и др., 2010].

Если объекты определены в языке представления данных как кортежи, возможны два варианта. В первом варианте – множество значений каждого признака одно и то же, объект можно представить как элемент декартовой степени множеств из At . Каждая компонента такого кортежа может состоять из одного значения признака, но может содержать и несколько значений, как в примере с частными признаками почерка. Очевидно, что такое описание эквивалентно представлению объекта в виде подмножества множества $\bigcup_{i=1}^n A_i \times V(A_i)$.

Для всех вышеприведенных представлений объектов в языке представления данных операция сходства может быть определена как операция пересечения множеств. Два объекта не сходны, если результат операции пересечения – пустое множество.

Но возможно, что множества значений элементов кортежа имеют разную природу. В этом случае имеется не одно множество значений признаков, а n – по количеству элементов кортежа.

Операция сходства для таких кортежей задается следующим образом (Z_1, Z_2 – объекты):

$$Z_1 \Pi Z_2 = \{z^1_1 \Pi_1 z^2_1, \dots, z^1_n \Pi_n z^2_n\},$$

где для каждого i операция сходства Π_i зависит от типа данных множества $V(A_i)$. Результатом такой операции сходства будет кортеж, компоненты которого есть результат операции сходства для каждого типа признаков, входящих в кортеж. Этот частный результат может быть не пустым, если значения признака совпадают или относятся к одному и тому же классу (в случае разбиения множества значений признака на классы). Если значения признака заданы числовыми интервалами, границы результирующего интервала для $[a^1, b^1]$ и $[a^2, b^2]$ можно определить как $[\max(a^1, a^2), \min(b^1, b^2)]$. Два объекта Z_1 и Z_2 , имеющие такое описание, не сходны, если в результирующем кортеже $Z_1 \Pi Z_2$ количество компонент, равных пустому элементу больше либо равно выбранному порогу. Этот порог может принимать значения от 1 до n .

Когда объекты описаны графами, операция сходства может быть задана как операция пересечения множеств упорядоченных или неупорядоченных пар. Если в задаче, где объекты представлены в виде обыкновенных графов, сходство должно отражать не просто наличие изоморфных подграфов, но важна также структурная близость, можно определить еще две операции сходства Π_1 и Π_2 на графах.

Для определения этих операций введем соответствие Галуа f, g между булеанами $B(V)$ и $B(C)$, где V – множество вершин графа, C – множество его клик:

$$\begin{aligned} - \forall \tilde{V} \in B(V) f(\tilde{V}) &= \{\hat{C} \in B(C), | \tilde{V} \subseteq c, c \in \hat{C}\} \\ - \forall \hat{C} \in B(C) g(\hat{C}) &= \{\cap_{c \in \hat{C}} c\}. \end{aligned}$$

Тогда подмножество вершин \tilde{V} , для которого $g f(\tilde{V}) = \tilde{V}$, замкнуто и подмножество клик \hat{C} , для которого $f g(\hat{C}) = \hat{C}$, замкнуто.

Ядро – это такое подмножество вершин V' , для которого $f(v_i) = f(v_j)$, $v_i, v_j \in V'$. Очевидно, что ядро можно определить через элементы C следующим образом: $\text{Ker}(\hat{C}) = \cap_{c \in \hat{C}} c \cap_{\bar{c} \in \bar{\hat{C}}} \bar{c}$. [Очан, 1942].

В соответствии с операцией Π_1 сходны графы, имеющие изоморфные графы клик. Поскольку для графов, имеющих изоморфные графы клик и одновременно являющихся или не являющихся клик-Хэлли графами (клик-Хэлли граф – это граф, у которого любое подмножество попарно пересекающихся клик имеет общее пересечение), существует биекция

между множествами клик такая, что замкнутое ядро переходит в замкнутое ядро, результатом операции сходства Π_1 графов G_1 и G_2 является индуцированный подграф, построенный на замкнутых ядрах в каждом графе. Очевидно, что такие подграфы изоморфны.

Операция Π_2 основана на том факте, что каждому обыкновенному графу однозначно соответствует факторизованный граф, т.е. граф, у которого все непустые ядра имеют мощность равную единице. Такой граф строится как фактор-граф по эквивалентности, задаваемой ядрами как классами на множестве вершин графа. Соответствие между графом и его фактор-графом является гомоморфизмом и корреспонденцией. В соответствии с операцией Π_2 два графа сходны, если их фактор-графы изоморфны и результат операции Π_2 – этот факторизованный граф, который является подграфом графов-операндов.

Для установления факта изоморфизма факторизованных графов двух графов можно воспользоваться следующим утверждением.

Утверждение.

Графы G_1 и G_2 имеют изоморфные факторизованные графы тогда и только тогда, когда существует биекция μ между множествами их клик такая, что $|Ker(\hat{C})| \neq \emptyset \Leftrightarrow Ker(\mu\hat{C}) \neq \emptyset$.

Доказательство.

Если графы G_1 и G_2 имеют изоморфные фактор-графы, то по теореме 1 из [Очан, 1942] между множествами клик этих графов существует биекция, переводящая непустое ядро (т.е. мощность которого равна единице) в непустое ядро. Поскольку графы G_1 и G_2 отличаются только количеством вершин в ядрах, эта биекция распространяется и на множества клик графов G_1 и G_2 и соответствующие по ней ядра одновременно пусты или не пусты.

Если же существует биекция μ между множествами клик графов G_1 и G_2 , такая, что $|Ker(\hat{C})| \neq \emptyset \Leftrightarrow Ker(\mu\hat{C}) \neq \emptyset$, то она распространяется и на их фактор-графы и мощности соответствующих ядер одновременно равны единице или нулю. По указанной теореме Очана фактор-графы графов G_1 и G_2 изоморфны.

Легко видеть, операция Π_2 задает более сильное структурное сходство, чем операция Π_1 .

В случае представления объектов алгебраическими системами, операция сходства – это пересечение алгебраических подсистем алгебраической системы, построенной на носителе $V(At)$. Поскольку подмножества носителя, являющиеся носителями подсистем алгебраической системы, замкнуты относительно операций сигнатуры системы, и пересечение таких замкнутых подмножеств замкнуто, пересечение подсистем алгебраической системы тоже подсистема. Заметим, что рассматриваются только

такие алгебраические системы, которые не имеют в сигнатуре 0-арных операций. Это является условием необходимым для того, чтобы пустое пересечение подсистем алгебраической системы тоже было подсистемой.

При определении операции сходства необходимо учитывать, что правильное с математической точки зрения формальное определение этой операции еще не есть гарантия успеха при решении задач, особенно, когда дело касается плохо формализованных предметных областей. Здесь, как уже говорилось, необходимо также учитывать не только аксиомы предметной области, но в значительной мере и экспертные знания, как правило, неформализуемые. Характерной чертой многих таких областей – психологии, социологии, криминалистики, даже отчасти медицины и др. – является субъективизм, присутствующий при сборе и описании данных. Поэтому, несмотря на разнообразие математических структур, используемых для описания объектов и возможности корректного определения операции сходства на каждой из таких структур, в некоторых задачах приходится отступать от шаблонов и привлекать эвристические методы, что естественно в условиях неполноты данных и знаний. Такие методы использовались при решении задачи датировки берестяных грамот, где операция сходства на объектах, представленных подмножествами признаков, определялась как объединение, а не пересечение множеств [Гусакова, 2004] или в задаче определения зависимости трудно подделываемых признаков подписей от темперамента, где операция сходства была определена по результатам пред-процессинга с привлечением характеристик, не входящих в объекты, на которых определялось сходство [Гусакова и др., 2022].

Заключение

Многолетняя история развития ИИ подтверждает, что это направление исследований по своим целям и задачам является экспериментальной наукой, методы и подходы которой особенно необходимы для построения объяснительных моделей в плохо формализованных областях². При этом проведенные средствами ИИ исследование, реализующее индуктивную («восходящую») стратегию анализа данных для построения эмпирической теории на его основе, с неизбежностью сталкивается с проблемой достаточности оснований для принятия результатов такого анализа – одной из ключевых проблем индукции. Выработка критерия такого принятия опирается на три концепции истины: корреспондентной – для верифицированных предсказаний, когерентной – для порождаемых причинных гипотез, прагматической – для выбора полезных процедур [Финн, 2023, с. 27–98].

² Впрочем, нельзя отрицать, что, к примеру, нейронные сети последних поколений демонстрируют свою эффективность и полезность во многих задачах современной физики, химии, минералогии и т.д.

Очевидно, что отсутствие формальных средств для представления знаний и не всегда ясно определяемая структура самого исследования не могут привести к формированию законченной картины мира, и построенные теории всегда будут носить частичный характер. Формирование эмпирических теорий в плохо формализованных областях оказывается плодотворным при условии эффективного междисциплинарного взаимодействия специалистов ИИ и конкретной предметной области. Такое взаимодействие предусматривает взаимную адаптацию видения этой области с последующим уточнением экспертной модели и развитием формального языка для представления знаний и инструментов их анализа, поддержанных эмпирической верификацией их адекватности и состоятельности [Климов и др., 2019]. Существенную роль для успешного использования инструментов ИИ играет экспертное управление параметризацией данных, направленное формирование и пополнение БФ и интерактивный выбор эффективных процедур анализа средствами партнёрских человеко-машинных интеллектуальных систем (в том числе – на этапе их разработки). Такая деятельность служит взаимному обогащению представителей точных и плохо формализованных наук, способствуя развитию как теоретических идей, так и практической их реализации.

Список литературы

- [Автоматическое порождение..., 2020] Автоматическое порождение гипотез в интеллектуальных системах / под общ. ред. В.К. Финна. – М.: Книжный дом «ЛИБРОКОМ», 2020 (изд. 2-е стереотип.). – 528 с.
- [Блинова и др., 2016] Блинова В.Г., Добрынин Д.А., Харчевникова Н.В. Сопоставление различных моделей исследования связи «структура – астмагенная активность» // НТИ. Сер. 2. – 2016 – № 2. – С. 23-28.
- [Гусакова, 2004] Подход к решению задач атрибуции исторических источников с помощью ДСМ-метода // Новости искусственного интеллекта. – 2004. – № 3. – С. 42-48.
- [Гусакова и др., 2010] Гусакова С.М., Комаров А.С. Интеллектуальная система для решения идентификационной задачи в почерковедении // Искусственный интеллект и принятие решений. – 2010. – № 4. – С. 49-54.
- [Гусакова и др., 2016] Гусакова С.М., Михеенкова М.А. Интеллектуальный анализ данных как инструмент формирования структуры социума // НТИ. Сер. 2. – 2016. – № 8. – С. 9-18.
- [Гусакова и др., 2022] Гусакова С.М., Огорельцева Н.В. Развитие интеллектуальной системы поддержки научных исследований в криминалистике // Труды Двадцатой Национальной конференция по искусственному интеллекту с международным участием КИИ-2022. – М.: РАИИ, 2022. – Т. 2. – С. 130-140.
- [Забезайло, 2022] Забезайло М.И. Три вопроса (на понимание), адресованные «товарищам по партии» // Труды Двадцатой Национальной конференция по искусственному интеллекту с международным участием КИИ-2022. – М.: РАИИ, 2022. – Т. 1. – С. 280-289.

- [Климов и др., 2019] Климов И.А., Климова С.Г., Михеенкова М.А. Формирование зоны общего языка в междисциплинарном взаимодействии: рефлексии по поводу опыта сотрудничества представителей точных и социальных наук // Социологический журнал. – 2019. – Т. 25, № 3. – С. 117-132.
- [Климова и др., 2020] Климова С.Г., Михеенкова М.А., Панкратов Д.В. ДСМ-метод как метод выявления детерминант социального поведения // Автоматическое порождение гипотез в интеллектуальных системах. – М.: Книжный дом «Либроком», 2020. – С. 410-427.
- [Митин и др., 2020] Митин А.В., Забежайло М.И. Инновационный подход к поиску новых веществ с заданными физ.-хим. свойствами на основе ИИ и квантово-химических расчетов. // Конференция памяти академика О.М. Белоцерковского. – М.: МФТИ, 2020.
- [Очан, 1942] Очан Ю.С. Некоторые вопросы эквивалентности семейства множеств // Известия Академии наук. – 1942. – № 6. – С. 171-188.
- [Поппер, 2019] Поппер К.Р. Вся жизнь – решение проблем. О познании, истории, политике. Ч. I: Вопросы познания природы: пер. с нем. – М.: УРСС: ЛЕНАНД, 2019. – 200 с.
- [Риккерт, 1998] Риккерт Г. Науки о природе и науки о культуре. – М.: Республика, 1998. 413 с.
- [Финн и др., 1993] Финн В.К., Михеенкова М.А. Некоторые проблемы обобщенного ДСМ-метода автоматического порождения гипотез // Семиотика и информатика. Вып. 33. – М.: Наука, 1993. – С. 136-163.
- [Финн и др., 2016] Финн В.К., Шестерникова О.П. О новом варианте обобщенного ДСМ-метода автоматизированной поддержки научных исследований // Искусственный интеллект и принятие решений. – 2016. – № 1. – С. 75-63.
- [Финн, 2023] Финн В.К. Интеллект, информационное общество, гуманитарное знание и образование. – М.: ЛЕНАНД, 2023 (Изд. стереотип.).
- [Brachman et al., 2004] Brachman R., Levesque H. Knowledge representation and reasoning. – Elsevier, 2004. – 416 p.
- [Chao et al., 2023] Chao Y., Xuejing Z., Hankz H. Z., Hai W., Weilin L. Reinforcement Learning with Knowledge Representation and Reasoning: A Brief Survey // arXiv:2304.12090.
- [Gao et al., 2023] Gao L., Guan L. Interpretability of Machine Learning: Recent Advances and Future Prospects // arXiv:2305.00537.
- [Gross et al., 2020] Gross E.R., Gusakova S.M., Ogorelstova N.V., Okhlopina A.N. The JSM-System of Psychological and Handwriting Research on Signatures // Automatic Documentation and Mathematical Linguistics. – 2020. – Vol. 54, No 5. – P. 260-268.
- [McCarthy, 2014] McCarthy J. Epistemological problems of artificial intelligence // Readings in Artificial Intelligence (Ed. By Webber, B.L., Nilsson, N.J.). – Ca.: Morgan Kaufmann, 2014. – P. 459-465.
- [Wang et al., 2022] Wang H., Li W., Jin X., Cho K., Ji H., Han J., Burke M.D. Chemical-reaction-aware molecule representation learning // Proceedings of International Conference on Learning Representations (Virtual), 25–29 April, 2022.

К ПРОБЛЕМЕ ОБЪЯСНЕНИЯ РЕЗУЛЬТАТОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

М.И. Забежайло (*m.zabzhailo@yandex.ru*)

Федеральный исследовательский центр

«Информатика и управление» РАН, Москва

Московский физико-технический институт

(национальный исследовательский университет), Долгопрудный

Обсуждаются возможности оценивать приемлемость результатов, формируемых системами искусственного интеллекта (ИИ) в процессе интеллектуального анализа данных (ИАД). Порожденные системой ИАД объяснения – один из удобных «инструментов» такого оценивания. Рассмотрены некоторые известные подходы к формированию объяснений. Особое внимание уделено каузальным объяснениям, в основе которых - извлекаемые в процессе ИАД из данных эмпирические зависимости причинно-следственного типа. Приведены некоторые алгоритмические свойства каузальных объяснений.

Ключевые слова: искусственный интеллект, интеллектуальный анализ данных, эмпирические зависимости в накапливаемых данных, объяснения в системах ИИ, анализ причинности.

Введение

Проблема доверия к результатам компьютерного анализа данных – наличия как соответствующих «инструментов», так и собственно порождаемой с их помощью оценки «приемлемости» результатов, формируемых компьютером в процессе анализа данных, – хорошо известна уже много десятилетий. Еще в эпоху так называемой *первой волны* ИИ-исследований и разработок, когда в центре внимания были так называемые системы, основанные на правилах (их часто также называют продукционными экспертными системами) стала активно обсуждаться проблема *интерпретации* и *объяснения* порождаемых такими ИИ-инструментами результатов.

Еще тогда достаточно быстро было осознано, что ответы на вопросы «КАК?» и «ПОЧЕМУ?» (системой ИИ получен тот или иной результат) имеют принципиально разную «природу». В большом количестве были идентифицированы дефекты смысла в трассировках «вывода» в продукционных экспертных системах (в «треках» последовательно использован-

ных продукций, показывающих как именно был получено конкретное заключение). Оказалось, что демонстрация того, «КАК» порождено заключение, в общем случае не ведет к неоспариваемому ответу на вопрос «ПОЧЕМУ?» (этому заключению можно доверять). Чуть позже – с развитием работ по автоматическому доказательству теорем (см., например, [Маслов, 1964], [Robinson, 1965] и др.) осознанная еще в продукционных экспертных системах взаимосвязь между синтаксической и семантической корректностью формируемых результатов получила ясно описываемое уточнение в виде теорем о полноте в формальных системах (т.е. утверждениях вида: *все, что выводимо – истинно, и наоборот*).

Сегодня ситуация принципиальным образом изменилась: необходимо иметь возможности оперировать в открытых мирах Big Data. Оказалось, что в таких задачах традиционные представления о доверии к результатам компьютерного анализа данных (их надежности), вообще говоря, неадекватны. Ключевым преткновением оказалась проблема *репрезентативности* обучающих выборок, пример которой легко видеть в классическом статистическом анализе данных. Действительно, хорошо известна схема: формируем генеральную совокупность, выборки из которой и будут предметом дальнейшего анализа и использования, при этом рассматриваем лишь репрезентативные выборки из генеральной совокупности, «обучение» на которых и обеспечивает нам требуемую надежность результатов, порождаемых системой анализа данных. Несложно убедиться, что в случае работы с открытыми – постоянно пополняемыми новой информацией – коллекциями данных под сомнение попадает само понятие *генеральной совокупности*.

В последние несколько лет активно развиваемым альтернативным подходом при работе с открытыми данными стали так называемые частичные эмпирические теории [TW AI, 2018], представляющие собою компактные описания накапливаемых эмпирических данных семействами формируемых из них зависимостей, причем так, что каждый элемент имеющейся коллекции данных может быть представлен как логическое следствие таких зависимостей. Вся текущая совокупность этих зависимостей и образует эмпирическую теорию, разумеется, динамически перестраиваемую с приходом в описываемую коллекцию новых данных. Экстраполируемость подобных зависимостей на вновь поступающие факты позволяет предлагать для них «объяснения» в контексте уже имеющейся эмпирической теории.

1. Немного об истории вопроса

Проблема объяснения в той или иной форме развивается в целом ряде как классических, так и современных исследований. С точки зрения логической специфики вопросов «КАК?» и «ПОЧЕМУ?» она достаточно де-

тально разобрана в серии работ, наиболее известной из которых, по-видимому, можно считать книгу Н. Белнапа и Т. Стила «Логика вопросов и ответов» [Белнап и др., 1981].

Объяснительные конструкции дедуктивного характера – так называемое *Scientific Explanation* (научное объяснение) – были предложены и подробно рассмотрены в свое время К. Гемпелем и П. Оппенхеймом [Hempel et al., 1948].

Ч.С. Пирс [Пирс, 2005] ввел в исследовательскую практику понятийную конструкцию так называемое *абдуктивного объяснения*. В предложенной им схеме объяснение фактам *B* строится поиском такой гипотезы *A*, что из *A* дедуктивно следует *B*, и при этом *A* – истинно. Пример конструктивной процедуры порождения таких *A* по заданным *B* можно найти, в частности, в [Финн, 2010 и др.].

В новейшие времена разносторонние исследования этой проблемы были предприняты в программе *Explainable Artificial Intelligence (XAI) – Объясняемый ИИ* – американского агентства перспективных военных исследований DARPA, начиная примерно с середины 2010-х годов [Gunning et al и др.]. В части полученных здесь результатов одним из наиболее продвинутых стало направление, в рамках которого объяснения формируемых искусственной нейронной сетью (ИНС) заключений строятся путем комплексирования данной ИНС с той или иной другой моделью машинного обучения (МО), например, регрессионного и т.п. типа. Значимые параметры комплексированной с ИНС модели МО и служат факторами, призванными обеспечить объяснение заключения исходной ИНС.

В последние годы широкое распространение (преимущественно – в задачах компьютерной обработки больших массивов текстовых документов) получили такие подходы, где конструкция, которая рассматривается в качестве объяснения, фактически представляет собою тот или иной вариант *интерпретации*. Например, анализируемому тексту сопоставляются определенные индексы (в частности, – ключевые слова или словосочетания, метки тематической направленности и т.п.), средствами которых и обеспечивается соответствующий вариант объяснения. Подобный вариант процедуры объяснения активно используется и в так называемых Больших Языковых Моделях (Large Language Models – см., в частности, [GPT-4], [Wu et al., 2023], [Zhao et al., 2023 и др.]).

К сожалению, в функционировании современных БЯМ в настоящее время достаточно заметны существенные смысловые дефекты, обусловленные утратой семантических взаимосвязей между, по-видимому, критически значимыми (однако не идентифицируемыми в «черном ящике» используемого комплекса ИНС) промежуточными результатами, задействованными при формировании ответа на запрос к БЯМ. Сегодня предпри-

нимаются попытки во все большей мере возложить на Большие Языковые Модели (как на вопросно-ответные конструкции) также и экспертные функции. При этом достаточно успешная автоматическая генерация объяснений-интерпретаций при работе с текстами на естественном языке (ЕЯ), когда значимые элементы семантики поданных на вход БЯМ текстов наследуются последовательно в ключевые слова, тематические индексы и т.п., что позволяет «оформить» процедуру построения объяснения в виде генерации соответствующего текста – эссе, реферата и т.п., соседствует с абсолютно неприемлемыми фактами генерации и «объяснения» для текстов иной природы, в частности, – для программного кода. Ситуация с применением БЯМ в задаче автоматической генерации текстов компьютерных программ (программного кода) на сегодняшний день абсолютно «сырая». Достаточно широко известны случаи, когда сформированные по «лекалам» обработки текстов на естественном языке (разумеется, при достаточно высоком уровне «прозрачности» их содержания) ответы на запросы о генерации текстов программ (в том числе – нуждающиеся в демонстрации корректности этих программ, а также о предъявлении сопутствующих такому программному коду объяснений его приемлемости), вообще говоря, не выдерживают серьезной критики. (И это несложно объяснить: задача отражения в индексах, сопоставляемых конкретному программному коду, *достаточных* «признаков» его корректности имеет совершенно иную семантическую «природу» и принципиальным образом отличается от задач «обычного» процессинга текстов на естественном языке).

2. Доверие к результатам ИАД в открытых мирах Big Data

Особую специфику в обсуждаемую проблематику внесла необходимость анализировать данные и принимать на этой основе решения в открытых мирах Big Data. Наряду с эффектом собственно размера (*Big*) приходится оперировать с учетом постоянно приходящих новых данных (эффектом *Open*). При этом не следует игнорировать и возможности прихода информации принципиально иного (по отношению к уже известной) характера, – того, что Н.Тaleb назвал эффектом «черного лебедя» [Тaleb, 2015].

При невозможности (в частности, – из-за эффекта *Open*) порождать доказательства надежности получаемых ИИ-системой результатов дедуктивными средствами, вполне работоспособными показали себя аргументационные схемы принятия текущих результатов интеллектуального анализа данных (ИАД) [Финн, 2020], [Вагин, 2019]. При этом порождаемые в процессе ИАД эмпирические зависимости зарекомендовали себя удобным «инструментом» объяснения накопленных данных: образующие соответствующую эмпирическую теорию зависимости («восстановленные» из

накапливаемых данных) позволяют получить каждый из фактов собранной на текущий момент коллекции данных как логическое следствие этой теории (разумеется, в ее текущем состоянии).

Вполне естественным в такой ситуации оказывается вопрос о наследуемости порождаемых подобным путем объяснений при расширении текущей коллекции накопленных данных новыми эмпирическими фактами. Основанием для такой наследуемости (устойчивости) при расширении коллекции анализируемых данных) может быть опора на анализ *причинности* – использование *причинно-следственных* связей, скрытых в обрабатываемых средствами ИАД данных.

Выявление (восстановление по накопленным эмпирическим данным) зависимостей причинно-следственного типа опирается на эвристику причинного (каузального) сходства. При поиске факторов влияния, воздействием которых обусловлено наличие заданного набора целевых характеристик – свойства – (а отсутствием их влияния – его отсутствие) у прецедентов анализируемой коллекции эмпирических данных, естественно рассчитывать, что такие совокупности факторов должны отражаться (пусть и в приближенном¹ виде) сходством описаний прецедентов, имеющих целевое свойство (их удобно называть примерами). При этом также естественно считать, что в описание ни одного из прецедентов, не обладающих целевым свойством (контрпримеров), ни одна из выделенных сходством «примеров» совокупностей факторов причинного влияния не входит.

Эвристика причинности может формализуется на базе уточнения понятия сходства прецедентов анализируемой коллекции данных средствами бинарной алгебраической операции: на формализованном представлении прецедентов задается идемпотентная, симметричная и ассоциативная бинарная операция [Кон, 1968] сходства; далее с ее помощью определяется бинарное отношение сходства описаний прецедентов; после чего на анализируемой коллекции прецедентов строятся классы сходства и классы эквивалентности (последние – по совпадающим значения вычисления операции сходства) прецедентов. Каждый из таких классов эквивалентности и порождает соответствующую эмпирическую зависимость.

Важным элементом процедурной схемы формирования объяснений причинно-следственными зависимостями оказывается идентификация артефактов переобучения. Принимая во внимание, вообще говоря, приближенный характер (используемых при формализованном описании прецедентов) «синтаксических» представлений возможных семантических особенностей наличия или отсутствия у некоторых из них целевого

¹ В частности – с точностью до используемого языка представления данных и знаний о прецедентах и «содержательности» накопленных в анализируемой коллекции прецедентов.

свойства, не следует исключать также и возможности порождения артефактов ИАД. (Полученных вычислениями, в которых не было ошибок, однако содержательно не интерпретируемых результатов). Потенциальная содержательная фальсифицируемость эмпирических зависимостей, восстанавливаемых из анализируемых данных, критически значима для формирования объяснений результатов ИАД.

В ряде случаев, дополнительные возможности для такой потенциальной (а затем уже и реально проверяемой на данных конкретной анализируемой коллекции прецедентов) фальсифицируемости может дать использование семантических особенностей изучаемой коллекции данных. В частности, совместное порождение семантически несовместимых в конкретной ситуации эмпирических зависимостей (отражающих несовместные в реальности содержательные эффекты) дает возможности организовывать взаимную фальсифицируемость соответствующих классов эмпирических зависимостей. Так, например, в работе [Забейко и др., 2021] описывается реальная ситуация, в которой в процессе диагностики эффекта так называемой *псевдопрогрессии* опухолей головного мозга человека (*ПсП*) в качестве оснований для фальсификации порождаемых в процессе ИАД из имеющихся эмпирических данных *ПсП*-зависимостей используются аналогичные (формируемые в том же математическом формализме) причинно-следственные зависимости, характеризующие эффект сокращения объема опухоли (*СОО*) в ходе штатной серии сеансов лучевой терапии. Явления увеличения (*ПсП*) и сокращения (*СОО*) объема опухоли у одного и того же пациента одновременно не совместимы.

3. Процедурные особенности формирования объяснений в ситуации с ИАД Big Data

При переходе к работе с Big Data процедуры формирования объяснения результатов ИАД должны быть соответствующим образом модифицированы. Так, в частности, в схеме Белнапа и Стила необходимо переходить к открытому множеству объектов, имеющих потенциальную возможность стать ответом на вопрос «*ПОЧЕМУ?*». Таким образом, необходимо оперировать лишь *частично-определенным* (заданным некоторым количеством примеров и контрпримеров) вопросно-ответным отношением, представляющим ответ на вопрос «*ПОЧЕМУ?*» и отражающим условие «*A не могут стать никакие B, за исключением C, ... и D*» [Белнап и др., 1981].

Процедурная техника частичных эмпирических теорий (о которой шла речь выше – см. Введение) позволяет в таких ситуациях открытости мира анализируемых данных формировать объяснения по отношению к заданному контексту – набору данных (накопленному множеству описаний прецедентов в его текущем состоянии). При этом причинно-следственные зави-

симости, порождаемые средствами ИАД на заданном контексте, дают возможности объяснять вновь регистрируемые факты (проверкой экстраполируемости уже имеющихся зависимостей на эти анализируемые факты).

Наконец, вполне работоспособной оказалась интерполяционно-экстраполяционная схема формирования ответа на вопрос «*ПОЧЕМУ?*» в текущих (и постоянно пополняемых - !) данных. Процедурные средства поиска устойчивых (наследуемых с приходом конкретных новых данных) интерполяционно-экстраполяционных зависимостей причинно-следственного характера можно использовать и при поиске таких задаваемых имеющимся контекстом объяснений, которые остаются устойчивы по отношению к заданным конкретным расширениям анализируемой коллекции прецедентов (выбранного для ИАД контекста в его текущем состоянии).

4. Общий вид схемы каузального объяснения в ИАД

В основе предлагаемой процедурной конструкции – постулат о *возможности* объяснения результатов ИАД зависимостями причинно-следственного характера, изначально скрытыми в накапливаемых данных. При этом считается, что неформальным обоснованием существования эффекта *устойчивости* объяснений (при пополнении текущего контекста новыми эмпирическими данными) является именно свойство *причинной* обусловленности анализируемых целевых характеристик прецедентов контекста. Контекст в каждом его текущем состоянии структурируется как множество примеров (прецедентов, обладающих целевыми характеристиками) и контрпримеров (прецедентов, не обладающих такими характеристиками). Каузальные (именно такие – см. Раздел 2 выше) эмпирические зависимости, восстанавливаемые из текущего контекста, позволяют разделить эту коллекцию прецедентов на пару непересекающихся множеств – примеров и контрпримеров – по схеме порождения каждого элемента текущих анализируемых данных как логического следствия в эмпирической теории, сформированной найденными в данных каузальными зависимостями.

Поиск факторов причинного влияния, формирующих соответствующие каузальные зависимости, реализует эвристику каузального сходства (см. выше – Раздел 2), включающую в себя также и идею запрета на вхождение целостных (обеспечивающих наличие целевых характеристик у соответствующих примеров) множеств факторов причинности в любой из контрпримеров. «Технически» эвристика каузального сходства оформляется в комплекс алгоритмов ИАД по той же процедурной схеме, что представлена выше – в Разделе 2 (см. также [Забейайло, 2022] и др.): от формального описания примеров и контрпримеров – к бинарной алгебраической операции сходства, далее – к отношению сходства на прецедентах

анализируемого контекста, затем – к формированию соответствующих классов сходства и их «разбору» на классы эквивалентности прецедентов (восстановлению структуры покрытий классов сходства выделенными из них классами эквивалентности прецедентов).

Существенным элементом всей схемы рассуждений при этом оказывается неоспариваемость (понимаемая как непротиворечивость каузального объяснения на текущем контексте, в том числе – его нефальсифицируемость в смысле представленной выше в Разделе 3 конструкции, формируемой на базе семантически несовместных содержательных особенностей и эффектов в анализируемых данных).

Разумеется, из множества порождаемых на расширениях анализируемого контекста каузальных объяснений естественно особо выделять такие, которые устойчивы – наследуемы при расширении текущего контекста новыми эмпирическим данными.

Подводя итоги, приведем краткое описание нескольких полезных утверждений об устойчивости каузальных объяснений результатов ИАД в открытых мирах:

- (1) Каузальная репрезентативность текущего контекста – возможность порождения из него причинно-следственных зависимостей, позволяющих каждый включенный в него прецедент описать как логическое следствие в эмпирической теории, образованной такими зависимостями, – может быть установлена алгоритмом полиномиальной² вычислительной сложности.
- (2) Наследуемость каузальной репрезентативности текущего контекста при его заданном расширении новыми прецедентами также может быть установлена алгоритмом полиномиальной вычислительной сложности.
- (3) Задача о числе каузальных объяснений, порождаемых на заданном контексте перечислительно полна.
- (4) Существование в заданном контексте каузального объяснения для заданного нового факта может быть установлена алгоритмом полиномиальной вычислительной сложности.
- (5) Наследуемость свойства каузальной объяснимости заданного факта (описания прецедента) при заданном расширении используемого конкретного контекста может быть установлена алгоритмом полиномиальной вычислительной сложности.

² От текущих размеров анализируемого контекста.

Заключение

Процедурные средства компьютерного порождения формализованных объяснений (в том числе – каузальных контекстных объяснений) – один из результативных инструментов оценки достаточности оснований для принятия получаемых в процессе ИАД выводов и заключений.

В современных ИИ-исследованиях и разработках «быстрый» ответ на фундаментальный вопрос «*Что есть истина?*» в отношении порождаемых ИИ-системой решений и рекомендаций зачастую оказывается чрезвычайно затруднённым (а в ряде случаев, по-видимому, вообще нереализуемым в рамках имеющихся у нас ресурсов).

В свое время К.Поппер обратил внимание исследовательского сообщества на то, что «... мы не застрахованы от заблуждений и, ... что стремление к достоверности (или даже к высокой вероятности) ошибочно. Однако, отсюда не следует, что было бы ошибочным стремиться к истине. Наоборот, понятие заблуждения подразумевает понятие истины как образец, которого мы, впрочем, можем не достигать. Признание погрешимости знания означает, что, хотя мы можем жаждать истины и даже способны обнаруживать ее ..., мы тем не менее никогда не можем быть уверены до конца, что действительно обладаем истинной. Всегда имеется возможность заблуждения, и только в случае некоторых логических и математических доказательств эта возможность столь незначительна, что ею можно пренебречь [Поппер, 1983]».

Математически состоятельный «механизм» формирования каузальных контекстных объяснений как раз и оказывается таким полезным «инструментарием» при решении целого ряда значимых прикладных задач.

Список литературы

- [Белнап и Стил, 1981] Белнап Н., Стил Т. Логика вопросов и ответов. – М.: Прогресс, 1981. – 288 с.
- [Вагин, 2019] Вагин В.Н. Элементы теории аргументации и её роль в интеллектуальном анализе данных // В кн.: Знания и убеждения в интеллектуальном анализе данных. – М.: Физматлит, 2019. – С. 260-377.
- [Забежайло и Трунин, 2021] Забежайло М.И., Трунин Ю.Ю. О значимости эмпирического противоречия в оценке надежности результатов интеллектуального анализа данных // НТИ. Сер. 2. – 2021. – № 6. – С. 14-20.
- [Забежайло, 2022] Забежайло М.И. Устойчивость эмпирических зависимостей и проблема объяснения результатов интеллектуального анализа данных // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник научных трудов XI Международной научно-практической конференции (ИММВ-2022, Коломна, 16-19 мая 2022 г.). В 2-х т. Т. 2. – М.: РАИИ, 2022. – С. 50-59.
- [Кон, 1968] Кон П.М. Универсальная алгебра. – М.: Мир, 1968. – 359 с.

- [**Маслов, 1964**] Маслов С Ю. Обратный метод установления выводимости в классическом исчислении предикатов // ДАН СССР. – 1964. – Т. 159, № 1. – С. 17–20.
- [**Пирс, 2005**] Пирс Ч.С. Рассуждение и логика вещей. Лекции для Кембриджских конференций 1898 года. – М.: РГГУ. 2005. – 371 с.
- [**Поппер, 1983**] Поппер К. Предположения и опровержения. Рост научного знания // Логика и рост научного знания. – М., 1983. – С. 380-392.
- [**Талеб, 2015**] Талеб Н.Н. Черный лебедь. Под знаком непредсказуемости. – М: КоЛибри, 2015. – 736 с.
- [**Финн, 2010**] Финн В.К. Индуктивные методы Д.С. Милля в системах ИИ // ИИ и принятие решений. – 2010. – Ч. I: № 3. – С. 3 -21; – Ч. II: № 4. – С. 14 -40.
- [**Финн, 2020**] Финн В.К. Стандартные и нестандартные логики аргументации // В сб.: Многозначные логики и их применения. Т. 2: Логики в системах искусственного интеллекта / под ред. В.К. Финна. – М.: URSS, 2020. – С. 59-91.
- [**Gunning et al.**] Gunning D, Vorm E., Wang J. Y., Turek M. DARPA's explainable AI (XAI) program: A retrospective. – <https://onlinelibrary.wiley.com/doi/full/10.1002/ai12.61>.
- [**GPT-4**] GPT-4. – Open AI. - <https://openai.com/research/gpt-4>.
- [**Hempel and Oppenheim, 1948**] Hempel C.G.; Oppenheim P. Studies in the Logic of Explanation // Philosophy of Science. – Apr., 1948. – Vol. 15, No. 2. – P. 135-175. – <http://links.jstor.org/sici?sici=0031-8248%28194804%2915%3A2%3C135%3ASITLOE%3E2.0.CO%3B2-E>.
- [**Robinson, 1965**] Robinson J.A. A Machine-Oriented Logic Based on the Resolution Principle // Communications of the ACM. – 1965. – Vol. 5. – P. 23-41.
- [**TW AI, 2018**] DARPA Sets Up Fast Track for Third Wave AI. – Jul 26, 2018. – <https://defence.pk/pdf/threads/darpa-sets-up-fast-track-for-third-wave-ai.569563/>.
- [**Wu et al., 2023**] Wu Sh., Irsoy O., Lu S., Dabrovolski V., Dredze M., Gehrman S., Kambadur P., Rosenberg D., Mann G. BloombergGPT: A Large Language Model for Finance. – <https://arxiv.org/abs/2303.17564>.
- [**Zhao et al., 2023**] Wayne Xin Zhao, Kun Zhou et al. A Survey of Large Language Models. – <https://arxiv.org/abs/2303.18223>.

О НЕКОТОРЫХ АКТУАЛЬНЫХ МИФАХ СОВРЕМЕННОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

М.И. Забежайло (*m.zabzhailo@yandex.ru*)

Федеральный исследовательский центр
«Информатика и управление» РАН, Москва
Московский физико-технический институт
(национальный исследовательский университет), Долгопрудный

М.А. Михеенкова (*m.mikheyenkova@yandex.ru*)

В.К. Финн (*v.k.finn@yandex.ru*)

Федеральный исследовательский центр
«Информатика и управление» РАН, Москва
Российский государственный гуманитарный университет, Москва

Рассмотрены несколько широко распространенных мифов, сопутствующих современным исследованиям и разработкам в области искусственного интеллекта (ИИ). Обсуждаются некоторые суждения относительно свойств и возможностей так называемых Больших Языковых Моделей, сильного ИИ, а также об универсальном характере подходов на базе искусственных нейронных сетей. Сформулированы три принципа организации научной экспертизы, позволяющие отделять собственно ИИ-исследования и разработки от сопутствующих им конструкций мифологического характера

Ключевые слова: искусственный интеллект, исследования и разработки, интеллектуальный анализ данных, экспертиза.

Введение

Развитие искусственного интеллекта (ИИ) как области исследований и разработок с самых первых шагов сопровождалось в том числе и различными легендами и мифами, привлекавшими внимание публики. Достаточно хорошо известно обилие различных *спекуляций* (от английского *speculate* – раздумывать, строить догадки, размышлять) на тему *может ли машина мыслить?* На протяжении десятилетий – даже после известных предложений А. Тьюринга относительно возможностей оценки конкретных результатов работы компьютера как успешной имитации интеллектуального поведения человека ([Turing, 1950] и др.), а также констатации,

190

что мы, по-видимому, как не имели, так и не имеем на данный момент неоспариваемых представлений о том, что такое *мыслить*, – эта проблематика активно пополняется все новыми суждениями. (Справедливости ради следует отметить, что подавляющее большинство таких вербальных конструкций если и имеют отношение к ИИ как собственно области научных исследований и технологических разработок, то ничтожно малое). Не менее плодотворной областью оказалась проблематика «понимания» компьютером естественного языка, где невзирая на отсутствие единого и неоспариваемого представления о смысле термина «*понимание*» (даже в ситуациях, когда речь идет о человеке), постоянно генерируются суждения, вообще говоря, не предполагающие каких-либо возможностей объективной проверки их неоспариваемости в честной дискуссии.

Сегодня область ИИ-мифологии активно пополняется новыми направлениями, примерами которых могут, в частности, стать проблематика так называемого «*сильного*» (или *общего* – от английского *general*) ИИ, множественность трактовок, вкладываемых в термин *объясняемый ИИ*, представления об универсальности и *всемогуществе* искусственных нейронных сетей (ИНС) и др.

Формируя те или иные оценочные суждения относительно подобных понятийных конструкций, было бы полезно уточнить их смысловую структуру (в частности, выделив наиболее существенные фрагменты характеризующих их семантических особенностей и убедившись при этом в их однозначном толковании).

Представляется целесообразным в рамках данной Конференции уделить внимание обсуждению и этих вопросов, в частности, для того, чтобы попытаться сформулировать (разумеется, если получится) соответствующие оценочные суждения, а также солидарную и неоспариваемую в честном сопоставлении аргументов точку зрения собранного здесь экспертного сообщества.

1. Немного истории. Общая мифология, порожденная известным вопросом «*Может ли машина мыслить?*»

В свое время даже первые опыты «освоения» (прояснения смысла используемых понятийных конструкций, постановки и анализа вопросов для исследования, ...) проблематики «*Может ли машина мыслить?*» показали, что наряду с такими бесспорно значимыми понятиями как *мыслить*, уточнения требуют наши представления еще о целом ряде базовых терминов и понятий. Так, в частности, было показано ([Маккарти и др., 1972]), что для успешного развития ИИ критически значимы представления как об объекте моделирования (а что, собственно, скрывается за термином *интеллект?*), его функциональных возможностях (например, «инструментах») порожде-

ния новых знаний из уже имеющихся данных и знаний), так и возможностях используемого инструментария моделирования – компьютерных средств и технологий. (В частности, уточнение наших представлений о возможностях и границах корректности переноса на цифровую вычислительную машину, как *синтаксически* организованного обработчика символьных конструкций, требующих учета нюансов и особенностей *семантики* анализируемых явлений). Разработчики ИИ-систем и технологий пошли именно в этом направлении, а обширная практика развития мифов и легенд на тему «Может ли машина мыслить?» – в своем.

Не менее поучительна история эволюции представлений о «понимании» компьютером естественного языка. Более полувека назад Йозеф Вейценбаум программой ELIZA продемонстрировал [Вейценбаум, 1982] возможности имитации разговорного общения, реализуемые (заметим – полностью игнорируя при этом проблему *понимания* текста или речи!) сугубо синтаксическими компьютерными средствами. Следует отметить, что при всей своей процедурной простоте (и даже – наивности) вот уже несколько десятилетий такое ИИ-решение успешно используется, например, в психиатрических клиниках при работе с тяжелыми пациентами. В современных телефонных ассистентах задача понимания речи заменена задачей принятия решения выбором из фиксированного набора альтернатив (обычно – либо прямым указанием на вариант, который следует выбрать, либо распознаванием идентификатора соответствующей альтернативы).

Не уточнив термин *понимание*, можно (даже не сужая проблему до «понимания» письменных или же устных текстов *компьютером*) развивать обширное поле спекуляций (см. выше во Введении), пополняя его все новыми мифами и легендами соответствующего характера. Опыт волшебных сказок (см. например, известную формулу «... *пойди туда, не знаю куда, и принеси то, не знаю, что*»), демонстрирует, что это может быть увлекательной и вполне востребованной деятельностью. (При этом вопрос о том, какое отношение такая деятельность может иметь к научным исследованиям и разработкам, по-видимому, будет иметь сугубо риторический характер).

2. Несколько популярных современных мифов, связываемых с проблематикой ИИ

Актуальная сегодня проблематика ИИ-исследований и разработок характеризуется рядом специфических особенностей, связанных, в частности, с необходимостью принимать те или иные управленческие решения на основе анализа Big Data, выполняемого в режиме жестких ограничений по времени. В подобных ситуациях результативными часто оказываются конкретные проблемно-ориентированные эвристики экспертов, упакован-

ные в соответствующие прикладные ИИ-системы и решения. Ключевым аспектом применения таких ИИ-инструментов оказывается возможность убедиться в корректности порождаемых результатов (т.е. получить в каждом конкретном случае ответ на классический вопрос «*Что есть истина?*»). При этом традиционные средства верификации получаемых результатов могут оказаться попросту не применимыми. (Например, как по зафиксированной в камере Вильсона вилочке треков проверить и убедиться, что в данном эксперименте наблюдался именно бозон Хиггса?). Таким образом, особую актуальность приобретает проблема выработки *критерия достаточности оснований* для *принятия* решений, предлагаемых ИИ-системой в каждом конкретном случае. Именно эта возможность (в том числе – наличие соответствующих инструментальных средств) оценки достаточности оснований для принятия порождаемых решений оказывается критически значимым шагом при движении к ответу на вопрос «*Что есть истина?*» и в данном конкретном случае применения ИИ-систем. В свою очередь, отсутствие «инструментов» такого типа может служить своего рода демаркационной линией между областью ИИ исследований и разработок с одной стороны и сферой ИИ-мифологии с другой. Обсудим подробности рассматриваемой ситуации на нескольких актуальных примерах.

3. Большие Языковые Модели как прорыв в области ИИ-исследований и технологий

Весною нынешнего года всеобщее внимание привлекли к себе так называемые Большие Языковые Модели (а англоязычной версии – Large Language Models, использующие Reinforcement learning from Human Feedback – обучение с подкреплением от человека-эксперта). Ряд экспертов даже заговорил о выдающемся прорыве в области ИИ-исследований и разработок в связи с успехами БЯМ и демонстрирующих впечатляющие результаты ИИС-технологиях (в частности – все тех же RLHF, диффузных и др.).

Однако, несложно убедиться, что успехи использования БЯМ [GPT-4], [Wu et al., 2023], [Zhao et al., 2023] в тех или иных конкретных ситуациях в значительной мере определяются тем, сколь жестко и однозначно заданы требования к формируемому ими результату. Так в задачах генерации по тексту изображения (сопоставляемой ему картинке) или же, наоборот, текста по изображению (текстового описания картинке), как и в задачах генерации текста по текстам (формирования эссе на заданную тему) приемлемым может считаться достаточно широкий набор результирующих объектов. (Хотя даже в этом случае известны достаточно показательные конфузы [S-Pic]).

При переходе к более сложным приложениям, в частности, – к генерации ответов на вопросы, ситуация коренным образом изменяется. Если при работе с «простыми» запросами (поиск датировок, фактов, географических объектов, котировок и т.п.) результаты в подавляющем большинстве случаев остаются на достаточно высоком уровне (все фактически сводится к поиску соответствующей сущности в обучающих данных), то при ответах на более сложные запросы (в т.ч. требующие рассуждений и вывода, сложных математических расчетов и т.п.) ситуация принципиально иная. Критически значимым оказывается вопрос о том, в какой мере сгенерированный средствами БЯМ ответ может восприниматься как *истинный*? (Например, как специалист по электродинамике сплошных сред должен воспринимать полученный от БЯМ ответ на вопрос о тех или иных количественных характеристиках теплопереноса с помощью сжимаемой жидкости, рассчитанных в модели Навье-Стокса, или же как специалист по кибербезопасности должен отнестись к сформированному БЯМ ответу на вопрос об оценках устойчивости ИНС заданной архитектуры к внешним вредоносным атакам, например, FGSM¹-типа?)

Как в подобных ситуациях убедиться если не в истинности, то хотя бы в достаточности оснований для принятия результатов, предлагаемых БЯМ в виде ответа на заданный ей запрос? Если на запрос Q используемая языковая модель LLM порождает ответ $LLM(Q) = A$, имеются ли (в частности, – математически корректные) возможности продемонстрировать истинность утверждения «формула $[LLM(Q) = A]$ – истинна» (т.е. «*A есть истинный ответ на вопрос Q , сформированный средствами языковой модели LLM*»)? Некоторые представления о характере трудностей, которые могут возникнуть при проверке истинности (поиске доказательства) формул такого вида, позволяет оценить опыт использования математической техники автоматического доказательства теорем [Маслов, 1964], [Robinson, 1965]. Как быть с корректностью преобразований семантики (от запроса – к ответу) в процессе синтаксических манипуляций, выполняемых БЯМ? В только что названной выше области – автоматической дедукции – такие «гарантии» семантической корректности при не содержащих ошибок синтаксических преобразованиях дают утверждения класса теорем о полноте. А как быть в случае работы с БЯМ'ами (в том числе – с учетом постоянного пополнения накапливаемых в них данных новыми сведениями)?

¹ Fast Gradient Sign Method – метод, в котором атакуется механизм оптимизации функции потерь ИНС малыми вредоносными влияниями на знак градиентного приращения.

Не вызывает сомнений, что обозначенные выше проблемы с обеспечением семантической корректности синтаксических преобразований анализируемых данных, выполняемых компьютерной системой, хорошо известны профессионалам. В таком случае, чем же можно было бы объяснить сегодняшний ажиотаж вокруг разработки и использования БЯМ? Комментируя текущую ситуацию, некоторые эксперты обращают внимание на явно обозначившиеся в последние годы серьезные барьеры на сложившемся уже примерно за полвека пути развития больших поисковых систем. С приходом в середине 90-х годов прошлого века эпохи Big Data ставший стандартом *de facto* тренд развития реляционных систем управления базами данных (RDBMS) успешно трансформировался в технологии так называемых Хранилищ Данных (Data Warehouse), где естественный рост объемов хранимой информации отрабатывался за счет технологической организации мета-данных, связывающих в единое Хранилище отдельные БД. Однако дальнейший рост объемов хранимых (и требующих поддержки «быстроты» ответов на каждый запрос) сведений столкнулся с процедурным барьером нормализации хранимых данных (комплекс так называемых процедур ETL – Extract, Transfer and Load) – с ограничениями, обусловленными сложностью вычислений при поиске ответа на запрос (в частности, доказуемо трудно-разрешимыми задачами в поиске ответов на запрос средствами реляционной алгебры, необходимостью регулярной перестройкой системы ключей и зависимостей при пополнении хранимой в DWH информации новыми сведениями и др.).

В текущей ситуации ИНС-технологии поддержки функционирования БЯМ позиционировали себя вполне результативной альтернативой уже упоминавшимся RDBMS- и DWH-технологиям при работе со сверхбольшими (и постоянно растущими) объемами данных. При этом лидеры этого направления (Google, Microsoft, Bloomberg и разработчики китайского национального поисковика) оказались и основными бенефициарами в перераспределении глобального рынка больших поисковых систем, 1% которого оценивается сегодня экспертами примерно в \$2B.

4. Сильный Искусственный Интеллект

На самых разных уровнях сегодня с той или иной степенью подробности обсуждается идея так называемого сильного ИИ (иногда также используется термин *artificial general intelligence* – *общий ИИ*). Введением в обиход этого термина исследовательское сообщество обязано двум вполне квалифицированным (в своих областях знания) специалистам – философу Джону Сёрлу и физика Роджеру Пенроузу. Первый определил сильный ИИ как «...соответствующим образом запрограммированный компьютер с нужными входами и выходами и будет разумом, в том смысле, в

котором человеческий разум – это разум» («... *the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states*») [Searle, 1980]. Фактически, аналогичное определение поддержал и Нобелевский лауреат 2020 года по физике Р. Пенроуз [Пенроуз, 2011]. Термин завораживает заложенным в нём потенциалом, однако при попытке перейти от термина к понятию становится очевидным его метафорический – чтобы не сказать «спекулятивный» – характер.

Не вступая в прямую конфронтацию с этими уважаемыми специалистами, все же представляется целесообразным обратить внимание на тот факт, что представленная выше точка зрения на сильный ИИ (СИИ) не дает операционно-неоспариваемого определения понятия СИИ, позволяющего, в частности отделить СИИ от «несильного» или же сравнивать различные версии СИИ по силе. Общий обзор накопленных представлений о СИИ в различных областях – от компьютерных наук и машинного обучения до системной нейронауки, психологии и когнитивных наук, – а также описание существующих подходов к его созданию [Ведяхин и др., 2021] не позволяют сформулировать критерии для решения указанной проблемы, а лишь маскируют их отсутствие.

Более того, если встать на точку зрения, предложенную еще «отцами-основателями» ИИ (см., например, [Маккарти и др., 1972] и др.), и рассматривать ИИ как имитацию и усиление функций естественного интеллекта (ЕИ) средствами компьютерных технологий, попытки уточнить термин СИИ (и перевести обсуждение этого понятия из области мифологии в область научных исследований и разработок) упираются в несколько фундаментальных вопросов. В частности, если СИИ – это компьютер, обладающий всеми свойствами человеческого разума, то:

- Что именно и как будет соответствовать в его работе функциям высшей нервной деятельности ЕИ (и в частности, – интуиции, целенаправлению, эмоциональному интеллекту и т.п.)?
- Как именно этот инструмент синтаксической обработки данных будет обеспечивать обработку семантически «нагруженных» данных, взаимно-однозначное представление которых синтаксическими средствами затруднено (или же вообще невозможно)?

Проблема неэквивалентности ЕИ и решений, предлагаемых исследованиями в области СИИ, упирается, прежде всего, в огромный разрыв между характерным для последнего «ментально интеллектуальным или логико-математическим» интеллектом и сложными аспектами и различными типами человеческого интеллекта – эмоциональным, социальным, творческим, инновационным, морально-этическим [Cichoński et al., 2021]. И преодоление этого разрыва с помощью создания систем «множественного интеллекта» пока что видится проблематичным.

Приведенные соображения не ставят своей целью воспрепятствовать существованию (и процветанию) мифа о СИИ. Однако, дают определенные основания оставить (по крайней мере на время использования при разработке ИИ-систем и решений цифровых вычислительных машин) обсуждение этого понятия за пределами ИИ как области исследований и разработок.

5. Миф о «всемогуществе» ИНС: работа с малыми обучающими выборками данных

Выдающиеся достижения технологии ИНС в решении прикладных задач сегодня абсолютно бесспорны. Однако, даже принимая во внимание универсальный характер ИНС-подхода с точки зрения вычислимости по Тьюрингу (как известно [Нуõtyniemi, 1996], рекуррентные ИНС (РИНС) – полны по Тьюрингу), следует учитывать и ряд дополнительных «отягчающих» обстоятельств, сопутствующих их применению. Так, например, совершенно открытая область здесь – оптимизация вычислительной сложности РИНС-представлений конкретных алгоритмов анализа данных и поддержки принятия решений (проблема сложности вычислений при представлении произвольного алгоритма средствами РИНС). Тем не менее, следует отметить, что миф о «всесилии» и универсальности ИНС-подхода постоянно привлекает на свою сторону все новых сторонников и евангелистов.

Между тем, весьма чувствительным оказался, например, вопрос о результативности ИНС в некоторых конкретных случаях анализа данных. Речь – о проблеме результативности обработки малых обучающих выборок прецедентов (ведь реальная ситуация во многих исследовательских лабораториях выглядит примерно так: на текущий момент данных ровно столько, сколько накоплено, и с ними необходимо работать, получая аргументированные результаты).

Предлагаемое (см., например, [Wang et al., 2020], [Wu et al., 2023] и др.) и активно используемое в подобных ситуациях ИНС-решение – технологии так называемых Few-Shot Learning (FSL) и One-Shot Learning (OSL), где заранее пред-обученной ИНС дополнительно предлагается (в порядке «до-обучения») небольшое количество проблемно-ориентированных примеров (в предельном случае OSL – всего один). Так, в частности, имплементация FSL-подхода в большой языковой модели BloombergGPT [Wu et al., 2023] характерна тем, что семантические особенности предметной области (где большинство запросов касаются конкретных сведений о точно описанных сущностях – котировках, валютных курсах, биржевых индексах и т.п.) – обеспечивают в целом ряде случаев успех использования схемы FSL/OSL. Общих «навыков» системы BloombergGPT в поиске приемлемых (см. вы-

ше – Раздел 3) ответов на запросы такого типа, полученных в процессе предварительно проведенного обширного пред-обучения, оказывается достаточно, чтобы, в частности, по нескольким примерам значений кросс-курсов заданных валют на заданные даты отвечать на запрос о конкретном соотношении, например, доллара к евро на интересующую пользователя дату.

Однако эксперименты в более широком спектре предметных областей и задействованных при этом ИНС-инструментов анализа данных показывают, что в части случаев использование FSL/OSL-схемы позволяет улучшить результаты по отношению к использованию исходной пред-обученной ИНС, в то время как в части – наоборот, к их ухудшению. (среди очевидных причин такой нестабильности – зависимость от конкретного типа архитектуры ИНС-решения, от характера используемой на этапе пред-обучения «настроечной» выборки, сопоставимости с ее характеристиками «специфики» до-обучающих примеров и др.). Тем мне менее, на активную позицию евангелистов мифа о «всемогуществе» и универсальности ИНС-подходов это вряд ли повлияет существенным образом.

Заключение

Подводя итоги предпринятому здесь беглому анализу некоторых актуальных мифов, сопровождающих искусственный интеллект как область исследований и разработок, вернемся к исходной интенции: есть ли возможность разделять собственно научные исследования и разработки в области ИИ и сопутствующие им мифы и легенды? Не оспаривая позитивной роли подобной мифологии (например, в части привлечения внимания общества к реальным проблемам и трудностям разработки практически значимых ИИ-систем и решений), представляется полезным сформулировать три принципа организации профессиональной экспертизы, которая как раз и призвана надежно отделять области научных исследований и разработок от сферы сопутствующих им мифов и легенд.

1. В экспертное обсуждение необходимо вводить *все* возникающие в сообществе исследователей позиции и аргументы. Классическая точка зрения ([Mill, 1860], [Ортега-и-Гассет, 2016] и др.), утверждающая, что *монополии на истину* нет ни у кого, представляется неоспоримой.
2. Все включаемые в обсуждение экспертами термины и понятия должны быть четко определены (иметь операционально состоятельные определения, которые позволяют отделить обсуждаемую сущность от остальных, в том числе – и «близко расположенных»). Так, например, физики в свое время договорились (см., в частности, [Ландау, 1962], [Абрикосов, 1965] и др.) иметь дело лишь с наблю-

даемыми на опыте сущностями, каждый шаг преобразований которых средствами того или иного математического аппарата должен быть корректно интерпретируем в терминах и понятиях изучаемой предметной области (включая используемые – наблюдаемые на опыте – сущности).

3. Решение об экспертной оценке анализируемого в подобной экспертизе термина, понятия, явления должно формироваться в «конкурсе» аргументов – доводов ЗА и ПРОТИВ, в котором финальный результат есть неоспариваемая в честной дискуссии позиция, базирующаяся на максимальных по «силе» аргументах.

Часто наличие и корректное использование такой аргументационной схемы оценки достаточности оснований для принятия результатов оказываются критически значимым элементом научной экспертизы. Разумеется, это утверждение имеет прямое отношение и к оценке результатов современных ИИ-исследований и разработок, где формирование ответа на вопрос «*Что есть истина?*» – *Корректно или же некорректно решение, сформированное ИИ-системой?* – может оказаться (как, например, в случае с БЯМ – см. выше, Раздел 3) не только весьма ресурсо-затратным, но и в целом ряде случаев – вообще затруднительным.

Список литературы

- [Абрикосов, 1965] Абрикосов А.А. Академик Л.Д. Ландау. – М.: Наука. – 1965. – 48 с.
- [Ведяхин и др., 2021] Ведяхин А.А. и др. Сильный ИИ. На подступах к сверхразуму. – М.: Альпина. – 2021. – 232 с.
- [Вейценбаум, 1982] Вейценбаум Дж. Возможности вычислительных машин и человеческий разум. От суждений к вычислениям: пер. с англ. – М.: Радио и связь, 1982. – 368 с.
- [Ландау, 1962] Ландау Л.Д. Фундаментальные проблемы (В сб.: Смородинский Я.А. Теоретическая физика 20 века. – М.: Иностранная литература, 1962. – 443 с.). – С. 285-291.
- [Маккарти и Хейес, 1972] Маккарти Дж., Хейес Р. Некоторые философские проблемы в задаче построения искусственного интеллекта // В: кибернетические проблемы бионики. – М.: «МИР». – 1972. – С. 40-88; McCarthy J., Hayes P.J. Some Philosophical Problems from the Standpoint of Artificial Intelligence // Machine Intelligence. – 1969. – No. 4. – P. 463-502.
- [Маслов, 1964] Маслов С.Ю. Обратный метод установления выводимости в классическом исчислении предикатов // ДАН СССР. – 1964. – Т. 159, № 1. – С. 17-20.
- [Ортега-и-Гассет, 2016] Ортега-и-Гассет Х. Восстание масс. – М.: АСТ, 2016. – 256 с.
- [Пенроуз, 2011] Пенроуз Р. Новый ум короля. О компьютерах, мышлении и законах физики (The Emperor's New Mind. Concerning Computers, Minds and The Laws of Physics. 1989). – М.: УРСС, ЛКИ, 2011. – 402 с. – <https://avidreaders.ru/read-book/novyuy-um-korolya.html>.

- [Cichocki et al., 2021] Cichocki A., Kuleshov A.P. Future Trends for Human-AI Collaboration: A Comprehensive Taxonomy of AI/AGI Using Multiple Intelligences and Learning Styles // Computational Intelligence and Neuroscience. – Vol. 2021. – <https://doi.org/10.1155/2021/8893795>.
- [GPT-4] GPT-4. – Open AI. - <https://openai.com/research/gpt-4>.
- [Hyötyniemi, 1996] Hyötyniemi, H. Turing machines are recurrent neural networks // Proceedings of STEP'96 (1996). – Publications of the Finnish Artificial Intelligence Society: 13–24. – <https://research.aalto.fi/>.
- [Mill, 1860] Mill, John Stuart (1860). On Liberty (2 ed.). London: John W.Parker & Son. – Милль Дж. О свободе / Пер. с англ. А. Фридмана // Наука и жизнь. – 1993. – № 11. – С. 10-15; – № 12. – С. 21-26.
- [Robinson, 1965] Robinson J.A. A Machine-Oriented Logic Based on the Resolution Principle // Communications of the ACM. – 1965. – Vol. 5. – P. 23–41.
- [Searle, 1980] Searle, J. Minds, brains, and programs // Behavioral and brain sciences. – 1980. – Т. 3, № 3. – P. 417-457. – <http://cogprints.org/7150/1/10.1.1.83.5248.pdf>.
- [S-Pic] Лосось, идущий вверх по реке (изображение, сгенерированное БЯМ). – https://yandex.ru/images/search?img_url=https%3A%2F%2Fscientificrussia.ru%2Fimages%2Ffn%2F2pzn-full.jpg&lr=213&pos=4&rpt=simage&source=serp&text=%D0%BA%D0%B0%D1%80%D1%82%D0%B8%D0%BD%D0%BA%D0%B0%20%D0%BB%D0%BE%D1%81%D0%BE%D1%81%D1%8C%2C%20%D0%B8%D0%B4%D1%83%D1%89%D0%B8%D0%B9%20%D0%B2%D0%B2%D0%B5%D1%80%D1%85%20%D0%BF%D0%BE%20%D1%80%D0%B5%D0%BA%D0%B5.
- [Turing, 1950] Turing, A.M. Computing machinery and intelligence // Mind, vol. LIX. – 1950. – No. 236. – P. 433-460.
- [Wang et al., 2020] Wang Y., Yao Q., Kwok J., Ni L.M. Generalizing from a Few Examples: A Survey on Few-Shot Learning. – 2020. – arXiv:1904.05046. – <https://arxiv.org/pdf/1904.05046.pdf>.
- [Wu et al., 2023] Wu Sh., Irsoy O., Lu S., Dabrovolski V., Dredze M., Gehrmann, S., Kambadur P., Rosenberg D., Mann G. BloombergGPT: A Large Language Model for Finance. – <https://arxiv.org/abs/2303.17564>.
- [Zhao et al., 2023] Wayne Xin Zhao, Kun Zhou et al. A Survey of Large Language Models. – <https://arxiv.org/abs/2303.18223>.

ИССЛЕДОВАНИЕ ВЫБОРОК WOS С ПОМОЩЬЮ НЕЙРОСЕТЕВЫХ КЛАССИФИКАТОРОВ

Н.А. Назаров (*straider105@gmail.com*)

В.О. Толчеев (*tolcheevvo@mail.ru*)

Национальный исследовательский университет «МЭИ», Москва

В работе проводится анализ текстовой коллекции, составленной из базы данных Web of Science (WOS). Исследование направлено на оценку качества классификации научных статей с помощью современных нейросетевых классификаторов (все статьи заданы в виде библиографических описаний, которые включают названия и аннотации). Несмотря на лидерство нейросети BERT по многим направлениям обработки естественного языка, ее результаты на коллекции WOS существенно зависят от «сложности» выборки (количества классов и их сбалансированности). Так, на «простой» выборке (7 классов) значение F1-меры составило 0.908, а на «сложной» выборке (134 класса) – 0.815. В работе также исследованы менее ресурсозатратные нейросетевые подходы. В частности, высокие показатели качества на коллекции WOS достигнуты при использовании модели FastText для представления документов и сверточной нейросети CNN для классификации («простой» выборки $F1=0.888$ и «сложной» выборки $F1=0.766$). При этом совместное применение нейросетей FastText и CNN позволяет существенно снизить время обучения и скорость классификации по сравнению с BERT.

Ключевые слова: классификация текстов, FastText, BERT, линейный классификатор, CNN, LSTM, машинное обучение.

Введение

Классификация представляет собой одну из важнейших задач в области машинного обучения, находящую широкое применение в обработке и анализе различных типов данных. В частности, классификация текстовых документов играет ключевую роль в таких сферах, как обработка естественного языка, информационный поиск и текстовый анализ. Однако при работе с текстовыми данными возникают серьезные проблемы, связанные с пересечением классов, высокой размерностью признакового пространства и сложностью учета контекста. Особые трудности возникают в случае роста числа классов и их несбалансированности. В ряде прикладных

задач при описании больших объемов информации используются сложные иерархические модели. В частности, такая ситуация возникает при обработке и анализе научных публикаций, распределенных по нескольким крупным предметным областям. Для решения этих нетривиальных задач разрабатываются специализированные модели представления текстовых документов (МТД) на основе нейросетевых подходов.

Целью данной работы является анализ современных нейросетевых классификаторов, таких как BERT, в контексте классификации научных статей из Web of Science (WOS) по разным категориям и областям знания, с оценкой качества классификации на основе F1-меры, выявлением возможных слабых сторон модели BERT, предложением альтернативных решений при работе со сложными и несбалансированными выборками данных, а также разработкой рекомендаций для улучшения и оптимизации нейросетевых классификаторов с учетом взаимосвязи между качеством классификации, методами обработки данных (МТД), скоростью работы и требованиями к памяти.

Требования к производительности и памяти нейросетевых моделей классификации включают ряд критериев. Среди ключевых факторов выделяются высокая скорость работы модели, эффективное использование памяти и оптимизация количества параметров модели. Для оценки производительности моделей применяются такие показатели, как время обучения, время предсказания и количество параметров модели. Эффективное использование памяти требует учета размера модели и использования памяти на графических процессорах (GPU). Оптимизация этих параметров может способствовать ускорению процесса обучения и снижению требований к памяти, что, в свою очередь, обеспечит высокую производительность и эффективность работы нейросетевых моделей.

Обзор нейросетевых методов классификации

Нейронные сети представляют собой эффективные модели машинного обучения и являются одним из основных способов классификации текстовых документов. Обычно применяются следующие архитектуры: сети прямого распространения (включая сверточные сети) и рекуррентные сети, а также их модификации и комбинации [Lai et al., 2015].

Полносвязные сети прямого распространения, такие как многослойный перцептрон, представляют собой нелинейные модели, способные формировать сложные разделяющие поверхности и обобщать входные данные. Тем не менее, они могут показывать низкое качество классификации при высокой размерности данных и сложных зависимостях, а также обладают высокой вычислительной сложностью.

Сверточные нейронные сети (CNN) изначально разрабатывались для обработки изображений, однако они также успешно применяются в классификации текста. CNN определяют закономерности в текстовых структурах и схожесть между ними, но могут терять информацию о структуре данных. CNN применяются к последовательностям произвольной длины и выявляют скрытые закономерности. Однако они имеют большое число гиперпараметров, что усложняет их настройку.

Рекуррентные сети (RNN) успешно используются для анализа текста и классификации, сохраняя большую часть информации о структуре и моделируя сложные последовательности. Они демонстрируют хорошие результаты в задачах языкового моделирования, определения частей речи и машинного перевода. В то же время, RNN сталкиваются с проблемами исчезающего и взрывающегося градиента. В связи с этим обычно используется нейросеть LSTM – модифицированная версия RNN, обучающая долгосрочным зависимостям и преодолевающая проблему исчезающего градиента. Она основана на использовании состояния ячейки и регулировании удаления информации с помощью фильтров. Широкое распространение получила также нейросеть BI-LSTM двунаправленная версия LSTM, обрабатывающая информацию о последовательности в обоих направлениях, сохраняя будущую и прошлую информацию. Преимущества RNN включают применение к последовательностям любой длины и сохранение информации о последовательности. Недостатки – проблемы исчезающего и взрывающегося градиента (решаемые в LSTM) и высокая вычислительная сложность обучения.

Также существуют модели, представляющие собой модификации и комбинации рекуррентных и сверточных нейронных сетей, направленных на решения специфических задач классификации текстовых документов. Они обладают своими уникальными характеристиками и преимуществами, которые могут обеспечить более точную классификацию и анализ данных в зависимости от требований конкретной задачи. В качестве примера рассмотрим C-LSTM и RMDL. C-LSTM использует сверточную сеть для извлечения из входной последовательности шаблонов, которые потом передаются в рекуррентную сеть. RMDL представляет собой параллельное соединение многослойного персептрона, сверточной сети и рекуррентной сети.

В процессе обучения нейронной сети для классификации текстовых документов дополнительно используется техника регуляризации Dropout. Это позволяет предотвратить переобучение модели и улучшить ее обобщающую способность. Dropout применяется в определенных слоях сети (например, после полносвязных слоев или между сверточными слоями), исключая случайную долю нейронов на каждом шаге обучения. В результате, нейронная сеть становится менее чувствительной к шуму в данных и лучше обобщает на новые, ранее не встречавшиеся примеры.

Модели представления текстовых документов

Для анализа естественного языка в машинном обучении используются векторные представления текстов [Amajd et al., 1989]. Эти модели разделяются на классические, основанные на анализе частоты слов, и нейросетевые. Недостатками нейросетевого подхода являются сложность подбора параметров и меньшая интерпретируемость. Предобученные модели помогают избежать необходимости обучения новых моделей с нуля. Развитие вычислительных мощностей позволило обучать более сложные модели на больших данных. В нейросетевых алгоритмах используются вложения слов, которые могут быть контекстуальными или без контекста. Наиболее популярные модели включают Word2Vec, Doc2Vec, FastText, ELMo, GPT, GPT-2, GPT-3 и BERT [Kilimci et al., 2019].

В работе исследованы оба вида моделей (с учетом и без учета контекста). Рассмотрим более подробно некоторые из них.

В качестве модели без учета контекста в работе исследуется FastText, в которой каждое слово представляется в виде n -грамм символов, что позволяет лучше понимать значение коротких слов и распознавать суффиксы и префиксы.

Модели с учетом контекста, исследованные в данной работе, основаны на нейросетевой архитектуре Transformer. В основе этой архитектуры лежит комбинация из слов кодировщика и декодировщика, которые применяют механизмы внимания и слои прямого распространения для обработки последовательностей.

Одной из ключевых особенностей архитектуры Transformer является введение позиционного кодирования, которое добавляет дополнительное измерение к каждому слову, указывая на его местоположение в контексте. Это позволяет учитывать порядок слов в последовательности без использования рекуррентных механизмов.

После обработки на слоях кодировщика, результаты передаются декодировщику для дополнительной обработки и формирования выходной последовательности.

К числу наиболее известных моделей на основе Transformer относятся нейросети GPT и BERT, а также их модификации.

Модель GPT (Generative Pre-trained Transformer), разработанная OpenAI, представляет собой однонаправленную языковую модель, использующую архитектуру Transformer. Исследования показали, что обучение трансформеров на больших объемах данных приводит к выдающимся результатам в задачах обработки естественного языка. Модель GPT использует структуру декодера Transformer, состоящую из 12 блоков, и может быть точно настроена на задачу с небольшим объемом данных в предметной области.

GPT-2 – улучшенная версия GPT – также применяет структуру декодера Transformer, но с увеличенной максимальной длиной последовательности до 1024, что обеспечивает большую гибкость. GPT-2 предлагает четыре предварительно обученные модели с разным количеством блоков декодера, позволяя исследователям выбирать наиболее подходящую модель для своих задач.

В нашей работе мы применяем GPT-2, поскольку эта модель обеспечивает гибкий выбор архитектур и размеров, подходящих для разнообразных задач классификации текстов. Благодаря предварительному обучению на обширных корпусах текста GPT-2 может быть успешно использована для векторизации текста и извлечения признаков, что способствует улучшению результатов классификации.

Модель BERT (Bidirectional Encoder Representations from Transformers), созданная разработчиками Google, является в настоящее время одной из самых эффективных моделей на основе архитектуры Transformer. BERT представляет собой двунаправленную модель, обучаемую на большом корпусе текста с использованием задач маскированного языкового моделирования и предсказания следующего предложения. Эта модель была адаптирована для решения множества задач, включая классификацию текстов, и показала выдающиеся результаты благодаря своей способности улавливать контекст и семантику текста [Vaswani et al., 2017].

BERT имеет несколько модификаций и адаптаций, разработанных для различных задач обработки естественного языка. Например, RoBERTa, ALBERT и DistilBERT являются вариациями BERT, которые оптимизированы для разных сценариев использования, таких как улучшение качества предсказаний, снижение количества параметров и ускорение обучения.

В задаче текстовой классификации архитектура трансформера и модели на ее основе, такие как BERT, могут обеспечить более точные результаты, учитывая их способность улавливать контекст и семантику текста. Однако они также могут быть более требовательными к вычислительным ресурсам и времени обучения, что может быть критичным для некоторых приложений. Важно выбирать подходящую модель представления текстовых документов и архитектуру нейросети для решаемой задачи, учитывая специфические требования и ограничения [Кусакин и др. 2022].

На практике обычно используются предобученные модели GPT и BERT, натренированные на больших корпусах текстов. Для адаптации предобученной модели к конкретной задаче классификации необходимо провести дообучение на специально сформированных выборках, отражающих специфику решаемой задачи.

Дообучение (fine-tuning) представляет собой технику машинного обучения, при которой предварительно обученная модель адаптируется к новой задаче путем дополнительного обучения на специфических для этой

задачи данных. В контексте текстовой классификации дообучение применяется для адаптации модели, натренированной на больших корпусах текста, к конкретной задаче классификации, связанной с определенной областью знаний или набором категорий.

Преимуществами использования дообучения являются улучшение производительности и снижение времени обучения модели по сравнению с построением модели с нуля. Это достигается за счет инициализации весов модели значениями, полученными в ходе предварительного обучения на больших корпусах текста, что позволяет модели лучше понимать общие закономерности в языке и семантику текста.

В рамках данной работы дообучение может быть применено для адаптации предобученных моделей, таких как BERT, к задаче классификации научных текстовых документов.

Оценка качества классификации

Анализ качества классификации является ключевым этапом исследования, так как он дает возможность оценить и сопоставить разные модели и методы. В данной работе для измерения качества классификации используются такие метрики, как точность (precision), полнота (recall) и F1-мера [Флах, 2015]. Сравнивая результаты по этим метрикам, можно определить наиболее подходящие методы для решения задачи классификации текстовых документов [Козлов и др. 2022].

Кроме того, важным аспектом оценки является анализ быстродействия предсказания и обучения различных моделей. Это позволяет определить, какие из них обеспечивают наиболее эффективное использование вычислительных ресурсов и могут быть применены в условиях ограничений по времени и мощности. Сравнение времени обучения и предсказания разных моделей может помочь выбрать наиболее подходящий подход, обеспечивающий баланс между качеством классификации и вычислительной эффективностью.

Описание выборки

Одним из примеров сложных датасетов для исследования является коллекция научных документов, полученная из реферативной базы научных публикаций Web of Science (WOS). Эта междисциплинарная англоязычная коллекция характеризуется иерархической двухуровневой структурой, что представляет дополнительные трудности для проведения анализа и классификации.

Коллекция WOS-46985 состоит из 46 985 библиографических описаний, извлеченных из базы данных Web of Science. В ней реализована двухуровневая система классификации, где первый уровень включает в

себя следующие предметные области (домены): Информатика, Электротехника, Психология, Механика, Гражданская Инженерия, Медицина и Биохимия. Второй уровень представлен 134 поддоменами.

Существуют также две упрощенные версии данной коллекции: WOS-11967 и WOS-5736, которые являются подмножествами WOS-46985. Однако для проведения исследования выбрана полная версия датасета, содержащая 46 985 статей. Средняя длина текстовых документов составляет 229 слов. Датасет предлагает разнообразные данные, охватывающие множество областей знаний, что способствует эффективному обучению и тестированию модели с учетом различных тематик.

Схема классификации

Рассмотрим схему классификации, использующуюся в исследовании (рис. 1).

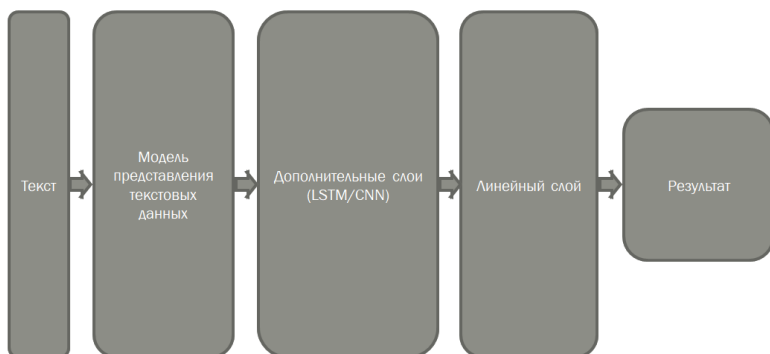


Рис. 1. Схема исследуемой модели классификации

В данной работе рассматривается подход к классификации текстовых документов, в котором на вход подаются тексты, а затем они преобразуются в векторное пространство с помощью модели представления текстовых данных, таких как FastText, GPT-2 или BERT. Модель представления текста возвращает набор векторов для каждого слова из текста. В зависимости от выбранного классификатора, последующая обработка данных может варьироваться.

Для линейного слоя классификации мы берем вектор из последнего слоя модели представления текста, который после прохождения через пулер, обрабатывающий агрегированные данные последнего скрытого состояния, передается через линейный слой с применением Dropout. На выходе получаем вектор с числом элементов, равным количеству классов.

При использовании классификатора на основе сверточных нейронных сетей (CNN) последнее скрытое состояние модели представления текста адаптируется для сверточных слоев и применяется функция активации ReLU. Затем выполняется глобальный макс-пулинг, выходные векторы объединяются и передаются через линейный классификатор с применением Dropout.

При использовании классификатора на основе долгосрочной краткосрочной памяти (LSTM) вектора последнего скрытого состояния модели представления текста передаются в LSTM-слой, который обрабатывает последовательность. Затем к последнему вектору скрытого состояния применяется линейный слой с Dropout.

Таким образом, представленный подход позволяет использовать различные модели представления текстовых данных и классификаторы для анализа текстовых документов, что может способствовать улучшению эффективности обработки данных.

Результаты классификации

Полученные результаты классификации по доменам первого уровня представлены в табл. 1.

Таблица 1

Модель	F1 Мера	Время обучения (с.)	Время предсказания (с.)
BERT + линейный слой	0.906	8241	170.6
BERT + CNN	0.908	8385	182.0
BERT + LSTM	0.904	9096	237.0
GPT-2 + линейный слой	0.858	3413	68.0
GPT-2 + CNN	0.873	3715	70.0
GPT-2 + LSTM	0.867	4596	90.0
FastText + линейный слой	0.794	2038	27.0
FastText + CNN	0.888	866	29.0
FastText + LSTM	0.811	2033	59.0

В табл. 1 время обучения и время предсказания были получены при обучении на графическом процессоре RTX 3060 Ti.

Все рассмотренные модели нейросетевых классификаторов успешно справились с задачей классификации статей из датасета WOS-46985. Модели, основанные на BERT, показали наилучшие результаты по F1-мере, с максимальным значением 0.908 для BERT + CNN. Однако они также требовали наибольшего времени обучения и предсказания, что может быть критическим фактором при выборе модели для некоторых приложений.

GPT-2 показал хорошие результаты с F1-мерой до 0.873 (GPT-2 + CNN), однако существенно уступает моделям на основе BERT. В то же время GPT-2 быстрее обучается и предсказывает классы по сравнению с BERT.

Модели на основе FastText дали разные результаты F1-меры (от 0.794 до 0.888). Особо отметим, что при использовании CNN в качестве классификатора получился лучший, после BERT результат по F1-мере, при этом время обучения и время предсказания также самые низкие из всех моделей. Это делает сочетание двух нейросетевых подходов FastText + CNN потенциально полезными для быстрой и качественной классификации.

Полученные результаты классификации по доменам второго уровня представлены в табл. 2.

Таблица 2

Модель	F1 Мера	Время обучения (с.)	Время предсказания (с.)
BERT + линейный слой	0,807	10954	183.4
BERT + CNN	0.815	12368	213.5
BERT + LSTM	0.782	12321	247.3
GPT-2 + линейный слой	0.698	3514	71.6
GPT-2 + CNN	0.758	3841	95.7
GPT-2 + LSTM	0.691	4635	101.4
FastText + линейный слой	0.604	4189	31.2
FastText + CNN	0.766	3360	35.3
FastText + LSTM	0.612	6573	74.2

При классификации на метках второго уровня (134 класса) на выборке WOS были получены следующие результаты:

1. Модели на основе BERT продемонстрировали лучшие результаты по F1-мере, достигая максимального значения 0.815 для BERT + CNN. Однако, как и в предыдущем эксперименте, эти модели требуют больше времени для обучения и предсказания.

2. GPT-2 показал результаты ниже, чем BERT, с максимальной F1-мерой 0.758 для GPT-2 + CNN. Однако обучение и предсказание GPT-2 проводится существенно быстрее по сравнению с BERT.

3. Модели на основе FastText показали наименьшее время обучения и предсказания среди всех моделей. Использование данной модели в сочетании с CNN также дало лучший результат по F1-мере после модели BERT.

Таким образом, при классификации на «сложной выборке» (134 класса) результаты по F1-мере существенно ниже (почти на 10 процентов) чем для «простой» выборки, содержащей 7 классов.

Заключение

Исследование нейросетевых моделей BERT, GPT-2 и FastText на датасете WOS-46985 с учетом меток первого и второго уровня показало следующие результаты:

1. Модели, использующие BERT, обеспечивают наилучшее качество классификации на «простых» и «сложных» выборках, однако требуют значительного времени обучения и времени предсказания.

2. GPT-2 является компромиссом между качеством классификации и временем работы модели, обеспечивая хорошие результаты при меньших временных затратах.

3. Модель FastText в сочетании со сверточной нейросетью CNN наиболее близка по F1-мере к результатам BERT, при этом обладая существенно лучшим быстродействием.

Проведенный в работе анализ различных архитектур нейросетевых классификаторов, таких как BERT, GPT-2 и FastText, позволяет более обоснованно выбирать метод классификации, исходя из характеристик датасета и требований конкретной задачи.

Как представляется, адаптация нейросетевых классификаторов к сложным и несбалансированным структурам данных является ключевым аспектом дальнейшего развития и применения этих методов. В настоящее время для ряда «сложных» выборок не удастся достичь высокого качества категоризации даже при использовании самых современных ресурсозатратных нейросетевых архитектур. На наш взгляд, перспективными направлениями исследований в этой области может стать ансамблирование моделей и применение мета-обучения с учетом специфики решаемой задачи.

Список литературы

- [Kilimci et al., 2019] Kilimci Z.H., Akyokuş S. The Analysis of Text Categorization Represented With Word Embeddings Using Homogeneous Classifiers // IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA). – 2019.
- [Флах, 2015] Флах П. Машинное обучение – Наука и искусство построения алгоритмов, которые извлекают знания из данных. – М.: ДМК-пресс, 2015.
- [Amajd et al., 1989] Amajd M., Kaimuldenov Z. and Voronkov I. Text classification with deep neural networks // Proc. CEUR Workshop. – 1989. – Vol. 17.
- [Lai et al., 2015] Lai, S., Xu L., Liu K., Zhao J. Recurrent Convolutional Neural Networks for Text Classification // In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, TX, USA, 2015.
- [Vaswani et al., 2017] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need // In 31st Conference on Neural Information Processing Systems (NIPS 2017). – Long Beach, CA, USA, 2017.
- [Кусакин и др. 2022] Кусакин И.К., Цурупа А.М., Алмакаев А.В., Романов А.Ю. Использование BERT для классификации коротких научных – М.: НТИ-2022, 2022.
- [Козлов и др. 2022] Козлов П.А., Мохов А.С., Назаров Н.А. [и др.]. Сравнительный анализ бинарных классификаторов на массиве научных публикаций // Заводская лаборатория. Диагностика материалов. – 2022.

УДК 004.912, 004.853

ПОДХОД К ПОСТРОЕНИЮ ДАТАСЕТА ДЛЯ ЗАДАЧИ ИЗВЛЕЧЕНИЯ АРГУМЕНТАТИВНЫХ ОТНОШЕНИЙ*

Е.А. Сидорова (*lsidorova@iis.nsk.su*)

Ю.А. Загоруйко (*zagor@iis.nsk.su*)

И.С. Кононенко (*irina_k@cn.ru*)

А.С. Серый (*alexey.seryj@iis.nsk.su*)

П.М. Чагина (*p.chagina@gmail.com*)

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

В работе рассмотрены методы генерации данных для обучения нейросетевых моделей и решения задач анализа аргументации. Подготовка данных включает создание аннотированного корпуса и генерацию датасета для конкретной задачи и выбранной модели обучения. Корпуса, размеченные экспертами вручную, можно пополнять синтетическими текстовыми данными, полученными на основе рефразирования или обратного перевода текстов базового корпуса. Особенностью задачи генерации набора данных является необходимость выделять в тексте компоненты и указывать их роль в составе аргумента (посылка, тезис, контраргумент). Рассмотрена роль индикаторов аргументации в задаче подготовки данных и проведено статистическое исследование информативности индикаторов на подготовленном корпусе ArgNet.

Ключевые слова: анализ аргументации, аргументативное отношение, создание размеченных данных, датасет, индикатор аргументации, сегментация текста.

Введение

Компьютерный анализ аргументации заключается в преобразовании неструктурированного текста в “цепочки” или графы связанных структурированных аргументов. Данный процесс включает такие этапы, как выявление фрагментов текста, содержащих аргументацию (Argument Detection), установление связей между компонентами аргументации (Argumentative Relation Prediction), выделение компонентов аргумента и

* Работа выполнена при финансовой поддержке РФФ (проект № 23-11-00261).

установление их ролей в составе аргумента (Argument Component Classification), классификация аргументов в соответствии со схемами аргументации, представляющими типовые модели рассуждения и убеждения (Argument Classification).

С ростом интереса научного сообщества к проблеме анализа аргументации (Argument Mining) стало появляться все больше аннотированных корпусов для разных текстовых жанров и предметных областей. Были созданы аннотированные корпуса на разных языках: а) англоязычный корпус Cornell eRulemaking (CDCP) [Park et al., 2018], состоящий из комментариев пользователей форума, посвященного электронному законодательству; б) корпус Университета Дармштадта PersEssays на немецком языке [Stab et al., 2014], включающий студенческие аргументативные эссе, новостные тексты и научные статьи; в) Потсдамский корпус микротекстов ArgMicro на немецком и английском языках [Peldszus et al., 2017] и др.

Русскоязычных корпусов с аргументативной разметкой гораздо меньше: а) в [Fishcheva et al., 2019] описан первый русскоязычный корпус, полученный переводом с английского языка корпуса ArgMicro; б) в [Fishcheva et al., 2021] авторы развили свой подход, расширив имеющийся корпус за счет автоматического перевода корпуса PersEssays и разработав объединенную схему разметки для возможности одновременно использовать тексты из обоих корпусов в соответствии с моделью; в) в [Сидорова и др., 2020], [Шина et al., 2021] описаны корпуса для анализа аргументации в научно-популярном дискурсе, аннотированные вручную экспертами в соответствии с онтологией аргументации и схемами аргументации Д. Уолтона; г) в работе [Kotelnikov et al., 2022] приведено описание процедуры сбора и ручного аннотирования корпуса текстов из социальных сетей, использованного для проведения соревнования RuARG-2022 – первого соревнования русскоязычных систем анализа аргументации. К сожалению, последний корпус пригоден только для задач предсказания или классификации позиций пользователей по определенным вопросам (stance detection и stance classification), а не для задач обнаружения аргументации в текстах.

Всего, по оценке авторов работы [Lawrence et al., 2019], на 2019 год общий объем размеченных данных составлял примерно 2,2 миллиона слов. На сегодняшний день большинство исследований проводится на указанных выше корпусах, из чего можно сделать вывод, что объем корпусов значительно не увеличивался. При этом основными языками исследований по-прежнему остаются английский и, в меньшей степени, немецкий, а сами данные представляют собой множество корпусов различных объемов, жанров и с различной разметкой, что делает комбинирование корпусов невозможным или очень затруднительным.

Целью нашего текущего исследования является подготовка датасетов для решения задач анализа аргументации в целом и извлечения аргументативных отношений, в частности. Подготовка таких датасетов включает два этапа: во-первых, создание корпуса текстов с аргументативной разметкой, во-вторых, генерацию наборов данных для конкретной задачи и выбранной модели обучения. Создание корпуса, помимо непосредственной разметки текстов, может включать методы автоматического пополнения, такие как перевод аннотированных текстов с других языков или синтез текстов на основе перефразирования или обратного перевода, когда текст переводится на промежуточный язык, как правило, английский, а затем обратно. Особенностью задачи генерации набора данных является необходимость выделять компоненты и указывать их роль в составе аргумента (посылка, тезис, контраргумент).

1. Создание и пополнение корпуса с аргументативной разметкой

При решении задач анализа аргументации мы ориентируемся на использование нейросетевых подходов, в которых объем обучающих данных напрямую влияет на качество работы системы. Основой для создаваемого набора данных послужил корпус текстов ArgNet, аннотированных специалистами на платформе ArgNetBank Studio [Сидорова и др., 2020]. Разметка аргументации в данном корпусе в среднем обладает хорошим качеством, но его пополнение – это длительный и дорогостоящий процесс. Для увеличения набора данных применялись методы автоматического перевода и перефразирования.

1.1. Особенности разметки корпуса ArgNet

Разметка проводилась в соответствии со стандартом AIF и набором схем аргументации, созданных на основе модели Д. Уолтона [Walton et al., 2008]. В рамках данной модели, аргумент – это структура, связывающая набор посылок с заключением, а тип связи каждой посылки с аргументом определяется схемой аргументации.

Корпус состоит из нескольких подкорпусов, созданных специалистами из разных областей знаний и с разными целями. Общий объем всех подкорпусов составляет 308 текстов, количество размеченных аргументов – 6740. Однако в корпусе была отмечена общая неоднородность разметки, что усложняет модель данных (проблема выделения границ утверждений) и потенциально снижает качество автоматического извлечения аргументации. Поэтому в итоговый корпус отобраны тексты в соответствии со следующими принципами.

1. Исключены тексты с абстрактной аргументацией.

2. Из текстов с неполной разметкой (отмеченных вручную) удалены фрагменты, не покрытые аргументацией.
3. Тексты с циклической разметкой проверены на корректность, циклы исключены из разметки, тексты с необоснованными циклами удалены полностью.
4. Тексты, размеченные несколькими аннотаторами, рассматриваются как эталонные, все такие аннотации считаются равноправными.

Всего корпус ArgNet включает 190 текстов, количество размеченных аргументов в нем – 6611.

1.2. Автоматическое пополнение корпуса

В первую очередь корпус ArgNet был обогащен за счет перевода корпуса ArgMicro, выполненного авторами [Fishcheva et al., 2019]. Данный корпус создавался для решения задач stance detection, т.е. классификации утверждений по их позиции, как правило “за” или “против”, по отношению к заданной теме. Вариант корпуса ArgMicro, размещенный в базе AIFdb (<http://corpora.aifdb.org/Microtext>), имеет графовую структуру, приближенную к таковой у ArgNet. Перевод был наложен на граф, который затем добавлен к имеющимся данным. Было добавлено 576 новых утверждений и 443 аргумента из 112 текстов. Расширенный корпус был обозначен как ArgNet+ArgMicro.

Другим источником данных стал немецкоязычный корпус SIMT [Romberg et al., 2021], собранный из обсуждений и предложений жителей по благоустройству городов в Германии. Корпус состоит из пяти частей. Каждое утверждение отмечено как “основная позиция” (major position, *mpos*), “посылка” или “суждение” (premise, *prem*) или утверждение без выраженной аргументации (non-arg). Перевод данного корпуса автоматическим переводчиком GoogleTranslate из пакета `deep_translator` для языка Python позволил составить русскоязычный корпус ruSIMT из примерно 16 тысяч утверждений с аналогичной разметкой для решения задач Argument Detection.

Перефразирование – это сравнительно малозатратный способ увеличить количество данных. Перефразировать текст – значит переписать его так, чтобы новый текст отличался от первоначального по форме, но был близок ему по смыслу. В случае автоматического перефразирования следует также следить за тем, чтобы новый текст выглядел как можно более естественным, т.е. похожим на написанный человеком. Перефразирование при этом может достигаться как, непосредственно, переписыванием оригинального утверждения, так и обратным переводом. В результате обратного перевода получаются достаточно естественные утверждения, близкие по смыслу к оригинальным.

Корпус ArgNet+ArgMicro был расширен за счет перефразирования. Каждое утверждение было переписано тремя разными методами. Первые два основаны на работах Сбера по обучению языковых моделей MT5 и GPT2 для перефразирования русских текстов [Fenogenova, 2021]. Третий – это перевод с русского на английский и обратно при помощи трансформерной модели от компании Facebook [Ng et al., 2019]. Таким образом, ожидалось увеличение первоначального корпуса в четыре раза.

1.3. Оценка качества сгенерированных корпусов

Чтобы оценить расширенный корпус, необходимо провести анализ смыслового и визуального сходства, а также осмысленности полученных парафраз, с точки зрения человека.

Для оценки смыслового сходства парафраз с оригиналом, применялась языковая модель cointegrated/LaBSE-en-ru, обученная векторизовать предложения так, чтобы векторы близких по смыслу предложений были близки геометрически. Это сокращенная версия модели LaBSE от Google. Смысловое сходство утверждений вычислялось как косинусное расстояние между соответствующими векторами в данной модели. На корпусе оценивались метрики MeaningSim и MeaningRand. Первая – это среднее смысловое сходство оригинальных и перефразированных утверждений. MeaningRand вычисляется как среднее попарное смысловое сходство оригинальных предложений и случайных парафраз. Она показывает, насколько однообразны по смыслу фразы из корпуса.

Метрика BLEU [Papineni et al., 2002], успешно применяемая при оценке результатов автоматического перевода, определяет визуальное сходство двух текстов. Она учитывает долю n-грамм, совпадающих с эталоном. Вычисление BLEU не является коммутативной операцией. Здесь интерес представляет случай, где оригинальный текст выступает в качестве эталона. Также вычислялась метрика BLEU_AVG – средние BLEU, вычисленные в обоих направлениях.

Сложнее всего сравнить осмысленность полученных текстов. Одним из способов является оценка перплексии – метрики, применяемой при сравнении языковых моделей. Если взять хорошую языковую модель, то высокая перплексия на заданном утверждении будет означать, что модель предскажет его с низкой вероятностью, т.е. такое утверждение не является естественным для русского языка с точки зрения модели. Мы использовали 2 языковых модели: ai-forever/ruT5-base и ai-forever/ruGPT3small_based_on_gpt2 от Сбера. Для каждого корпуса вычислялась средняя перплексия по каждой из моделей. Итоговая метрика Pplx – это среднее двух полученных оценок.

Результаты для расширений корпуса ArgNet+ArgMicro приведены в табл. 1.

Таблица 1

Результаты оценки качества перефразированных текстов

Метод Метрика	RUT5 Paraphraser	GPT2 Paraphraser	Back Translation
MeaningSim	0.914	0.5394	0.89144
MeaningRand	0.31317	0.3149	0.31538
BLEU	0.53778	0.20217	0.49598
BLEU_AVG	0.53753	0.20123	0.49737
Pplx_O		3.97978	
Pplx	4.3715	4.42185	4.02729

Из показателей MeaningSim можно видеть, что методы на основе RUT5 и обратного перевода не отходили слишком далеко по смыслу от оригинала. Этого нельзя сказать о модели на основе GPT2. Она обучалась на параллельном корпусе новостных заголовков Яндекса, вследствие чего генерируемые парафразы часто похожи на такие заголовки, пример чего показан в таблице 2. Очевидно, что вариант, представленный моделью на основе GPT2, значительно отличается по смыслу от оригинала и двух других парафраз и напоминает заголовок из новостной ленты.

Таблица 2

Пример перефразирования предложения

Оригинал	<i>Тем не менее, каждый должен участвовать в финансировании государственных теле- и радиокompаний в равной мере.</i>
RUT5 Paraphraser	<i>Однако все должны принять участие в финансировании гостелерадиокompаний в той или иной степени, в которой они могут быть финансированы.</i>
GPT2 Paraphraser	<i>Путин считает участие всех россиян в госфинансировании равнозначным.</i>
Back Translation	<i>Несомненно, все должны вносить равный вклад в финансирование государственных телерадиокompаний.</i>

Низкие MeaningRand показывают, что корпус обладает смысловым разнообразием. Из значений BLEU можно заключить, что новые тексты примерно наполовину похожи на оригиналы. Другими словами, они отличаются от них достаточно, чтобы их можно было считать новыми и, в то же время, сходны по стилистике и языку. Данный факт является положительным моментом в контексте задач обнаружения аргументации, т.к. перефразированные утверждения с большой долей вероятности несут ту же мысль, что и оригинал, немного отличаясь по форме. Pplx_O и Pplx – это оценки осмысленности оригинальных и новых текстов соответственно.

Результаты показывают, что новые данные получились в целом достаточно осмысленными. Pplx новых текстов находится примерно на одном уровне и не имеет значительного отличия от Pplx_0.

Тем не менее, парафразы, полученные методом на основе GPT2, не были включены в расширенный корпус. Исследование аргументации подразумевает поиск определенных связей между утверждениями. Расширяя имеющийся набор данных автоматически переписанными утверждениями, необходимо придерживаться общего смысла, заданного в оригинале, так как при значительных изменениях сохранение аргументативных связей не гарантировано. Модель GPT2 в среднем генерирует утверждения, более далекие по смыслу от оригинала, по сравнению с другими двумя моделями. Этим же объясняются и относительно низкие BLEU. Также, среди текстов, полученных с помощью GPT2, чаще встречаются малоосмысленные примеры. Лучшие показатели продемонстрировал метод обратного перевода. Чуть хуже, но примерно на том же уровне, результаты у модели на основе RUT5. Тексты, полученные этими двумя моделями, были включены в расширенный корпус. Таким образом, удалось добиться увеличения первоначального набора данных в три раза.

2. Генерация датасета на основе индикаторов аргументации

При генерации данных для задачи извлечения аргументативных отношений необходимо выделять в тексте утверждения, которые могут являться компонентами аргументов. В дальнейшем для пар таких утверждений решается задача классификации, т.е. принимается решение о наличии или отсутствии аргументативной связи, ее направленности и типе. Обычно в качестве утверждений рассматриваются предложения, а на вход классификатору подаются пары соседних предложений (т.е. берется окно в 2 предложения).

Однако, в представлении опытных аннотаторов, это является существенным упрощением, так как при сегментации возможна разная степень детализации, за которой стоит различное видение структурированности аргументации. В общем случае выделение аргументативных дискурсивных единиц (argumentative discourse units, ADU) строится на базе элементарных дискурсивных единиц, однако, как показано в [Kononenko et al., 2020], между ними нет взаимно однозначного соответствия.

В данной работе для уточнения границ утверждений предлагается использовать индикаторы аргументации. Индикаторы аргументации – это используемые в дискурсе языковые средства (слова и конструкции), указывающие на наличие аргументации. При ручной разметке текста индикаторы служат подсказками для определения структуры аргументации: они помогают установить наличие аргументов и их составляющих, выделить в

тексте границы утверждений, реконструировать связи между утверждениями, соотнести аргумент с определенной схемой рассуждения (формой вывода, выражающей взаимосвязь посылок и заключений).

При автоматическом анализе текста и выявлении аргументации индикаторы также играют существенную роль, что показано, например, в исследовании [Van Eemeren et al. 2007], где демонстрируется возможность рассмотрения широкого класса языковых выражений, сигнализирующих о применении той или иной схемы аргументации.

Для экспериментального исследования этого вопроса на основе современных нейросетевых подходов необходимо создать данные, явным образом включающие информацию об индикаторах.

2.1. Применение индикаторов для построения датасетов

В экспериментах по извлечению аргументативных отношений индикаторы играют роль значимых признаков. Мы использовали два подхода к подготовке данных для обучения ML-моделей.

В первом подходе нейронной сети предлагается самостоятельно обучиться выделять индикаторы, для этого в тексты вносятся специальные пунктуационные символы, или токены, маркирующие конструкции, найденные с помощью шаблонов. Такой способ выделения именованных сущностей был успешно применен для детектирования аргументативных предложений в работе [Alibaeva et al., 2022]. В нашем случае использовалась простая маркировка, меняющая входной текст на *_индикатор_*.

... **Автор статьи говорит о** том, что ...

Другие варианты маркировки включают указания типа или аспекта, что требует проведения более строгой классификации индикаторов, например, в соответствии со схемой аргументации. На текущий момент такое разграничение провести не удалось.

Второй подход основан на свойстве индикаторов служить коннекторами в общей структуре дискурса, тем самым фиксируя границы утверждений, играющих роль посылок, заключений или контраргументов в структуре аргументации.

<³...[*Более того*, <¹ *относя подобную лексику в большей мере к литературному языку*]^A, [А.П. Черкасова объясняет <² *их популярность даже за пределами Магриба*]^B >¹, [например, *djellaba – джеллаба (название одежды), balgha – бэльга (традиционная обувь)*]^C >² >³ ...

В данном примере выделенные индикаторы маркируют границы утверждений, входящих в структуру трех аргументов. Аргумент <1> выделяется индикатором *А.П. Черкасова объясняет* и обнаруживается в пределах содержащей его клаузы, представляющей посылку В об утверждении экспертом

некоторого знания и заключение (подчеркнуто), и клаузы левее индикатора, представляющей вторую посылку А – об области экспертизы. Аргумент <2> выделяется индикатором *например*, который включает посылку С в виде группы примеров справа от индикатора и заключение слева. Таким образом, в данном тексте заключение поддержано двумя посылками – экспертным мнением и примерами. Аргумент <3> опирается на индикатор *более того*, который указывает на наличие где-то в левом контексте (в данном тексте, в пределах абзаца) утверждения-заключения, поддержанного некоторой посылкой и параллельно посылкой, представленной заключением, содержащемся в правом контексте индикатора. Заметим, что данный текст содержит индикаторы, которые в большинстве случаев однозначно детектируют позиции и роли посылок и заключений. Однако, существует немало индикаторов, для которых распознавание не столь однозначно.

Таким образом, второй способ подготовки данных включает следующие аспекты:

- использование индикаторов для выделения утверждений внутри предложений;
- фиксацию контекстов индикаторов, потенциально способных быть связанными аргументативными отношениями;
- фиксацию направленности аргументативной связи.

Для каждого вхождения индикатора контекст его использования разделяется на три части: основное утверждение, включающее индикатор, а также левый и правый контексты. Фиксируя такие потенциально аргументативные высказывания в соответствии с расположением относительно индикатора, мы не просто выделяем аргументативные предложения, а конкретно посылки “в поддержку” и соответствующие заключения, либо атакующие и атакуемые утверждения, по сути представляющие посылки “против” и соответствующие заключения, что позволяет говорить о наличии или отсутствии аргументативной связи.

Так, для индикаторов в приведенном выше примере будут сформированы следующие строки таблицы:

- (1) pattern = эксперт_мнение | main = B | left_context = A | right_context = C | arg_left = 1 | arg_main = 2 | arg_right = 1 | rel_left = 1 | rel_right = 1
- (2) pattern = например | main = C | left_context = A+B | right_context = “...” | arg_left = 3 | arg_main = 1 | arg_right = 0 | rel_left = 2 | rel_right = 0
- (3) pattern = более_того | main = A | left_context = “...” | right_context = B+C | arg_left = 3 | arg_main = 1 | arg_right = 3 | rel_left = 0 | rel_right = 2.

В первой строке автоматически фиксируются две аргументативные связи (поля `rel_left` и `rel_right`), однако для второй связи (рассуждение от примера) маркер экспертного мнения не является индикатором. Вторая связь также фиксируется в наборе для маркера *например*, что абсолютно корректно. Первая связь (рассуждение от эксперта) повторно фиксируется

набором для маркера *более того*, который в данном случае является индикатором дополнительной аргументации некоторого тезиса, представленного в контексте слева.

2.2. Статистическая оценка индикаторов на корпусе

Статистическое исследование индикаторов проводилось на размеченном экспертами корпусе ArgNet, в котором из 6611 выделенных аргументов 2624 включали индикаторы (т.е. около 39,7%).

Текущий словарь индикаторов включает около 260 шаблонов, однако в корпусе встретилось всего 144 индикатора, из них больше 10 раз только 58 индикаторов. Для каждого индикатора были подсчитаны следующие параметры:

- **count** – количество вхождений в корпус;
- **qm** – доля аргументативных утверждений;
- **q_rel** – доля аргументативных связей.

Данные характеристики демонстрируют степень информативности маркеров, при этом показатель *qm* можно соотнести с задачей выявления аргументации (аргументативных утверждений), а *q_rel* – непосредственно с задачей выявления аргументативных отношений.

В табл. 3 приведены 20 наиболее информативных маркеров по показателю *qm* среди маркеров с частотой больше 10.

Таблица 3

Характеристики индикаторов аргументации

pattern	count	qm	q_rel	pattern	count	qm	q_rel
<i>предположим что</i>	11	0,91	0,55	<i>с другой стороны</i>	19	0,74	0,37
<i>следовательно</i>	16	0,88	0,38	<i>свидетельствовать о</i>	14	0,71	0,36
<i>с одной стороны</i>	14	0,86	0,57	<i>не только</i>	108	0,71	0,41
<i>напротив</i>	12	0,83	0,83	<i>так как</i>	57	0,70	0,35
<i>в частности</i>	40	0,83	0,48	<i>результаты исследования</i>	10	0,70	0,40
<i>так зпт</i>	90	0,80	0,56	<i>X приводит к Y</i>	13	0,69	0,38
<i>в связи с</i>	15	0,80	0,67	<i>поскольку</i>	60	0,68	0,47
<i>в таблице</i>	14	0,79	0,21	<i>например</i>	410	0,68	0,43
<i>на рисунке</i>	22	0,77	0,27	<i>с целью</i>	15	0,67	0,40
<i>таким образом</i>	77	0,74	0,51	<i>X говорит о Y</i>	18	0,67	0,39

Из приведенных данных видно, что индикаторы, достаточно хорошо выявляющие аргументацию, совсем необязательно будут хорошо идентифицировать аргументативное отношение ($qm \geq q_rel$). Это вполне объяснимо для таких индикаторов как *в таблице*, *на рисунке* и достаточно неожиданно для индикаторов *следовательно* или *в частности*.

Стоит отметить, что низкочастотные индикаторы нередко более специфичны и лучше выявляют как наличие аргументации, так и наличие аргументативных связей. Однако пока, на небольших объемах размеченных данных, это статистически подтвердить невозможно.

Заключение

Задача автоматического извлечения аргументации предполагает решение нескольких подзадач, таких как выявление фрагментов текста, содержащих аргументацию, сегментация найденных фрагментов на аргументативные дискурсивные единицы, установление наличия и вида связи между парами таких единиц и определение их ролей в составе аргументов (посылка за/против, заключение). О сложности извлечения аргументативных отношений свидетельствует тот факт, что, как указано в обзоре [Lawrence et al., 2019], лучшие современные решатели показывают для этой задачи невысокое качество.

Недостаточное качество работы систем извлечения аргументативных связей обусловлено рядом общих и частных проблем, которые в отношении подготовки датасетов в текущем исследовании можно охарактеризовать следующим образом:

- объем размеченных данных все еще недостаточен для проведения реальных экспериментов по извлечению аргументации из русскоязычных текстов;
- улучшение качества разметки данных имеет столь же важное значение, как и их объем, в связи с чем необходима тщательная проработка методики разметки, создание которой, в свою очередь, невозможно без накопления достаточного опыта аннотирования.

Базовым компонентом при подготовке данных является сегментация текста. В большинстве корпусов принята сегментация по предложениям. Однако в экспертной разметке границы утверждений не всегда совпадают с границами предложений. Метод сегментации на основе индикаторов имеет свои ограничения:

- не все индикаторы аргументации являются коннекторами (ср. операторы в [Anscombe et al., 1983]);
- многие индикаторы семантически неоднозначны и имеют различные сферы действия;
- существуют ситуации смещения индикатора-коннектора внутрь представляющего пропозицию сегмента;
- в рамках одного предложения возможны несколько индикаторов.

Таким образом, требуется дальнейшее исследование индикаторов аргументации и методов сегментации на их основе.

Список литературы

- [Alibaeva et al., 2022] Alibaeva K., Loukachevitch N.V. Analyzing COVID-related Stance and Arguments using BERT-based Natural Language Inference // Computational Linguistics and Intellectual Technologies: Proc. of the International Conference “Dialog-2022”. – Moscow: RSUH, 2022. – Issue 21.
- [Anscombre et al., 1983] Anscombre J.-C., Ducrot O. L’argumentation dans la langue, Bruxelles. 1983.
- [Fenogenova, 2021] Fenogenova A. Russian paraphrasers: Paraphrase with transformers // Proc. 8th Workshop on Balto-Slavic Natural Language Processing, Ukraine. 2021.
- [Fishcheva et al., 2021] Fishcheva I., Goloviznina V., Kotelnikov E. Traditional Machine Learning and Deep Learning Models for Argumentation Mining in Russian Texts // arXiv preprint arXiv:2106.14438. – 2021.
- [Fishcheva et al., 2019] Fishcheva I., Kotelnikov E. Cross-Lingual Argumentation Mining for Russian Texts // Analysis of Images, Social Networks and Texts: Proc. 8th International Conference AIST 2019. – Springer International Publishing, 2019.
- [Ilina et al., 2021] Ilina D., Kononenko I., Sidorova E. On Developing a Web Resource to Study Argumentation in Popular Science Discourse // Computational Linguistics and Intellectual Technologies: Proc. of the Int. Conference “Dialog-2021”, Russia. 2021.
- [Kononenko et al., 2020] Kononenko I.S., Sidorova E.A., Akhmadeeva I.R. Comparative analysis of rhetorical and argumentative structures in the study of popular science discourse // Computational Linguistics and Intellectual Technologies: Proc. of the International Conference “Dialog-2020”, Russia. 2020.
- [Kotelnikov et al., 2022] Kotelnikov E., Loukachevitch N., Nikishina I., Panchenko A. RuArg-2022: Argument Mining Evaluation // Computational Linguistics and Intellectual Technologies: Proc. of the International Conference “Dialog-2022”, Russia. 2022.
- [Lawrence et al., 2019] Lawrence J., Reed C. Argument Mining: A Survey // Computational Linguistics. – 2019. – 45 (4).
- [Ng et al., 2019] Ng N., Yee K., Baevski A., Ott M., Auli M., Edunov S. Facebook FAIR’s WMT19 news translation task submission // arXiv preprint:1907.06616. – 2019.
- [Papineni et al., 2002] Papineni K., Roukos S., Ward T., Zhu W. Bleu: a method for automatic evaluation of machine translation // Proc. 40th annual meeting of the Association for Computational Linguistics, USA. 2002.
- [Park et al., 2018] Park J., Cardie C. A corpus of eRulemaking user comments for measuring evaluability of arguments // Proc. Eleventh International Conference on Language Resources and Evaluation (LREC 2018), Japan. 2018.
- [Romberg et al., 2021] Romberg J., Conrad S. Citizen Involvement in Urban Planning-How Can Municipalities Be Supported in Evaluating Public Participation Processes for Mobility Transitions // Proc. 8th Workshop on Argument Mining. – 2021.
- [Stab et al., 2014] Stab C., Gurevych I. Identifying Argumentative Discourse Structures in Persuasive Essay // Empirical Methods in Natural Language Processing (EMNLP): Proc. of the Int. Conf., Doha, Qatar. – 2014.
- [Walton et al., 2008] Walton D., Reed C., Macagno F. Argumentation schemes. – Cambridge University Press, 2008.
- [Сидорова и др., 2020] Сидорова Е.А., Ахмадеева И.Р., Загоруйко Ю.А., Серый А.С., Шестаков В.К. Платформа для исследования аргументации в научно-популярном дискурсе // Онтология проектирования. – 2020. – Т. 10, № 4 (38).

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ГЕНЕРАЦИИ ОСМЫСЛЕННОГО СТИХОТВОРНОГО ТЕКСТА

С.Д. Харитонов (*pintastd@gmail.com*)

Д.Г. Буханов (*dgbukh@gmail.com*)

Белгородский государственный технологический университет
им. В.Г. Шухова, Белгород

В работе рассмотрены современные подходы к генерации стихотворного текста. Проведен сравнительный анализ различных языковых моделей, а также предложена и реализована архитектура NLP конвейера для автоматической генерации осмысленного стихотворного текста на базе языковой модели, генетического алгоритма и классификатора смысловой связности текста.

Ключевые слова: обработка естественных языков, генерация стихотворного текста, языковая модель, генетический алгоритм, сверточные нейронные сети.

Введение

Системы, эмулирующие интеллект человека, способны выполнять функции деятелей искусства. Например, программы пишут картины [Liu, 2020], музыку [Tan et al., 2021], стихотворный текст и прозу [Kreps et al., 2021].

Особый интерес представляют задачи создания стихотворных текстов. При их генерации необходимо сочетать различные подходы из широкого спектра наук, а именно методы обработки естественных языков (Natural Language Processing, NLP), технологии искусственного интеллекта и некоторые общие алгоритмы, нацеленные на оптимизацию и повышение качества результата процесса.

Стихотворный текст в отличие от прозы должен обладать не только связностью и корректностью грамматики и семантики предложения (заданного естественным языком), но и следовать заданному ритмическому рисунку, а также содержать рифму.

Одной из наиболее часто используемых технологий для решения NLP задач является технология языковых моделей, которые базируются на архитектуре трансформер [Graterol et al., 2021], [Chernyavskiy et al., 2021]. Эти модели позволяют манипулировать текстом и генерировать текст на основе входной последовательности токенов [Grefenstette, 1999].

Преимуществом данной технологии является общая смысловая связность генерируемого текста и «понимание» языковой моделью семантики языка. Это позволяет в автоматическом режиме расставлять знаки препинания и учитывать все морфологические особенности слов и контекст их употребления.

При обучении таких моделей, а также при их непосредственном использовании, входящий и выходящий текст представляется в виде токенов размер которых зависит от настроечных параметров. Это позволяет более корректно использовать различные формы слов.

Языковые модели на основе архитектуры трансформер поддерживаются большим сообществом разработчиков. Примером таких моделей являются GPT [Floridi et al., 2020] и BERT [Koroteev, 2021].

Недостатком языковых моделей – трансформеров, является необходимость обучения их на большом объеме данных. Например, GPT-3 [Dale, 2021] обучена на 500 миллиардов токенов, а RoBERT на 160ГБ данных [Gillioz et al., 2020]. Стихотворный текст, является результатом исключительно творческого процесса, он менее распространен чем проза, соответственно данных для обучения недостаточно. Вопрос недостаточного объема данных решается путем дообучения [Beheitt et al., 2022] существующих моделей. В начале модель обучается на большом корпусе прозы для «усвоения» знаний о семантике и морфологии языка, а после этого дообучается на меньшем по размеру корпусе стихотворений. В результате чего языковая модель способна генерировать стихотворения. Но их ритмические и фонетические составляющие имеют низкое качество.

Для повышения качества генерируемого стихотворного текста существует подход [Ghazvininejad et al., 2016]. В ходе обучения языковые модели на основе трансформеров разделяют текст на токены в зависимости от частоты встречаемости их в тексте. Для увеличения информации, содержащейся в одном токене, осуществляется деление на токены по слогам. При этом ударные и безударные слоги маркируются и по-разному обрабатываются языковой моделью. Например:

В ле|с`у ро|д`и|лась `е|лоч|ка.

Слоги «су» и «с`у» становятся разными токенами (технически имеют разные числовые индексы), и в ходе обучения модель учится использовать их при генерации по-разному.

Описанный подход решает проблему отсутствия у модели информации о фонетике и ритмике стихотворения, но не решает проблему рифмы в строках. Проблема с рифмой связана с тем, что генерация строки осуществляется слева направо по одному токену, в результате чего снижается количество возможных вариантов слов к концу строфы. Для решения этой проблемы используется обучение модели на основе инвертированных

строк, написанных справа налево. Таким образом уже в начале генерации строфы языковая модель способна подобрать корректное рифмующееся окончание. После генерации стихотворного текста его необходимо инвертировать по токенам. Например:

д`е во | шей ю ст`а ра на | в | И
н`ы би глу | из | ст`он | вный з`ы При
зд`е ве | ч`ит зву | г`ул | ле д`а | А
сн`ы ве | гл`ас | и | взд`ох | вый п`ер | То.

Таким образом, используя различные модификации процесса обучения и использования языковых моделей на основе архитектуры трансформер, достигается значительное увеличение качества генерируемого стихотворного текста [Santillan et al., 2020, De la Rosa et al., 2021]. Но данный подход не позволяет контролировать ритмический рисунок итогового произведения, его изменение и схему рифмовки строчек или частей строф.

Языковые модели на основе архитектуры трансформер позволяют генерировать осмысленный и качественный стихотворный текст, но отсутствие возможности тонкой настройки значительно сужает возможности использования и вынуждает конечного пользователя воспроизводить циклы генерации до необходимого результата заранее неизвестное количество раз.

При генерации стихотворных текстов используют менее тяжеловесные языковые модели на основе рекуррентных нейронных сетей такие как LSTM [Lippi et al., 2019] и GRU [Talafha et al., 2019]. В ходе использования которых возникают аналогичные проблемы. Они являются следствием того, что технологии базируются на том, что входной и выходной сигнал являются последовательностью токенов.

Для решения данной проблемы возможно обучать под каждый набор параметров (схема рифмовки, контекст произведения и стихотворный размер) отдельную языковую модель, однако такая узкая специализация приводит к сокращению объема обучающей выборки, и не решает вопрос задания контекста произведения. При использовании искусственных нейронных сетей контекст необходимо задавать первыми строчками произведения.

Таким образом существующие подходы генерации стихотворного текста имеют следующие недостатки:

- отсутствие настроечных параметров ритмического рисунка, схемы рифмовки и контекста произведения;
- потребность в заранее сочиненном начале стихотворения;
- нехватка данных для обучения.

Для решения описанных ранее проблем предлагается разработка алгоритма выбора слов для продолжения строфы. Помимо формальных параметров, таких как ритмический рисунок и рифма, проверка которых так же

возможна, как и формальным алгоритмом, так и другими различными технологиями обработки естественных языков, алгоритму формирования стихотворного текста необходимо предоставить набор слов уместных на конкретной позиции в предложении, а также подходящих по контексту произведения. Для этого необходимо определить контекст использования слова.

Одним из способов задания контекста использования слов является их векторное представление. При этом расстояние между векторами является показателем сходства контекстов использования двух слов. Существуют различные алгоритмы векторизации и вычисления разницы между их контекстами использования:

- RNNLM [Li et al., 2018];
- Word2vec [Goldberg et al., 2014], [Karani, 2018];
- GloVe [Pennington et al., 2014], [Sharma et al., 2017];
- fastText [Athiwaratkun et al., 2018].

RNNLM (рекуррентная нейросетевая языковая модель) принимает на вход последовательность слов и базируется на алгоритме непрерывного мешка слов [Zhang et al., 2010]. Из мешка выбираются вектора близких по контексту слова (стоящие рядом), после чего модель определяет центральное слово, сжимает вектора и передает их в скрытый слой с softmax функцией активации, на котором решается, какие вектора будут использоваться дальше. Недостатком ее является отсутствие хранения информации о долгосрочных связях слабо связанных по контексту использования слов.

Word2vec является переработкой предыдущей языковой модели учитывающая слова часто используемые в похожих контекстах. Из RNNLM модели был убран скрытый слой и изменена цель с определения центрального слова, на предсказание пропущенного. Для этого при обучении используются скип-граммы, это фрагменты текста с пропущенным словом, на основе, которого языковая модель пытается предсказать пропущенное слово исходя из контекста.

Результат векторизации слов с помощью Word2vec является интерпретируемым, который позволяет вычитать и складывать вектора. Помимо интерпретируемости, данная модель обладает теми же преимуществами, что и RNNLM модель.

GloVe (Global Vectors) модель, как и предыдущая опирается на интерпретируемость векторов слов. Она минимизирует разницу между произведением векторов слов и логарифмом вероятности их совместного появления с помощью стохастического градиентного спуска, в результате чего получается связывать различные понятия, например модели и марки машин, города и страны. Данная модель учитывает то, что одно и то же слово в разных контекстах может обладать различной смысловой нагрузкой, однако обладает теми же недостатками что и предыдущие модели при обработке редких и неизвестных слов.

FastText представляет собой комбинацию Word2Vec и RNNLM модели: для векторизации используются одновременно skip-gram, негативное семплирование и алгоритм непрерывного мешка. В качестве архитектуры выступает feed-forward архитектура с одним скрытым слоем. Основным отличием данной модели является разбиение слов на n-граммы в зависимости от настроечных параметров. Это позволяет модели лучше работать с неизвестными, редкими и устаревшими словами, при этом сохраняя преимущества Word2Vec и RNNLM моделей.

1. Интеллектуальная система генерации стихотворного текста

Разработан конвейер, использующий один из способов векторизации слов для соблюдения контекста произведения и алгоритм выбора слов, подходящих по стихотворному размеру и схеме рифмовке.

Исходя из определения стихотворения выделяются следующие основные параметры, описывающие его:

- стихотворный размер произведения;
- схема рифмовки окончаний стихов (пары рифмующихся строк);
- контекст произведения.

Как контекст произведения принимается набор ассоциаций, возникающих у человека при прочтении произведения, в данном случае он задается набором слов. Например, для эпоса это может быть такой набор слов как: битва, меч, гладиатор.

Каждая языковая модель имеет разную сферу применения и качество предсказания в зависимости от многих параметров. Для осознанного выбора наиболее подходящей языковой модели был проведен их анализ по следующим критериям.

- размер вокабуляра считанного моделью из исходного текста;
- занимаемая моделью память;
- время обучения;
- время выполнения запроса;
- качество итогового стихотворного текста, полученного с применением языковой модели, оцениваемое GPT-3 моделью.

На рис. 1 приведены результаты экспериментов по обучению и использованию языковых моделей, обучение происходило на массиве данных размером 50 тысяч токенов, в качестве токенов использовались слова, размер окна был выбран в 5 токенов, размерность векторов слов 300.

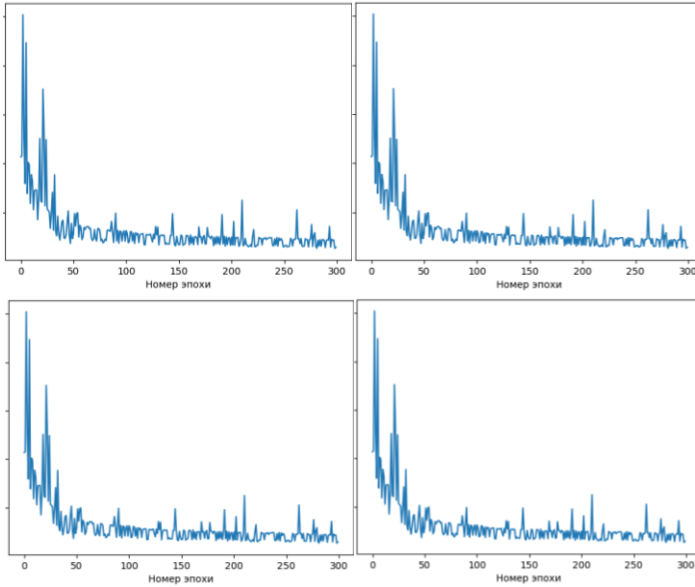


Рис. 1. Графики ошибки слева сверху Word2Vec, справа сверху FastText, слева внизу GloVe, справа внизу RNNLM

Таблица 1

Название	Размер словаря (слов)	Оценка занимаемой памяти (байт)	Время обучения (секунд)	Длительность запроса (микро секунд)	Оценка GPT 3 модели (%)
RNNLM	3780	63376600	3571	242	35%
Word2Vec		10962000	130	1377	41%
GloVe		10964900	31	210	53%
FastText		2411607920	170	25615	45%

Из результатов (табл. 1) экспериментов видно очевидное преимущество Glove модели над аналогичными, соответственно именно она будет использоваться при разработке системы.

В случае, когда алгоритм не может подобрать слово из близких по контексту или близкие по контексту слова плохо вписываются в канву генерируемого стихотворения, система предупреждает пользователя и выбирает случайное слово.

При эксплуатации системы на основе алгоритма выбора слов из близких по контексту была выявлена значительная проблема в связности стихотворения и логичности повествования. Для ее решения в конвейер были добавлены два элемента: пред обученный классификатор определения наличия смысловой связности стихотворного произведения и генетический алгоритм использующий его в качестве фитнес функции оценки экземпляров популяции нацеленный на максимизацию критерия количества смысловой связности, путем замены слов на более удаленные по контексту использования (мутации) и слияния двух стихотворения (скрещивание).

Все описанные компоненты были собраны в NLP конвейер, преобразующий входные параметры, описанные ранее в стихотворный текст им удовлетворяющий. Структура конвейера описана на рис. 2.

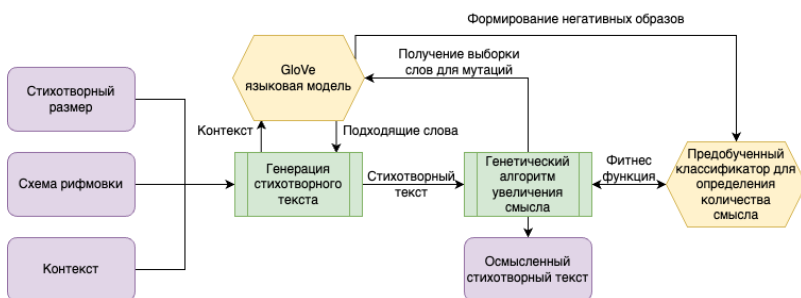


Рис. 2. Структура системы генерации стихотворного текста

Ключевым элементом является пред обученная модель для подбора слов, она используется на всех трех этапах конвейера, ввиду этого выбору подходящей модели векторизации слов было уделено особое внимание. Работа конвейера разделяется на две части. На первом этапе в соответствии входным параметрам формируется стихотворный текст из слов, рекомендованных языковой моделью и выбранных разработанным алгоритмом. Затем он поступает на вход генетического алгоритма увеличения смысловой связности, который превращает стихотворный текст в более осмысленный и связный.

2. Результаты работы

После реализации всего NLP конвейера для генерации стихотворного текста было проведено тестирование для оценки полученных результатов. Результаты оценивались на соответствие итогового произведения входным параметрам. Для каждого параметра были определены три наиболее часто встречаемых типа, а затем для каждого типа проведена генерация и оценка тысячи четверостиший.

При оценке точности соблюдения ритмического рисунка были выбраны такие стихотворные размеры как ямб, хорей и амфибрахий. В качестве контекста использовались слова, имеющие романтический контекст, а схема рифмовки определялась как рифма окончаний строк через одну. Результаты представлены в табл. 2.

Таблица 2

	Ямб	Хорей	Амфибрахий
Несоответствие стихотворному размеру	53	34	46
Несоответствие схеме рифмовки	26	41	31
Стихотворения без ошибок	921	925	923

Для вычисления качества формирования стихотворного текста была использована зависимость между полным количеством экспериментов и количества корректных генераций.

Из представленных результатов видно, что разработанная система в 92,3% случаев формирует корректные с точки зрения ритмики стихотворения:

- ямб – 92,1%;
- хорей – 92,5%;
- амфибрахий – 92,3%.

При оценке точности соблюдения схемы рифмовки были выбраны схемы рифмования «через строчку», «попарно», «попарно-через строчку». В качестве контекста использовались слова, имеющие романтический контекст, а в качестве стихотворного размера был выбран хорей, ввиду наилучших результатов в прошлом эксперименте. Результаты представлены в табл. 3.

Таблица 3

	Через строчку	Попарно	Через строчку
Нарушение рифмы в одной паре строк	63	45	53
Нарушение рифмы в двух парах строк	21	37	86
Стихотворения без ошибок по рифмовке	916	918	861

Из представленных результатов видно, что разработанная система в 89,8% случаев формирует корректные с точки зрения рифмы стихотворения:

- через строчку – 91,6%;
- попарно – 91,8%;
- попарно-через строчку – 86,1%.

При оценке сгенерированного стихотворного текста на соответствие контексту и смысловую связность были выбраны следующие контексты «лирический», «эпический» и «романтический». В качестве стихотворного размера был выбран хорей, а схема рифмовки «попарно», ввиду наилучших результатов в предыдущих экспериментах. Результаты представлены в табл. 4.

Таблица 4

	Лирический	Эпический	Романтический
Нарушен контекст	31	26	41
Стихотворение не имеет смысловой связности	312	273	276
Стихотворения без ошибок в контексте и смысле	657	701	683

Из представленных результатов видно, что разработанная система в 68% случаев формирует корректные с точки зрения контекста и смысловой связности стихотворения:

- лирический – 65,7%;
- эпический – 70,1%;
- романтический – 68,3%.

Пример сгенерированного стихотворного текста:

Не приеду, но сердцем приду
Там искусство встречено ночью
Где влюбился туда и иду
Я приду к тебе тихой ночью

Исходя из оценки результатов работы алгоритма следует, что разработанная система формирования стихотворных текстов на высоком уровне справляется с формальными частями стихотворения, такими как стихотворный размер и рифмовка, однако качество стихотворения в плане смысловой связности и соответствия контексту не так высоко. Из чего следует необходимость продолжать работу над генетическим алгоритмом повышения смысловой связности, а также в целом совершенствовать построенный NLP конвейер и его части.

Заключение

В работе рассмотрены различные современные подходы к генерации стихотворного текста, а также изучены и проанализированы различные способы векторизации слов, из которых в дальнейшем по средствам сравнительного анализа был выбран наиболее подходящий для использования

в разработанной системе формирования стихотворного текста. Предложена архитектура NLP конвейера для автоматической генерации стихотворного текста на основе заданных параметров таких как: стихотворный размер, схема рифмовки и контекст произведения.

Для реализации NLP конвейера был разработан алгоритм выбора из слов, подходящих по контексту максимально уместного слова на конкретное место в стихотворной строке. В дальнейшем ввиду низкой смысловой наполненности и связности генерируемого текста был предложен генетический алгоритм, нацеленный на повышение смысловой связности использующий в качестве фитнес функции пред обученный классификатор на основе сверточных нейронных сетей классифицирующий фрагменты текста на два класса: связанные смыслом и бессмыслицу.

В результате, разработанная система формирования осмысленного стихотворного текста способная удовлетворять заданные критерии со следующей точностью: ритмический рисунок – 92.3%, схема рифмовки – 89.8%, контекст – 68%.

Дальнейшим направлением развития исследования является совершенствование алгоритма выбора слов, а также разработка и реализация более эффективных схем скрещивания и мутаций в рамках генетического алгоритма. Разработанная структура NLP конвейера позволяет совершенствовать части системы изолированно, не влияя на другие. Помимо дальнейшего усовершенствования системы критически важно продолжать накапливать объемы текста для обучения интеллектуальных элементов системы.

Список литературы

- [Athiwaratkun et al., 2018] Athiwaratkun B., Wilson A. G., Anandkumar A. Probabilistic fasttext for multi-sense word embeddings //arXiv preprint arXiv:1806.02901. – 2018.
- [Beheitt et al., 2022] Beheitt M.E.G., Hmida M.B.H. Automatic Arabic Poem Generation with GPT-2 // ICAART (2). – 2022.
- [Chernyavskiy et al., 2021] Chernyavskiy A., Ilvovsky D., Nakov P. Transformers: “the end of history” for natural language processing? //Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part III 21. – Springer International Publishing, 2021.
- [Dale, 2021] Dale R. GPT-3: What’s it good for? // Natural Language Engineering. – 2021. – Vol. 27, No. 1.
- [De la Rosa et al., 2021] De la Rosa J. et al. Transformers analyzing poetry: multilingual metrical pattern prediction with transformer-based language models // Neural Computing and Applications. – 2021.
- [Floridi et al., 2020] Floridi L., Chiriatti M. GPT-3: Its nature, scope, limits, and consequences // Minds and Machines. – 2020. – Vol. 30.

- [**Ghazvininejad et al., 2016**] Ghazvininejad M. et al. Generating topical poetry // Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. – 2016.
- [**Gillioz et al., 2020**] Gillioz A. et al. Overview of the Transformer-based Models for NLP Tasks // 2020 15th Conference on Computer Science and Information Systems (FedCSIS). – IEEE, 2020.
- [**Goldberg et al., 2014**] Goldberg Y., Levy O. word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method // arXiv preprint arXiv:1402.3722. – 2014.
- [**Graterol et al., 2021**] Graterol W. et al. Emotion detection for social robots based on NLP transformers and an emotion ontology // Sensors. – 2021. – Vol. 21, No. 4.
- [**Grefenstette, 1999**] Grefenstette G. Tokenization // Syntactic Wordclass Tagging. – 1999
- [**Karani, 2018**] Karani D. Introduction to word embedding and word2vec // Data Sci. – 2018. – Vol. 1.
- [**Koroteev, 2021**] Koroteev M.V. BERT: a review of applications in natural language processing and understanding // arXiv preprint arXiv:2103.11943. – 2021.
- [**Kreps et al., 2021**] Kreps S., McCain R.M., Brundage M. All the news that's fit to fabricate: AI-generated text as a tool of media misinformation // Journal of experimental political science. – 2022. – Vol. 9, No. 1.
- [**Li et al., 2018**] Li S., Xu J. A recurrent neural network language model based on word embedding // Web and Big Data: APWeb-WAIM 2018 International Workshops: MWDA, BAH, KGMA, DMMOOC, DS, Macau, China, July 23–25, 2018, Revised Selected Papers 2. – Springer International Publishing, 2018.
- [**Lippi et al., 2019**] Lippi M. et al. Natural language statistical features of LSTM-generated texts // IEEE Transactions on Neural Networks and Learning Systems. – 2019. – Vol. 30, No. 11.
- [**Liu, 2020**] Liu X. Artistic reflection on artificial intelligence digital painting // Journal of Physics: Conference Series. – IOP Publishing, 2020. – Vol. 1648, No. 3.
- [**Pennington et al., 2014**] Pennington J., Socher R., Manning C. D. Glove: Global vectors for word representation // Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). – 2014.
- [**Santillan et al., 2020**] Santillan M.C., Azcarraga A.P. Poem generation using transformers and Doc2Vec embeddings // 2020 International Joint Conference on Neural Networks (IJCNN). – IEEE, 2020.
- [**Sharma et al., 2017**] Sharma Y. et al. Vector representation of words for sentiment analysis using GloVe // 2017 international conference on intelligent communication and computational techniques (icct). – IEEE, 2017.
- [**Talafha et al., 2019**] Talafha S., Rekabdar B. Arabic poem generation with hierarchical recurrent attentional network // 2019 IEEE 13th International Conference on Semantic Computing (ICSC). – IEEE, 2019.
- [**Tan et al., 2021**] Tan X., Li X. A tutorial on AI music composition // Proceedings of the 29th ACM international conference on multimedia. – 2021.
- [**Zhang et al., 2010**] Zhang Y., Jin R., Zhou Z. H. Understanding bag-of-words model: a statistical framework // International journal of machine learning and cybernetics. – 2010. – Vol. 1.

Секция 3

**ИНТЕЛЛЕКТУАЛЬНЫЕ АГЕНТЫ,
РОБОТЫ, ИНТЕЛЛЕКТУАЛЬНОЕ
УПРАВЛЕНИЕ,
КОМПЬЮТЕРНОЕ ЗРЕНИЕ**

УДК 004.89

**ОПТИМИЗАЦИЯ ТЕМПЕРАТУРНОГО РЕЖИМА
ОБЖИГА ФОСФОРИТОВЫХ ОКАТЫШЕЙ
С ИСПОЛЬЗОВАНИЕМ КОЭВОЛЮЦИОННОГО
АЛГОРИТМА ПЧЕЛИНЫХ КОЛОНИЙ***

В.И. Бобков (*vovabobkoff@mail.ru*)

О.В. Булыгина (*baguzova_ov@mail.ru*)

Филиал Национального исследовательского университета «МЭИ»
в г. Смоленске, Смоленск

В статье рассматривается научно-практическая проблема повышения энергоресурсоэффективности химико-технологических процессов производства фосфоритовых окатышей из отходов апатит-нефелиновых руд. Перспективным путем ее решения является поиск оптимального температурного режима нагрева шихты в обжиговой конвейерной машине. Специфика поставленной задачи оптимизации (нелинейность целевой функции, большая размерность поискового пространства, высокая вычислительная сложность) обуславливает целесообразность использования стохастических методов, в частности алгоритма пчелиных колоний. Для определения значений «свободных» параметров, существенно влияющих на сходимость алго-

* Работа выполнена в рамках государственного задания, проект № FSWF-2023-0012.

ритма, было предложено воспользоваться идеей коэволюции, заключающейся в параллельном запуске нескольких ABC-алгоритмов с различными «настройками» для каждой субпопуляции.

Ключевые слова: энергоресурсоэффективность, фосфоритовые окатыши, обжиговая конвейерная машина, температурный режим, коэволюция, алгоритм пчелиных колоний, нечеткая логика.

Введение

Одной из ключевых проблем развития отечественной горно-химической отрасли является низкая энергоресурсоэффективность технологических процессов. В современных условиях постоянного роста цен на энергоносители данная проблема приобретает особую значимость уже не только с точки зрения рационального использования природных ресурсов, но и с позиции экономической эффективности [Кудрин и др., 2022]. Особенно остро она стоит перед энергозатратными производствами, связанными с высокотемпературной обработкой горно-химического сырья (ярким примером является производство фосфоритовых окатышей из отходов апатит-нефелиновых руд, хранящихся в отвалах горно-обогажительных комбинатов, в обжиговой конвейерной машине).

Решением указанной проблемы может стать снижение объемов потребления энергоресурсов в химико-технологических процессах производства фосфоритовых окатышей [Пучков и др., 2022]. Перспективным способом его достижения является оптимизация температурных режимов нагрева шихты на основе результатов моделирования тепло- и массообменных процессов, протекающих в различных зонах обжиговой конвейерной машины.

Невозможность проведения натурного эксперимента приводит к необходимости использования методов компьютерного моделирования химико-технологических процессов высокотемпературного обжига фосфоритовых окатышей. Однако нелинейность целевой функции, большая размерность поискового пространства, высокая вычислительная сложность затрудняют применение традиционных методов детерминированного поиска. В этих условиях хорошие результаты показывают стохастические методы, преднамеренно вводящие элемент случайности в поисковый алгоритм. Сегодня широкое распространение получили популяционные алгоритмы, которые основаны на моделировании коллективного поведения различных живых организмов (муравьиной колонии, пчелиного роя, волчьей стаи и т.п.) и отличаются возможностью одновременной обработки нескольких вариантов решения поставленной задачи.

Таким образом, возникает актуальная научно-практическая задача разработки информационного обеспечения поддержки принятия решений по повышению энергоресурсоэффективности химико-технологических процессов производства фосфоритовых окатышей на основе оптимизации температурных режимов работы обжиговой конвейерной машины с использованием популяционных алгоритмов.

1. Постановка задачи оптимизации температурных режимов работы обжиговой конвейерной машины

Химико-энерготехнологическая система переработки отходов апатит-нефелиновых руд, хранящихся в отвалах горно-обогатительных комбинатов, состоит из последовательности ресурсоемких процессов, потребляющих большие объемы тепловой и электрической энергии.

На первой стадии производится подготовка минерального сырья для термической обработки: сначала выполняется измельчение отходов апатит-нефелиновых руд, а затем окомкование полученной однородной мелкодисперсной массы в сырые окатыши. Выходными параметрами этой стадии являются радиус r_o и влажность u_o сырых окатышей.

На второй стадии выполняется упрочение сырых окатышей в обжиговой конвейерной машине. Данный химико-технологический процесс происходит следующим образом [Bobkov et al., 2022]. На движущуюся колосниковую решетку (тележку) сначала укладывается слой готовых окатышей («постель»), поверх которого насыпаются сырые окатыши с радиусом r_o (высота укладки сырых окатышей h варьируется в диапазоне 0,25-0,35 м в зависимости от типа машины). Двигаясь на конвейере длиной l м со скоростью V м/с, многослойная масса окатышей последовательно проходит отдельные ($i=1,2,\dots,k$) вакуум-камеры длиной Δl , в которых осуществляется подача газа-теплоносителя с определёнными температурой T_{gi} и скоростью W_{gi} . Критерием завершенности химико-технологического процесса является минимум среднего по слою влагосодержания окатышей, который определяет их прочность.

На рис. 1 показана схема пребывания фосфоритовых окатышей в обжиговой конвейерной машине в процессе прохождения отдельных вакуум-камер, для каждой из которых формируется вектор (T_{gi}, W_{gi}) .

Себестоимость производства фосфоритовых окатышей в обжиговой конвейерной машине в значительной степени определяется энергозатратами. Оптимальным считается такой температурный режим, при котором требуемое качество обожженных окатышей достигается минимумом израсходованной тепловой и электрической энергии. Таким образом, задача оптимизации сводится к поиску минимальных энергозатрат при заданных характеристиках сырья и оборудования, т.е. $F(r_o, u_o, V, T_g, W_g) \rightarrow \min$.

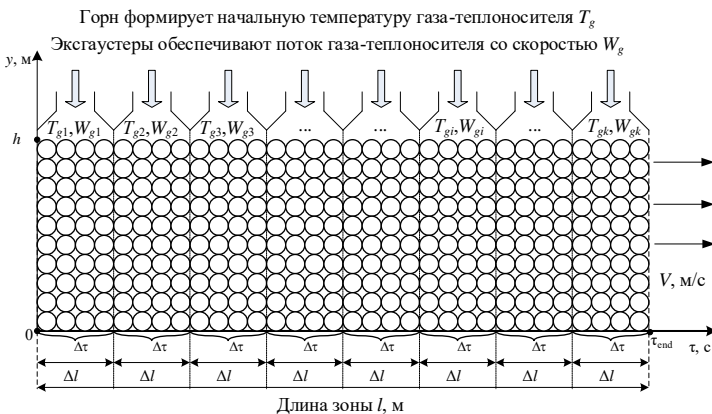


Рис. 1. Схема пребывания фосфоритовых окатышей в обжиговой конвейерной машине

Другими словами, для каждой вакуум-камеры обжиговой конвейерной машины необходимо найти такие значения (T_g, W_g) , которые обеспечат минимальные совокупные затраты на тепловую и электрическую энергию.

2. Использование коэволюционного алгоритма пчелиных колоний для решения поставленной задачи оптимизации

На сегодняшний день разработано множество алгоритмов, вдохновленных коллективным поведением живых организмов, которые могут успешно применяются для решения задач локальной и глобальной оптимизации [Xing et al., 2014]. Такие поисковые алгоритмы используют популяцию, представляющую собой множество взаимодействующих искусственных агентов, которые имитируют естественное поведение (например, по удовлетворению пищевых потребностей) [Карпенко, 2017].

Однако практическое применение популяционных алгоритмов встречается с определенными трудностями, связанными с выбором «свободных» параметров (авторами не были разработаны рекомендации по выбору их оптимальных значений), от которых в значительной степени зависит сходимость [Семенкин и др., 2010]. Простейшим способом их определения является многократное решение задачи, однако это неприемлемо для сложных задач многомерной оптимизации. В качестве альтернативы можно использоваться идея коэволюции, заключающаяся в параллельном запуске нескольких взаимодействующих субпопуляций, для каждой из которых применяются различные «настройки» [El-Abd, 2005]. Конкуренция и кооперация субпопуляций обеспечивает самонастройку алгоритма

под исследуемую задачу в ходе ее однократного решения, т.е. снимается проблема ручного подбора наилучших значений «свободных» параметров [Sun et al., 2010].

Для рассматриваемой задачи оптимизации в качестве базового алгоритма был выбран алгоритм пчелиных колоний (АВС), который моделирует поведение медоносных пчел в процессе поиска источников нектара (подробно описан в работах Д. Карабога [Karaboga, 2005] и Д. Фама [Pham et al., 2005]). Одной из отличительных характеристик данного алгоритма является наличие «коммуникационного» аспекта.

В коэволюционном АВС-алгоритме параллельно выполняется несколько канонических алгоритмов с разными значениями «свободных» параметров, т.е. одновременно эволюционируют несколько субпопуляций, которые борются за «общий» ресурс. Перераспределение ресурса происходит путем сокращения ресурса каждой проигравшей субпопуляции на значение штрафа и увеличения ресурса победившей субпопуляции на сумму потерь проигравших (таким образом, размер всей популяции остается неизменным). Сокращение ресурса проигравших субпопуляций происходит до минимально установленного значения.

Для определения размера штрафа разработано несколько подходов [Семенкин и др., 2011]. В рассматриваемой задаче расчет значения штрафа для каждой субпопуляции на каждой итерации выполнялся с помощью алгоритма нечетко-логического вывода Такаги-Сугено (его выбор обусловлен необходимостью анализа и статистической, и нечеткой информации). В качестве входов использовались удаленность итерации и доля попаданий лучшего представителя субпопуляции в допустимую или недопустимую область на каждом шаге [Borisov et al., 2023].

На рис. 2 показана блок-схема коэволюционного АВС-алгоритма. Тестирование его эффективности проводилось на примере оптимизации технологических параметров обжиговой конвейерной машины ОК-520/536Ф. В таблице 1 приведены показатели повышения энергоэффективности химико-технологических процессов обжига и качества готовых фосфоритовых окатышей, полученные с использованием различных алгоритмов.

Таблица 1

Результаты использования различных методов

	Прочность окатышей, кг/ок	Удельный расход энергии, ТУТ/т
Регламентный режим	51,2	1,93
Динамическое программирование	60,3	1,74
Канонический АВС-алгоритм	67,2	1,67
Коэволюционный АВС-алгоритм	71,8	1,61

Вышеописанный коэволюционный ABC-алгоритм с использованием высокоуровневого языка C++ был реализован в виде программного модуля, входящего в состав компьютеризированной системы поддержки принятия решений по оптимальному управлению энергоресурсоэффективностью химико-энерготехнологической системы переработки отходов апатит-нефелиновых руд [Мешалкин и др., 2021].

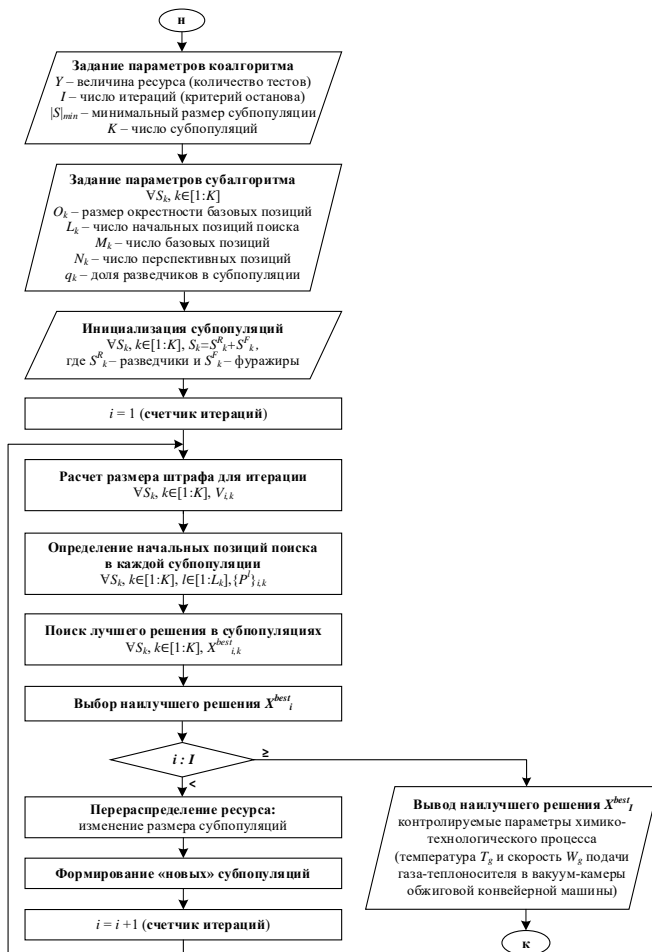


Рис. 2. Блок-схема коэволюционного алгоритма пчелиных колоний для оптимизации температурного режима работы обжиговой конвейерной машины

Заключение

В статье для повышения энергоресурсоэффективности производства фосфоритовых окатышей предложено осуществлять поиск оптимального температурного режима нагрева шихты в обжиговой конвейерной машины с использованием коэволюционного алгоритма пчелиных колоний. Разработанный Со-АВС-алгоритм отличается возможностью комплексного учета множества разнотипных характеристик химико-технологического процесса, а также быстрой настройкой «свободных» параметров алгоритма благодаря параллельному запуску нескольких субалгоритмов, использующих единый ресурс, и методов нечеткой логики для расчета значения штрафа.

Как представляется, управление химико-энерготехнологической системой переработки отходов апатит-нефелиновых с использованием полученных результатов позволит сформировать оптимальный режим работы обжиговой конвейерной машины и повысить качество обожжённых фосфоритовых окатышей.

Список литературы

- [Bobkov et al., 2022] Bobkov V., Bulygina O., Vereikina E. Using population algorithms to optimize the temperature regime of roasting phosphorite pellets // *Prikladnaya informatika = Journal of Applied Informatics*. – 2022. – Vol. 17, No. 6. – P. 94-101.
- [Borisov et al., 2023] Borisov V., Bulygina O., Vereikina E. The use of coevolutionary algorithms for optimizing the operating regimes of the roasting conveyor machine // *Prikladnaya informatika = Journal of Applied Informatics*. – 2023. – Vol. 18, No. 3. – P. 52-60.
- [El-Abd, 2005] El-Abd K.M. A taxonomy of cooperative search algorithms // *Proceedings of the Second international conference on Hybrid Metaheuristics*. – 2005. – Vol. 3636. – P. 32-41.
- [Karaboga, 2005] Karaboga D. An idea based on honey bee swarm for numerical optimization. – Technical report - TR06. Erciyes University, 2005.
- [Pham et al., 2005] Pham D.T., Ghanbarzadeh A., Koc E, Otri S, Rahim S., Zaidi M. The Bees Algorithm. Technical Note. Cardiff University, 2005.
- [Sun et al., 2010] Sun Y., Zhang L., Gu X. Co-evolutionary cultural based particle swarm optimization algorithm // *Life system modeling and intelligent. Communications in computer and information science*. – 2010. – Vol. 98, No. 1. – P. 1-7.
- [Xing et al., 2014] Bo Xing, Wen-Jing Gao. *Innovative computational intelligence: A rough guide to 134 clever algorithms*. – Cham: Springer, 2014. – 451 p.
- [Карпенко, 2017] Карпенко А.П. *Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой*. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2017.
- [Кудрин и др., 2022] Кудрин А.А., Коломиец А.Р. биржевые рынки энергоресурсов в России: защита от шоков или монополий? // *Современная конкуренция*. – 2022. – Т.16. № 2. – С. 34-50.

- [Мешалкин и др., 2021] Мешалкин В.П., Бобков В.И., Длин М.И., Федулов А.С., Шинкевич А.И. Компьютеризированная система принятия решений по оптимальному управлению энергоресурсоэффективностью химико-энерготехнологической системы переработки отходов апатит-нефелиновых руд // Теоретические основы химической технологии. – 2021. – Т. 55, № 1. – С. 67-75.
- [Пучков и др., 2022] Пучков А.Ю., Лобанева Е.И., Култыгин О.П. Алгоритм прогнозирования параметров системы переработки отходов апатит-нефелиновых руд // Прикладная информатика. – 2022. – Т. 17, № 1 (97). – С. 55-68.
- [Семенкин и др., 2010] Семенкин Е.С., Сергиенко Р.Б. Коэволюционный алгоритм для задач условной и многокритериальной оптимизации // Программные продукты и системы. – 2010. – № 4. – С. 23.
- [Семенкин и др., 2011] Семенкин Е.С., Шабалов А.А., Клешков В.М. Система поддержки принятия решений при управлении инновациями реструктурированного машиностроительного предприятия // Сибирский аэрокосмический журнал. – 2011. – № 5. – С. 207-210.

К ВОПРОСУ ОБ ОРИЕНТАЦИИ РОБОТОВ НА ОСНОВЕ ВИЗУАЛЬНЫХ ОРИЕНТИРОВ И КОМПАСА

И.П. Карпова (*karpova_ip@mail.ru*)
Национальный исследовательский университет
«Высшая школа экономики», Москва

В работе рассматривается метод использования визуальных ориентиров для запоминания пройденного пути мобильным роботом (аниматом), основанный на механизме навигации муравьев. Описывается модель представления маршрута и алгоритм преобразования прямого маршрута в обратный. Приведены результаты экспериментов по имитационному моделированию, описаны натурные эксперименты.

Ключевые слова: автономный мобильный робот, навигация муравьев, построение обратного маршрута, задача фуражировки.

Введение

Для мобильных роботов большое значение имеет ориентация в пространстве, запоминание маршрута, способность вернуться в точку отправления и умение повторить маршрут. Среди методов решения этих задач, не требующих построения карты, особое место занимают биологически инспирированные методы ориентации и навигации по аналогии с механизмами ориентации животных, чаще всего – насекомых. В данной работе моделируется поведение муравьев, занимающихся фуражировкой. Этот подход относится к моделированию социального поведения для решения задач групповой робототехники [Карпов, 2016].

Многие исследователи рассматривают принципы навигации муравьев как основу методов навигации и ориентации автономных мобильных роботов. Есть множество вариантов использования отдельных механизмов или морфологических особенностей муравьев: например, роботы, оснащенные специальным компасом для ориентации по поляризации солнечного света ([Lambrinos et al., 2000], [Dureyroux et al., 2019]); или имитация феромонного следа [Dorigo and Blum, 2005]. Для следования роботов по маршруту предлагаются методы на основе запоминания моментальных снимков окружающей среды во время изучения маршрута ([Vardy, 2006],

[Yu et al., 2011]), а также модели векторной навигации на основе искусственной нейронной сети [Goldschmidt et al., 2017], которые используются для запоминания и повторного прохождения агентами маршрута. Большое внимание уделяется организации памяти для навигации ([Augustine et al., 2012], [Mair et al., 2014]). Каждый из упомянутых методов обладает некоторыми недостатками: одни не учитывают визуальные ориентиры, играющие большую роль в ориентации насекомых; другие требуют построения топологической карты; третьи не реализуют возвращение в исходную точку маршрута.

Ранее автором был разработан метод ориентации без построения карты и без использования средств глобальной навигации [Карпова, 2022]. Метод включает механизм запоминания и интерпретации описания маршрута, позволяющий повторно пройти по запомненному маршруту. Недостатки этого метода – отсутствие возможности вернуться в точку отправления и сложности с обходом препятствий. На текущем этапе исследований разработан алгоритм преобразования прямого маршрута в обратный, а перемещение по маршруту включает обход препятствий.

Замечание о терминах. Объект исследования – искусственный автономный агент, действующий в виртуальной или реальной среде и имитирующий поведение живого организма [Wilson, 1987], поэтому в тексте чаще используется термин «анимат». Если речь идет о проведении натуральных экспериментов, то употребляется термин «робот».

1. Алгоритмы ориентации

Разработанный метод опирается на принципы фуражировочного поведения рыжих лесных муравьев (*Formica rufa*). Фуражировка включает поиск ресурса, перенос его в гнездо (возвращение в точку отправления) и повтор этого пути. Во время поиска муравей-разведчик запоминает свое положение относительно солнца, визуальные ориентиры, мимо которых он проходит, и примерное расстояние до них ([Длусский, 1967], [Cruse et al., 2011]). С помощью этой информации разведчик может вернуться в гнездо. Затем он мобилизует фуражиров и ведет их к ресурсу, или передает им сведения о запомненном маршруте, и фуражиры самостоятельно идут за ресурсом [Захаров и др., 2013]. Подробнее особенности поведения муравьев описаны в [Карпова, 2021].

Маршрут может быть представлен как последовательность отрезков, на которых анимат движется прямо. В начале каждого отрезка анимат-разведчик запоминает сцену, выбирает основной ориентир и направление движения относительно этого ориентира. Он передвигается в выбранном направлении, и после обхода ориентира переходит к следующему отрезку пути. Если он не видит ни одного ориентира, то поиск считается неудач-

ным, и анимат возвращается на «базу». Если он видит искомый ресурс, то этот ресурс (объект) является для него основным ориентиром сцены. Общий алгоритм действий анимата-разведчика приведен на рис. 1.

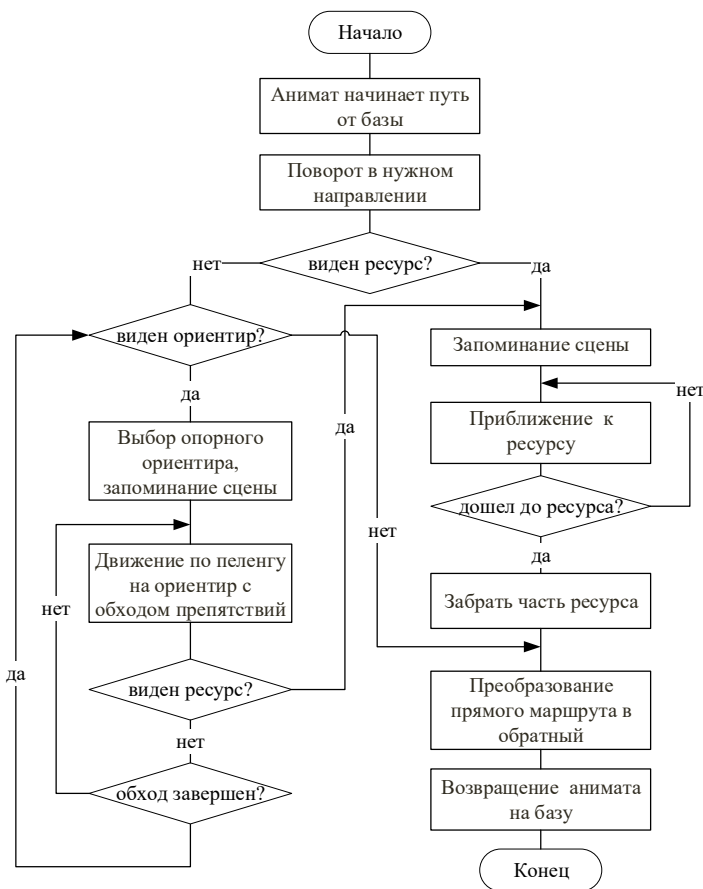


Рис. 1. Общий алгоритм действий анимата-разведчика при поиске ресурса

Далее анимат действует как фуражир и по похожему алгоритму (рис. 2). Он не ищет новый ориентир, а сопоставляет сцены запомненного маршрута с тем, что видит вокруг себя. При успешном сопоставлении сцен фуражир обходит опорный ориентир, используя процедуру движения по пеленгу с обходом препятствий, и переходит к следующей сцене. При неудаче за-

пускает процедуру поиска опорного ориентира в текущей или следующей сцене, аналогично муравью, потерявшему ориентир. Если фуражир не видит ни одного ориентира, он пытается вернуться на «базу».

На «базу» анимат возвращается практически по тому же алгоритму (рис. 2) с двумя отличиями: (1) вместо искомого ресурса анимат ищет «базу» и (2) при потере ориентации анимат продолжает поиск «базы», продвигаясь в том направлении, где предположительно находится «база». Это обеспечивает унификацию поведенческих процедур.

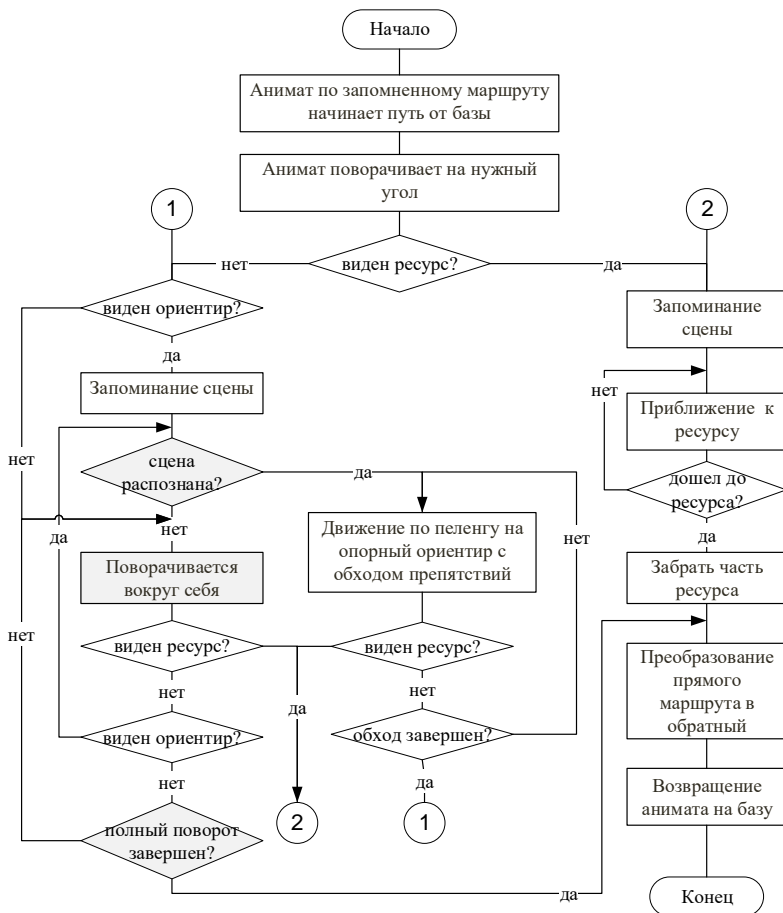


Рис. 2. Общий алгоритм действий анимата-фуражира при повторе пути

2. Мир анимата

Определим, как можно представить мир анимата, и какими возможностями он должен обладать для реализации такого поведения. На полигоне, по которому перемещаются аниматы, расположена «база», неподвижные объекты (ориентиры) и некоторые целевые объекты – «ресурс», который нужно перенести на «базу» (рис. 3). Путь – это последовательность перемещений анимата от одного целевого объекта к другому (целевые объекты – «база» или «ресурс»). Анимат может двигаться вперед и назад, останавливаться, поворачивать направо и налево. С помощью локатора он «видит» окружающие объекты внутри некоторой области видимости (сектор круга на рис. 3).

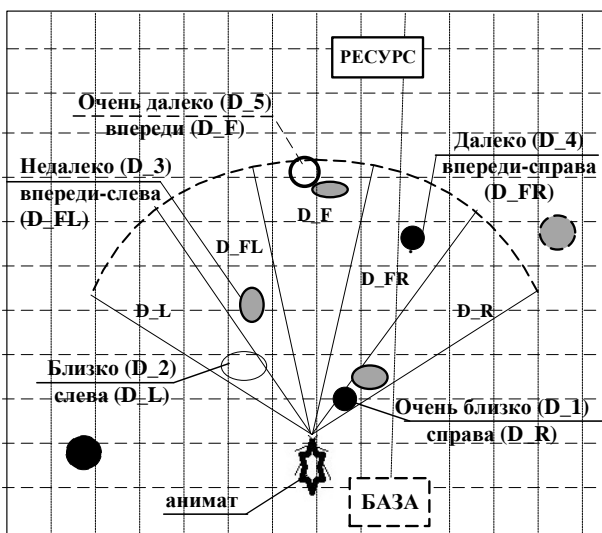


Рис. 3. Представление мира анимата: область видимости, направление, расстояние

Мир анимата должен быть представлен так, чтобы можно было при описании модели оперировать относительными категориями, без привязки к абсолютным физическим величинам. Для этого полигон виртуально разбивается на клетки, размер которых определяется характерным линейным размером робота (модельным или реальным) (рис. 3), а для понятий «расстояние» и «направление» вводятся порядковые шкалы и правила пересчета реальных значений в условные единицы (клетки).

3. Описание маршрута

Пусть в мире анимата существует множество Ω из N распознаваемых им объектов x_n : $\Omega = \{x_n\}$, $n=1, \dots, N$. Под *ориентиром* Q будем понимать один или несколько объектов, образующих компактную группу, т.е. расположенных близко друг к другу:

$$Q = \{x_i\} (\{x_i\} \subset \Omega): \forall x_k \in Q \exists x_m: x_m \in Q, \rho(x_k, x_m) \leq \Delta. \quad (3.1)$$

Здесь ρ – расстояние между объектами (например, евклидово), а Δ – верхняя граница диапазона «очень близко». Наличие одинаковых ориентиров приводит к необходимости распознавания сцен, а не отдельных ориентиров или объектов [Московский, 2015]. Сцена является совокупностью видимых ориентиров, пространственных отношений между ними, направления по компасу и временной составляющей. При этом сцена определяет поведение анимата и изменение его состояния.

Маршрут описывается как последовательность сцен $\Omega = \{S_j\}$, $j=1, \dots, L_p$, где L_p – длина маршрута. Каждая сцена S_j состоит из множества ориентиров $\{O_i^j\}$, один из которых находится в *фокусе внимания* и является опорным (относительно него совершаются все действия). Каждый ориентир O_i^j , в свою очередь, состоит из множества наблюдаемых объектов с их характеристиками (контекстом) V : $O_i^j = \{V_k\}$

$$V_k = \{\text{Id}, X, \{(\text{dir}_i, \text{dist}_i)\}, \text{Compas}, \text{Timer}\}, \quad (3.2)$$

где Id – идентификатор объекта (например, цвет); X – признак опорного ориентира (направление обхода ориентира); $\{(\text{dir}_i, \text{dist}_i)\}$ – множество пар, хранящих направление и расстояние до левой границы, центра объекта и правой границы соответственно; Compas – значение датчика «компас»; Timer – номера тактов в начале и в конце отрезка пути.

Для сопоставления сцен используются правила псевдофизической (пространственной) логики [Pospelov, 1987] и сравнение ориентиров с учетом контекста (подробнее см. [Карпова, 2022]).

4. Преобразование прямого маршрута в обратный

Для возвращения на базу описание маршрута Ψ преобразовывалось в обратный маршрут Ψ' . Общий принцип преобразования описания маршрута Ψ в обратный маршрут Ψ' заключается в следующем. Обратный маршрут состоит из тех же сцен, что и прямой, но в обратном порядке. Для прямого маршрута конечная точка – искомый ресурс, для обратного – «база». Ориентиры каждой сцены зеркально отображаются слева направо, направление движения меняется на противоположное. При этом короткие сцены пропускаются: это повышает устойчивость анимата (и робота) на маршруте. Алгоритм преобразования прямого маршрута (DirectRoute) в обратный (BackRoute):


```

BackRoute := [] # инициализация обратного маршрута пустым списком
Base := [сцена из одного ориентира – «база»]
Минимальное время нахождения на участке := 50
current_time := (Текущее модельное время)
период_2 := (период времени на последнем участке маршрута)
направление_2 := (направление движения на последнем участке м-та)
Цикл по Line из (DirectRoute) # в обратном порядке
# Начиная с предпоследнего участка, т.к. последняя сцена не будет
# видна после разворота анимата в противоположном направлении
Scene := [] # инициализация новой сцены описания маршрута
направление_2 := (развернуть на 180°)
направление_1 := (направление движения на участке Line)
период_1 := (период времени на участке Line)
time = current_time + период_2
Если (период_2) < (Минимальное время нахождения на участке),
то направление_2 := направление_1 # запомнить направление
    период_2 := период_1 # запомнить период
    current_time := time # запомнить новое текущее время
    continue # пропустить «короткую» сцену
к.с.
Цикл по Landmark из (Line) # в обратном порядке – слева направо
направление_БО = (направление обхода базового ориентира)
Если направление_БО = слева, то направление_БО := справа
иначе-если направление_БО = справа, то направление_БО := слева
к.с.
Orientir := [] # инициализация нового ориентира пустым списком
Цикл по Object из (Landmark) # в обратном порядке – слева направо
    Добавить Object к Orientir
    Пересчитать границы видимости для Object
к.п. по Object
# Для Orientir установить новые направления и период:
Установить направление обхода баз.ориентира (направление_БО)
Установить направление движения на участке (направление_2)
Установить период движения на участке (период_2)
Добавить Orientir к Scene
к.п. по Landmark
направление_2 := направление_1 # запомнить направление
период_2 := период_1 # запомнить период
current_time := time # запомнить новое текущее время
Добавить Scene к Route
к.п. по Line
# добавить конечную точку обратного маршрута – «базу»
направление_2 := (развернуть на 180°)

```


Эксперименты проводились в системе многоагентного моделирования *Kvoqum*, созданной в НИЦ «Курчатовский институт» [Карпов и др., 2018]. Полигон – область 200×200 клеток, время моделирования – 20000 тактов для каждого эксперимента. Исследовались различные конфигурации ориентиров на полигоне, и на каждой конфигурации проводилось 5 запусков в разных направлениях (всего – 50 экспериментов). Статистика моделирования приведена в табл. 1.

Таблица 1

Результаты имитационного моделирования

Кол-во экспериментов	Описание	Кол-во проходов в одном эксперименте	Кол-во возвратов в одном эксперименте	Общее кол-во неудач	Общее кол-во успешных проходов
35	нашел ресурс, вернулся, пошел 2-й раз, вернулся	4	2	0	70
3	нашел ресурс не сразу, вернулся, пошел 2-й раз, вернулся	от 6 до 12	от 3 до 6	0	12
5	нашел ресурс, вернулся, пошел еще раз, не нашел ресурс, не вернулся	3	1	5	5
2	Нашел ресурс, вернулся, 2-й раз не нашел, но вернулся	3	1	2	7
5	не нашел ресурс, вернулся, ходил повторно на поиски	от 14 до 48	от 7 до 24	2	87

Если не учитывать эксперименты, в которых анимат все время возвращался из-за того, что не видит ориентиров, то из 101-го прохода туда-обратно анимат не вернулся 7 раз (примерно 93% успешных проходов). Если считать полностью успешными эксперименты, в которых анимат дважды дошел до одного и того же ресурса, то таких было 38: это 84% от 45 экспериментов, в которых конфигурация ориентиров позволяла анимату успешно решить задачу.

6.2. Результаты натуральных экспериментов

Натурные эксперименты проводились на мобильных платформах – мини-роботах серии YARP. Мини-робот – это платформа с дифференциальным приводом, оснащенная дальномерами, гироскопом, камерой и бортовым вычислителем Raspberry Pi 4. Экспериментальный комплекс пред-

ставляет собой мобильную платформу, удаленный управляющий компьютер и полигон, на котором расположены ArUco-маркеры. Комплекс работает под управлением ROS (Robot Operating System). На бортовом вычислителе развернуты компоненты ROS, отвечающие за выполнение двигательных функций и обработку данных с камеры. Камера – основной источник сведений о наблюдаемых роботом объектах – ориентирах с нанесенными ArUco-маркерами. В экспериментах участвовали роботы на двух платформах серии YARP – платформе YARP-132 и YARP-133. Эти платформы имеют одинаковую архитектуру системы управления, но отличаются по габаритам и скоростям движения. Несмотря на это, итоговые результаты их работы по запоминанию и обработке маршрутов были идентичными. Это свидетельствует о независимости разработанной поведенческой модели от физической реализации. Для экспериментов использовался полигон Лаборатории робототехники «Курчатовского института» (рис. 5).



Рис. 5. Натурные эксперименты на полигоне (YARP-133)

При переносе моделирующей программы на реального робота была произведена калибровка камеры и настройка скоростей движения робота, а также написан специальный модуль для перевода данных камеры в формат системы Kvozum. Для исключения ложных срабатываний камеры на полигоне было введено ограничение на номера распознаваемых маркеров (от 1 до 9). Никаких других изменений в управляющую программу вносить не потребовалось. Таким образом, одна и та же программа использовалась и для имитационных экспериментов, и для управления роботом на реальном полигоне. В ходе натурных экспериментов роботы вели себя абсолютно аналогично аниматам.

Заключение

Сравнение полученных результатов с аналогичными работами вызывает определенные сложности. Условия проведения экспериментов сильно отличаются, поэтому сравнение статистических показателей было бы некорректным. Большинство работ по данной тематике апеллирует к точности прохождения маршрута [Lambrinos et al., 2000]; в данном же исследовании точность не играет существенной роли. Вместо панорамных камер (как в [Augustine et al., 2012], [Yu and Kim, 2011] и др.) наши роботы имели угол развертки камеры 120°. Поэтому сравнивать с аналогами можно только по качественным показателям.

По сравнению с работами, которые ограничивались имитационным моделированием, данный метод имеет преимущества, т.к. он проверен на различных робототехнических платформах и в реальных условиях. По отношению к тем работам ([Lambrinos et al., 2000], [Dall'Osto et al., 2021]), в которых роботы действуют в реальной среде, данный метод несколько проигрывает, т.к. для него необходимы определенные ориентиры, которые он умеет распознавать. Переход к произвольным ориентирам повлечет усложнение процедуры распознавания и увеличение объема памяти для хранения описания маршрута. Если же говорить о методе в целом, то к его достоинствам можно отнести невысокие требования к точности позиционирования, отсутствие требования уникальности ориентиров, работу с камерой с небольшим углом развертки, обход препятствий и возможность возвращения в точку отправления.

Благодарности. Автор выражает благодарность сотрудникам лаборатории робототехники НИЦ «Курчатовский институт», на базе которого проводились эксперименты, и, в первую очередь, руководителю лаборатории Карпову В.Э.

Список литературы

- [Augustine et al., 2012] Augustine M., Ortmeier F., Mair E., Burschka D., Stelzer A., Suppa M. Landmark-Tree map: A biologically inspired topological map for long-distance robot navigation. // IEEE International Conference on Robotics and Biomimetics, ROBOT 2012 – Conference Digest.
- [Cruse et al., 2011] Cruse H., Wehner R. No need for a cognitive map: Decentralized memory for insect navigation // PLoS Comput. Biol. – 2011, – No. 7 (3).
- [Dall'Osto et al., 2021] Dall'Osto D., Fischer T., Milford M. Fast and Robust Bio-Inspired Teach and Repeat Navigation // 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).
- [Dorigo et al., 2005] Dorigo M., Blum C. Ant colony optimization theory: A survey // Theoretical Computer Science. – 2005. – No. 344 (2-3).

- [**Dupeyroux et al., 2019**] Dupeyroux J., Viollet S., Serres J.R. An ant-inspired celestial compass applied to autonomous outdoor robot navigation // *Robotics and Autonomous Systems*. – 2019. – No. 117.
- [**Goldschmidt et al., 2017**] Goldschmidt D., Manoonpong P., Dasgupta S. A neurocomputational model of goal-directed navigation in insect-inspired artificial agents // *Frontiers in Neurorobotics*. – 2017. – No. 11 (APR).
- [**Lambrinos et al., 2000**] Lambrinos D., Möller R., Labhart T., Pfeifer R., Wehner R. A mobile robot employing insect strategies for navigation // *Robotics and Autonomous Systems*. – 2000. – No. 30 (1).
- [**Mair et al., 2014**] Mair E., Augustine M., Jäger B., Stelzer A., Brand C., Burschka D., Suppa M. A biologically inspired navigation concept based on the Landmark-Tree map for efficient long-distance robot navigation // *Advanced Robotics*. – 2014. – 28 (5).
- [**Pospelov, 1987**] Pospelov D.A. Fuzzy reasoning in pseudo-physical logics // *Fuzzy Sets and Systems*. – 1987. – No. 22 (1–2).
- [**Vardy, 2006**] Vardy A. Long-range visual homing. // *IEEE International Conference on Robotics and Biomimetics, ROBIO 2006*.
- [**Wilson, 1987**] Wilson S.W. Classifier Systems and the Animat Problem // *Machine Learning*. – 1987. – No. 2 (3).
- [**Yu et al., 2011**] Yu S.E., Kim D. Image-based homing navigation with landmark arrangement matching // *Information Sciences*. – 2011. – No. 181 (16).
- [**Длусский, 1967**] Длусский Г.М. Муравьи рода *Формика*. – М.: Наука, 1967.
- [**Захаров и др., 2013**] Захаров А.А., Длусский Г.М., Горюнов Д.Н., Гилев А.В., Зрянин В.А., Федосеева Е.Б., Гороховская Е.А., Радченко А.Г. Мониторинг муравьев *Формика*. – КМК, 2013.
- [**Карпов, 2016**] Карпов В.Э. Моделирование социального поведения в групповой робототехнике // *Управление большими системами*. – № 59.
- [**Карпов, 2011**] Карпов В.Э. Об одном механизме реконструкции схем // VI Международная научно-практическая конференция “Интегрированные модели и мягкие вычисления в искусственном интеллекте”, 2011. Сб. науч. трудов. В 2-х т. Т. 1. – М.: Физматлит.
- [**Карпов и др., 2018**] Карпов В.Э., Ровбо М.А., Овсянникова Е.Е. Система моделирования поведения групп робототехнических агентов с элементами социальной организации Кворум // *Программные продукты и системы*. – 2018. – № 31 (3).
- [**Карпова, 2021**] Карпова И.П. Организация маршрута анимата на основе визуальных ориентиров и распознавания сцен // *Мехатроника. Автоматизация. Управление*. – 2021. – 22 (10).
- [**Карпова, 2022**] Карпова И.П. Об одном биоинспирированном подходе к ориентации роботов, или настоящий «муравьиный» алгоритм // *Управление большими системами*. – 2022. – № 96.
- [**Московский, 2015**] Московский А.Д. Метод распознавания сцен для задачи навигации мобильных роботов. // *Мат. II Всерос. научно-практ. семинара «Беспилотные транспортные средства с элементами искусственного интеллекта (БТС-ИИ-2015)»* (9.10.2015г., г. СПб, Россия): Политехника-сервис.

УДК 004.896

ДИНАМИЧЕСКОЕ ФОРМИРОВАНИЕ МАРШРУТА ПЕРЕМЕЩЕНИЯ РОБОТА В НЕДЕТЕРМИНИРОВАННЫХ СРЕДАХ С ОБХОДОМ СТАЦИОНАРНЫХ И НЕСТАЦИОНАРНЫХ ПРЕПЯТСТВИЙ

Д.В. Котов (*dim.kot2009@yandex.ru*)

Военная академия Генерального штаба ВС РФ, Москва

О.Б. Лебедев (*lebedev.ob@mail.ru*)

Южный федеральный университет, Таганрог

Д.Д. Фугаров (*ddf_1@mail.ru*)

Донской государственный технический университет, Ростов-на-Дону

В работе описывается гибридный алгоритм динамического формирования маршрута перемещения робота в недетерминированных средах с обходом стационарных и нестационарных препятствий для двумерного пространства, основанный на интеграции волнового и муравьиного алгоритмов, позволяющий строить в реальном масштабе времени траектории минимальной длины с одновременной оптимизацией ряда критериев качества построенного пути. Выявление ограничений, препятствующих прокладке траектории из текущей позиции, осуществляется в процессе ее построения. Процесс прокладки траектории осуществляется пошагово. Вся траектория, связывающая исходную позицию робота с целевой позицией, является совокупностью отдельных участков. Временная сложность алгоритма зависит от времени жизни колонии l (число итераций), количества вершин графа n , числа муравьев m , и определяется как $O(l \cdot n^2 \cdot m)$.

Ключевые слова: планирование траектории, частичная неопределенность, двумерное пространство, стационарные и нестационарные препятствия, волновой алгоритм, муравьиный алгоритм, гибридизация.

Введение

В связи с неопределенностью ограничений, препятствующих прокладке траекторий, перемещение робота по траектории к заданной цели происходит одновременно с ее формированием (прокладкой). Наличие неоп-

ределенности обусловлено тем, что применительно к роботу, такие препятствия не локализованы, т.е. система управления предварительно не имеет никакой информации ни об их форме, ни о размере, ни, о параметрах их перемещения и расположении. Выявление ограничений, препятствующих прокладке траектории из текущей позиции, осуществляется в процессе построения траектории [Пшихопов и др., 2014], [Пшихопов и др., 2011]. Разработана «аттракторно-репеллерная» методика синтеза стратегий обхода недетерминированных препятствий. Траектория строится динамически с учетом динамически изменяемых ограничений. В связи с этим процесс прокладки траектории осуществляется пошагово. Последовательно на каждом шаге относительно текущей позиции робота формируется локальная зона видимости (ЛЗВ). Анализируется рельеф стационарных препятствий и формируются его изображения на карте в виде горизонталей, устанавливающих границы приближения к стационарным препятствиям. В пределах ЛЗВ с помощью радара робота осуществляется определение текущих координат (локализация) всех нестационарных перемещающихся препятствий (ПП) параметры их перемещения – скорость и курс [Гузик и др., 2016], [Пшихопов и др., 2015]. Далее экстраполируются параметры перемещения ПП и на их основе рассчитываются прогнозируемые позиции и предполагаемые траектории перемещения ПП через промежуток времени δt , которые наносятся на карту. После этого робот в течение заданного интервала времени δt перемещается в пределах ЛЗВ с учетом прогнозируемых позиций и траекторий перемещения подвижных препятствий и не подвижных объектов по кратчайшему пути от текущей позиции к новой позиции, которая после ее достижения объявляется текущей [Карпенко, 2016], [Caro et al., 2005].

В ряде работ [Hoefler et al., 2011], [Нейдорф и др., 2016], [Fateme et al., 2013], [Chen et al., 2013] описываются подходы на основе различных эвристических методов к решению данной задачи, которые пока не показали слишком хороших результатов. Наибольшее распространение для построения маршрута на плоскости получили волновые алгоритмы [Карпенко, 2016]. Карта местности при волновой трассировке разбивается на квадраты (ячейки). Формируется рецепторная модель карты местности, включающая множество ячеек. Классический алгоритм волновой трассировки включает два этапа. На первом этапе в процессе распространения волны от источника к цели ячейкам дискретного рабочего поля (ДРП) присваиваются весовые оценки, связанные с принятым критерием оптимальности. На втором этапе алгоритма осуществляется построение пути. Для этого следует, начиная от ячейки-цели, двигаться в направлении, противоположном направлению распространения волны, переходя от ячейки с большим весом к смежной ячейке с меньшим весом до тех пор, пока не будет достигнута ячейка-источник.

В последнее время для решения различных «сложных» задач всё чаще используются способы, основанные на применении методов случайного направленного поиска. В основе большинства этих алгоритмов лежат метаэвристики, заимствованные в природе [Карпенко, 2016], [Caro et al., 2005]. Анализ методов решения сложных прикладных задач показывает, что применение любого одного алгоритма оптимизации (как классического, так и популяционного) далеко не всегда приводит к успеху. Поэтому одним из путей повышения эффективности решения оптимизационных задач в настоящее время является разработка гибридных популяционных алгоритмов [Raidl, 2006].

1. Постановка задачи и схема базового алгоритма

Карта местности первоначально разбивается на множество ячеек с заданным шагом дискретизации и представляется в виде дискретного рабочего поля (ДРП). Модель такого рода называют **рецепторной**. Рецепторную модель карты можно представить в виде графа $G=(X,Y)$, где $X=\{x_i|i=1,2,\dots,n\}$ – множество вершин, моделирующих ячейки. Если ячейки смежны, то соответствующие этим ячейкам вершины графа связаны ребром. $U=\{u_j|j=1,2,\dots,l\}$ – множество ребер. Вершина x_i связана с x_j ребром, если соответствующие элементы связаны соединением. Вершины графа представляют все возможные позиции размещения робота и препятствий в пространстве поиска.

Первой задачей является поиск и своевременное обнаружение препятствий, в качестве которых могут выступать здания, деревья, линии электропередач, слишком мощные воздушные потоки и даже зоны действия ПВО противника. На основе полученной информации формируется локальная зона видимости (ЛЗВ), на которой отражается пространство, доступное для перемещения робота. Данные о препятствиях поступают от сенсорной системы робота. Координаты препятствий совмещают с вершинами графа. Вершины помечаются как препятствия, и происходит удаление ребер вокруг каждой вершины. Поиск оптимального пути происходит в пространстве сформированного графа.

Основными показателями при планировании траектории являются: *Длина (P_L)*, *Показатель безопасности (S_m)*, *Время выполнения задания (t_m)*, *Коэффициент успешности миссии (M)* [Пшихопов В.Х. и др., 2014], [Пшихопов В.Х. и др., 2015].

P_L – длина всей траектории, пройденной роботом от начальной точки до цели.

S_m – минимальное расстояние между любым датчиком робота и любым препятствием вдоль всей траектории. Этот показатель определяет максимальный риск в течение всего движения.

t_m – время, необходимое для завершения движения.

V – количество успешных миссий в недетерминированных средах со сложными препятствиями.

В общем случае планирование траектории производится в соответствии с алгоритмом, представленном ниже.

Алгоритм планирования траектории.

1. $t=0$.
2. Робот помещается в текущую позицию (ТП) $p(t)$ локальной зоны видимости.
3. На основе данных сенсорной системы робота генерируется локальная зона видимости $z(t)$ относительно точки $p(t)$.
4. Все сгенерированные ранее ЛЗВ объединяются в одну объединенную ЛЗВ (ОЛЗВ) $Oz(t)$. Препятствия за пределами ОЛЗВ $Oz(t)$ не рассматриваются.
5. По информации сенсорной системы робота формируются координаты расположения в пределах ОЛЗВ множества стационарных и нестационарных препятствий в момент времени $t=0$: $K=\{k_{oi} | i=1, 2, \dots, n\}$, $k_{oi}=(x_{oi}, y_{oi})$.
6. Множество препятствий K разбивается на подмножества $K_0=K_1 \cup K_2$:
 K_1 – множество неподвижных препятствий;
 K_2 – множество подвижных препятствий в момент времени $t=0$;
 $K_0=\{k_{oi} | i=1, 2, \dots, n\}$, $k_{oi}=(x_{oi}, y_{oi})$.
7. Для каждого подвижного препятствия фиксируются параметры передвижения: скорость и направление передвижения каждого препятствия.
8. На карте ЛЗВ отражается размещение всех неподвижных препятствий и начальных позиций подвижных препятствий.
9. Методами экстраполяции рассчитываются прогнозируемые координаты $K_p=\{k_{pi} | i=1, 2, \dots, n\}$, $k_{pi}=(x_{pi}, y_{pi})$ подвижных препятствий через промежуток времени δt и траектории в виде отрезков, считывающие начальную и прогнозируемую позиции.
9. На карте ЛЗВ отражается размещение всех стационарных и нестационарных препятствий (подвижных и неподвижных) в моменты времени t и $(t+\delta t)$.
10. Множество позиций, в которых размещены объединенное множество всех препятствий в моменты времени t и $(t+\delta t)$, рассматривается, как множество стационарных (неподвижных препятствий).
11. Устанавливаются ограничения на возможное сближение робота с препятствиями и на длину S_i перемещения робота по траектории на шаге t .
12. Формируется граф поиска решений $G=(X, U)$, соответствующий текущей ЛЗВ.

13. На графе поиска решений $G=(X,U)$ волновым и муравьиным алгоритмами между точками $p(t)$ и *цель*, строится без чрезмерного сближения с препятствиями траектория перемещения робота.

14. На построенной траектории выбирается новая текущая позиция $p(t+1)$ робота, расположенная на расстоянии S_t от позиции $p(t)$.

15. Если точка *цель* расположена между позициями $p(t)$ и $p(t+1)$, то траектория сформирована полностью и переход к 17, иначе переход к п.16.

16. Удаляются все спрогнозированные на шаге t позиции. Переход к п.3.

17. Конец работы алгоритма.

Отметим, что реальное расположение ПП на шаге $t+1$ *может отличаться от ранее прогнозируемого* на шаге t , поэтому на каждом шаге используются реальные координаты размещения ПП.

Интегральная оценка траектории рассчитывается по следующей формуле:

$$W=k_1P_L + k_2S_m + k_3t_m + k_4/V. \quad (1)$$

2. Построение на основе гибридизации волнового и муравьиного алгоритмов траектории в условиях частичной неопределенности

В общем случае интеграция основана на сращивании механизмов волнового и муравьиного алгоритмов и реализуется при выполнении двух основных этапов. На первом этапе путем распространения на ДРП волны от источника к цели формируется пространство поиска – R , представляющая совокупность связанных ячеек ДРП, достигнутых волной, с присвоенными им весами. На втором этапе муравьиным алгоритмом на базе сформированного пространства поиска R прокладывается маршрут.

Формируется рецепторная модель карты местности, включающая множество ячеек $E=\{e_i|i=1,2,\dots,n_e\}$. Каждая ячейка имеет координаты (x_i, y_i) . Задаются ячейка, которая помечается как исходная позиция $p(t)$, в которую помещается робот и ячейка, которая рассматривается как целевая позиция (*цель*). Изначально все ячейки считаются свободными.

Процесс планирования траектории итерационный и включает четыре этапа, выполняемых на каждой итерации t . На каждой итерации строится отдельный участок траектории, являющийся продолжением ранее построенного участка. А вся траектория является совокупностью отдельных участков, связывающая исходную позицию робота с целевой позицией.

На первом этапе по данным радара робота, размещенного на данной итерации t в точке (ячейке) $p(t)$ (текущая позиция), формируется локальная зона видимости: определяются ее границы на модели карты местности и выявляются, содержащиеся в ней препятствия. По информации сенсорной системы робота формируются координаты расположения множества препятствий в момент времени $t=0$: $K_0=\{k_{0i}|i=1,2,\dots,n\}$, $k_{0i}=(x_{0i}, y_{0i})$.

На втором этапе множество препятствий K разбивается на множество K_1 неподвижных объектов и множество подвижных объектов K_2 .

Для каждого подвижного препятствия радаром фиксируются параметры передвижения: скорость и направление передвижения каждого препятствия. Методами экстраполяции рассчитываются прогнозируемые координаты $K_p = \{k_{pi} / i=1, 2, \dots, n\}$, $k_{pi} = (x_{pi}, y_{pi})$ подвижных препятствий через промежутки времени δt .

На карте ЛЗВ отражается размещение всех препятствий стационарных и нестационарных препятствий (подвижных и неподвижных).

Ячейки с препятствиями помечаются. Координаты точки $p(t)$ определяются на предыдущей итерации ($t-1$). Данные радара, полученные на итерации t , объединяются с данными, полученными на основе экстраполяции.

На третьем этапе путем распространения на ДРП волны от точки источника (точка $p(t)$ к точке *цель* формируется пространство поиска траектории – $ППТ(t)$, представляющая совокупность связанных ячеек ДРП, достигнутых волной, с присвоенными им весами.

На четвертом этапе муравьиным алгоритмом на сформированной области $ППТ(t)$ прокладывается траектория $M(t)$ из точки *цель* к точке $p(t)$ с новой текущей позицией робота.

3. Муравьиный алгоритм прокладки маршрута

После распространения волны на ДРП от источника (точка $p(t)$ к точке *цель* и присвоение веса ячейкам ДРП муравьиным алгоритмом на ДРП из точки *цель* в точку $p(t)$ строится маршрут $M(t)$ [Лебедев, 2013].

Поиск решения осуществляется популяцией агентов на графе поиска решений $G=(X, U)$. Вершины множества X соответствуют ячейкам области R . Две вершины $x_i \in X$ и $x_j \in X$ связаны ребром u_k , если соответствующие им ячейки на модели местности в виде ДРП смежны и возможен переход соединения из одной ячейки в другую.

Каждая итерация l муравьиного алгоритма включает три этапа.

Предварительно на ребрах графа поиска решений $G=(X, U)$ откладывается начальное количество феромона: $\zeta = Q/n_u$, где $n_u = |U|$, Q – управляющий параметр.

На первом этапе каждый агент a_k популяции находит решение. Агент a_k конструктивным алгоритмом строит на графе G маршрут $M_k(t)$ из вершины *цель* в вершину $p(t)$.

Рассчитывается оценка $F_k(t)$ k -го решения, полученного агентом a_k на t -ой итерации.

На втором этапе каждый агент a_k откладывает феромон на ребрах маршрута, построенного на графовой модели пространства поиска, в количестве пропорциональном оценке $F_k(t)$ k -го решения $\tau_k(t) = \delta / F_k(t)$. В ра-

боте используется циклический (ant-cycle) метод муравьиных систем [Лебедев, 2013]. В этом случае феромон откладывается совместно агентами на ребрах графа после полного формирования решений всей популяцией агентов.

На третьем этапе осуществляется общее испарение (уменьшение) количества феромона на всех ребрах графа поиска решений $G=(X,U)$ на некоторую величину, в соответствии с формулой $h_j=h_j(1-\sigma)$, где σ – коэффициент обновления, h_j – суммарное количество феромона, отложенного муравьями на ребре $u_j \in U$.

Построение муравьем траектории осуществляется на базе сформированного пространства поиска $ППТ(t)$ после распространения на ДРП волны от источника к точке *цель*, с присвоением весов свободным ячейкам. Подобно волновому алгоритму муравей строит на $ППТ(t)$ маршрут, представляющий совокупность ячеек с последовательно уменьшающимся весом. Всеми муравьями будут построены маршруты одной минимальной длины, но отличающимися конфигурациями. В этом случае оценкой маршрута $M_k(t)$, построенного агентом a_k служит оценка его конфигурации.

$$F_k = \alpha f_{rk} + \beta f_{gk}, \quad (2)$$

где α, β – управляющие параметры, которые подбираются экспериментально;

f_{rk} – число поворотов маршрута M_k ;

f_{gk} – число содержащих препятствия ячеек, смежных ячейкам маршрута M_k .

Чем меньше f_{rk} , тем меньше показатель *Время выполнения задания* – время, необходимое для завершения движения.

Чем меньше f_{gk} , тем выше *Показатель безопасности* – число позиций ПО на минимальном расстоянии между ПО и любым препятствием вдоль всей траектории.

Цель оптимизации – минимизация значения F_k .

Отметим, что показатели P_L, S_m, t_m, V не являются объектами оптимизационных процедур. Объектом оптимизации является показатель $F_k = \alpha f_{rk} + \beta f_{gk}$, минимизация которого на базе эвристик гибридного алгоритма планирования косвенно способствует минимизации интегральной оценки траектории: $W=k_1P_L+k_2S_m+k_3t_m+k_4V$.

4. Экспериментальные исследования

Для проведения экспериментов была использована процедура синтеза контрольных примеров с известным оптимумом по аналогии с известным методом PEKU (Placement Examples with Known Upper bounds on wirelength).

Для задачи планирования траектории созданы четыре набора примеров с известным оптимумом по длине траектории и оптимальным значением F_{opt} оценки конфигурации траектории F_k . Оптимизация проводилась по критерию минимальной длины траектории и оптимального значения оценки конфигурации и опиралась на методику, изложенную в работе [Пшихопов и др., 2014], [Пшихопов и др., 2015].

На основе обработки экспериментальных исследований была построена средняя зависимость качества решений от числа итераций (рис. 1). Оценкой качества служит величина F_k/F_{opt} , где F – оценка полученного решения. Исследования показали, что число итераций, при которых алгоритм находил лучшее решение, которое лежит в пределах 68-80. Из графика видно, что в среднем на 76-й итерации решение близко к оптимальному.

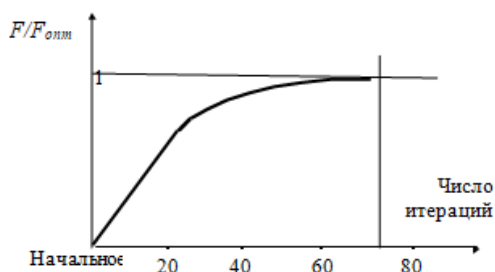


Рис. 1. Средняя зависимость качества решений от числа итераций

Общая оценка временной сложности при любом подходе к гибридации лежит в пределах $O(n^2)$ – $O(n^3)$.

Сравнение значений критерия, полученных гибридным алгоритмом на тестовых примерах с известным оптимумом, показало, что у 60% примеров полученное решение было оптимальным; у 25% примеров решения отличались от оптимальных не более чем на 4%; у 15% примеров решения отличались от оптимальных не более, чем на 2%.

Сравнительный анализ производился на тестовом наборе сцен, приведенном в работе [Пшихопов и др., 2014]. Сравнение интегральной оценки F_T разработанного гибридного алгоритма с известными алгоритмами [Fateme et al., 2013], [Chen et al., 2013], [Raidl, 2006], [Лебедев, 2013], [Veselov et al., 2020], [Веселов и др., 2021], [Городецкий, 2016] по траектории W показало, что при меньшем времени работы у полученных с помощью разработанного алгоритма решений значения F_T меньше в среднем на 6%.

Заключение

В работе решалась задача ситуационной прокладки маршрута перемещения робота в условиях неопределенности и наличия стационарных и нестационарных препятствий. Несмотря на достаточно большое число разработанных моделей, и алгоритмов управления подвижными объектами в условиях неопределенности исследователи часто сталкиваются с рядом проблем, к числу которых относятся трудность в обосновании качества результатов анализа, учитывающего специфику конкретной задачи.

К числу перспективных тенденций относятся разработка гибридных биоинспирированных методов. В гибридных алгоритмах, объединяющих различные, либо одинаковые алгоритмы, но с различными значениями свободных параметров, преимущества одного алгоритма могут компенсировать недостатки другого. В работе предлагается гибридный алгоритм, основанный на интеграции волнового и муравьиного алгоритмов, позволяющий строить в реальном масштабе времени траектории минимальной длины с одновременной оптимизацией ряда других критериев качества построенного пути. Отличительная особенность заключается в том, что перемещение подвижного объекта по траектории происходит производится практически одновременно с ее формированием (прокладкой), в режиме *on-line*. Ключевая проблема, которая была решена в данной работе это методика сращивания механизмов муравьиного и волнового алгоритмов. Источником усовершенствования может быть более детальная проработка вопросов, связанных с формированием зоны видимости и выбором (сменной) текущих исходных позиций, что ускорит процесс достижения подвижным объектом целевого состояния.

Список литературы

- [Пшихопов и др., 2014] Пшихопов В.Х. и др. Интеллектуальное планирование траекторий подвижных объектов в средах с препятствиями. – М.: Физматлит, 2014. – 197 с.
- [Пшихопов и др., 2011] Пшихопов В.Х., Медведев М.Ю. Управление подвижными объектами в определенных и неопределенных средах. – М.: Наука, 2011. – С. 194–207.
- [Гузик и др., 2016] Гузик В.Ф., Переверзев В.А., Пьявченко А.О., Сапрыкин Р.В. Принципы построения экстраполирующего многомерного нейросетевого планировщика интеллектуальной системы позиционно-траекторного управления подвижными объектами // Известия ЮФУ. Технические науки. – 2016. – С. 67-80.
- [Пшихопов и др., 2015] Пшихопов В.Х., Медведев М.Ю., Гуренко Б.В. Алгоритмы адаптивных позиционно-траекторных систем управления подвижными объектами // Проблемы управления, 2015. – С. 66-74.
- [Карпенко, 2016] Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учебное пособие. Второе издание. – М: Изд-во МГТУ им. Н.Э. Баумана, 2016. – 448 с.

- [Caro et al., 2005] Caro G.D., Ducatelle F., Gambardella L.M. AntHocNet: An adaptive nature inspired algorithm for routing in mobile ad hoc networks. // European Transactions on Telecommunications, 2005. – pp. 443-455.
- [Hoefler et al., 2011] Hoefler T., Snir M. Generic Topology Mapping Strategies for Large-scale Parallel Architectures. – University of Illinois at Urbana-Champaign Urbana, 2011. – P. 75-85.
- [Нейдорф и др., 2016] Нейдорф Р.А., Полях В.В., Черногоров И.В., Ярахмедов О.Т. Исследование эвристических алгоритмов в задачах прокладки и оптимизация маршрутов в среде с препятствиями // Известия ЮФУ. Технические науки, 2016. – С. 127-143.
- [Fatemeh et al., 2013] Fatemeh K.P., Fardad F., Reza S.N. Comparing the Performance of Genetic Algorithm and Ant Colony Optimization Algorithm for Mobile Robot Path Planning in the Dynamic Environments with Different Complexities // Journal of Academic and Applied Studies. – 2013. – P. 29-44.
- [Chen et al., 2013] Chen S., Eshaghian M.M. A fast recursive mapping algorithm // Department of computer and information science. – 2013. – P. 219-227.
- [Raidl, 2006] Raidl G.R. A Unified View on Hybrid Metaheuristics // Lecture Notes in Computer Science. – 2006. – P. 1-12.
- [Лебедев, 2013] Лебедев О.Б. Модели адаптивного поведения муравьиной колонии в задачах проектирования. – Изд-во ЮФУ, 2013. – 199 с.
- [Veselov et al., 2020] Veselov G.E., Lebedev B.K., Lebedev O.B. Management of Behavior of a Swarm of Robots Applicable to the Tasks of Monitoring a Some Territory // Artificial Intelligence Methods in Intelligent Algorithms, 2020. – P. 324-332.
- [Веселов и др., 2021] Веселов Г.Е., Лебедев Б.К., Лебедев О.Б. Управление движением группы мобильных роботов в колонне // Научно-практический журнал Информатизация и связь. – 2021. – С. 7-11.
- [Городецкий, 2016] Городецкий В.И. Управление коллективным поведением роботов в автономной миссии// Робототехника и техническая кибернетика. – 2016. – С. 40-54.

ФОРМИРОВАНИЕ УМЕНИЙ АГЕНТА ПО ПРИНЦИПУ ДОСТИЖИМОСТИ В ОБУЧЕНИИ С ПОДКРЕПЛЕНИЕМ

А.К. Латышев (*latyshev.ak@phystech.edu*)
Московский физико-технический институт
(национальный исследовательский университет), Москва

А.И. Панов (*panov@airi.net*)
Институт искусственного интеллекта AIRI, Москва
Федеральный исследовательский центр
«Информатика и управление» РАН, Москва

В работе рассматривается способ формирования набора умений агента без использования информации о поставленной перед ним цели. В обучении с подкреплением подобные способы относят к методам внутренней мотивации. Предложенный в настоящей работе алгоритм берет за основу вычисление взаимной информации между действиями агента и его состояниями. Такой подход позволяет конструировать умения, приводящие к наиболее равномерному достижению возможных состояний. В работе представлены формируемые алгоритмом умения на примере непрерывного управления движением агента. Проанализирована применимость получаемых умений для решения задач обучения с подкреплением.

Ключевые слова: формирование умений, взаимная информация, достижимость, внутренняя мотивация, обучение с подкреплением.

Введение

Построение систем искусственного интеллекта, способных к автономному обучению в течении длительного периода – одна из важнейших задач в теории интеллектуальных агентов [Тарасов, 2002], [Colas et al., 2022]. Формирование умений – переиспользуемых подстратегий, говоря другими словами, универсальных¹ сценариев поведения агента – является одной из важнейших проблем, стоящих перед исследователями, развивающими область обучения с подкреплением. Агент, обладающий широ-

¹ в контексте многократного использования для достижения целей в среде

ким набором разнообразных заранее сформированных умений, способен значительно быстрее и эффективнее учиться достигать цели по сравнению с агентом без каких-либо умений.

В задаче формирования умений можно выделить два подхода. В первом агент учится достигать определенные состояния в среде, выбор которых опирается на стратегию более высокого уровня (Director [Hafner et al., 2022]), прошлый опыт (HER [Andrychowicz et al., 2017]), исследовательский опыт (LEXA [Mendonca et al., 2021]), обученный граф взаимодействий со средой (SRICS [Zadaianchuk et al., 2022]). Второй подход направлен на конструирование умений без привязки к целевому состоянию. Основное внимание уделяется выполняемым действиям и получаемым при этом состояниям. Умения в таком случае оптимизируют функционал, характеризующий связь действий и состояний, например, взаимную информацию (DIAYN [Eysenbach et al., 2018], DADS [Sharma et al., 2020]).

Для получения универсальных умений в конкретной среде важны методы, формирующие умения без использования информации о вознаграждении в этой среде. Такие методы входят в состав обширной группы, использующей подход внутренней мотивации. Это подход обучения агентов в режиме самостоятельного определения вознаграждающего внутреннего сигнала, который близок к тому способу, как устроена внутренняя мотивация у человека в соответствии с некоторыми моделями [Ryan et al., 2000].

Настоящая работа предлагает метод формирования умений на основе принципа достижимости состояний, определяемого через взаимную информацию между последовательностью действий агента и конечными состояниями. В данной работе рассматриваются примеры формируемых умений для задач управления физическим телом агента в симуляторе. В этих задачах среды являются детерминированным и предполагают наличие единственного агента.

1. Постановка задачи

Классическим подходом к рассмотрению задачи обучения искусственного агента является формализм марковского процесса принятия решений (МППР) – $\langle S, A, T, R, \rho, \gamma \rangle$. Агент последовательно исполняет действия $a_t \in A$ в среде и получает вознаграждение $R : S \times A \times S \rightarrow \mathbb{R}$, соответствующее переходу из текущего состояния $s_t \in S$ в следующее $s_{t+1} \in S$. Пространства состояний $S \in \mathbb{R}^2$ и действий $A \in \mathbb{R}^l$ в настоящей работе являются подпространствами непрерывного векторного пространства со стандартной L_2 нормой, где параметр l зависит от конкретной среды (см. раздел 3). Динамика среды определяется моделью переходов

$T : S \times A \times S \rightarrow \mathbb{R}$, задающей распределение $T(s_{t+1} | s_t, a_t)$. Задача алгоритма обучения с подкреплением найти оптимальную стратегию $\pi^* : S \times A \rightarrow \mathbb{R}$, максимизирующую дисконтированную отдачу:

$$\pi^* = \arg \max_{\pi} \mathbb{E}_{\rho(s_0), \pi(a_t | s_t), T(s_{t+1} | s_t, a_t)} \sum_{t=0}^{\infty} \gamma^t R(s_t, a_t, s_{t+1}), \quad (1.1)$$

где ρ определяет распределение начальных состояний агента.

Формирование умений агента состоит в определении множества умений $Z = \{z = \langle a^0, a^1, \dots, a^{n-1} \rangle | a^i \in A\}$, где каждый элемент множества определяет последовательность из n действий. При обучении агента с умениями решается следующая задача МППР: $\langle S, Z, T, R_n, \rho, \gamma \rangle$. Агент выбирает умение z_t в соответствии со стратегией $\pi_z : S \times Z \rightarrow \mathbb{R}$ для текущего состояния s_t и исполняет в среде последовательность действий, определенную умением $z_t = \langle a_t^0, a_t^1, \dots, a_t^{n-1} \rangle$. Вознаграждение в таком случае определяется как сумма вознаграждений всех n переходов:

$$R_n(z_t, s_t^0, \dots, s_t^{n-1}, s_{t+1}^0) = \sum_{i=0}^{n-1} R(s_t^i, a_t^i, s_{t+1}^{i+1}), \quad (1.2)$$

а задачей алгоритма обучения агента является найти оптимальную стратегию выбора умений (как и (1.1), но с другим вознаграждением (1.2)):

$$\pi_z^* = \arg \max_{\pi_z} \mathbb{E}_{\rho(s_0), \pi_z(z_t | s_t), T(s_t^{i+1} | s_t^i, a_t^i)} \sum_{t=0}^{\infty} \gamma^t R_n(z_t, s_t^0, \dots, s_{t+1}^0). \quad (1.3)$$

Для формирования умений необходим критерий отбора последовательностей действий, образующих Z . В настоящей работе такой критерий опирается на теоретико-информационный подход. Взаимодействие агента со средой определяет канал передачи информации: «агент» – «среда». Последовательность действий агента определяет сообщение $\alpha_n = \langle a^0, a^1, \dots, a^{n-1} \rangle \in A^n$, отправленное по каналу, а состояние, в котором окажется агент по исполнению этих действий, $s_n \in S$ является сообщением, полученным по каналу. Пропускная способность C этого канала (*empowerment*, [Klyubin et al., 2005]) позволяет судить о возможности агента повлиять на свои будущие состояния:

$$C(s) = \max_{\omega(a^n | s)} I(A_{\omega}^n, S_{\omega} | s), \quad (1.4)$$

где A_{ω}^n – случайная величина последовательностей действий с распределением вероятностей $\omega(\alpha^n | s)$, S_{ω} – случайная величина с распределением:

$$p_{\omega}(s_n | s = s_0) = \int_{A^n} \omega(\alpha^n | s) \int_{S^{n-1}} \prod_{i=0}^{i=n-1} T(s_{i+1} | s_i, \alpha_i^n) ds_1 \dots ds_{n-1} d\alpha^n, \quad (1.5)$$

$I(A_{\omega}^n, S_{\omega} | s)$ – взаимная информация, а s – текущее начальное состояние.

Распределение $\omega(\alpha^n | s) = \arg \max I(A_{\omega}^n, S_{\omega} | s)$ имеет очень полезное в обучении с подкреплением свойство – оно выделяет те последовательности действий, которые позволяют достичь максимальный объем состояний.

Точное вычисление пропускной способности для непрерывных состояний и действий требует вычисления интегралов по A^n и S , что является крайне сложной задачей, особенно когда интегрируемые функции определяются неявно через взаимодействие со средой (1.5).

Преодолеть проблему вычисления интегралов позволяет Монте-Карло интегрирование Юнга [Jung et al., 2011], которое является модификацией стандартного алгоритма Блахут-Аримото [Salge et al., 2014] для вычисления пропускной способности информационного канала. Алгоритм позволяет работать с любого вида зависимостями между действиями и состояниями (T и (1.5)), так как формирует выборку достижимых состояний (случайной величины S_{ω}) на основе фиксированного конечного набора последовательностей действий $\{\alpha_i^n | \alpha_i^n \in A^n, i \in \{0 \dots d-1\}\}$.

В настоящей работе предложенный нами способ формирования навыков (см. раздел 2) построен на итеративной фильтрации с получением наиболее репрезентативного такого набора, определяющего наибольшую приближенную пропускную способность (1.4). Получаемый набор – множество сформированных навыков.

2. Формирование умений агента

Процесс формирования умений состоит из следующих шагов:

- 1) инициализация множества умений Z_0 ;
- 2) определение тренировочного множества умений D_i ;
- 3) шаг работы Монте-Карло интегрирования Юнга на D_i ;
- 4) топ- h фильтрация умений из D_i и внесение их в Z_i ;
- 5) возврат к шагу 2 (повторять N итераций);
- 6) топ- k фильтрация итогового множества Z_N .

2.1. Инициализация множества умений Z

Первый шаг формирования умений состоит в инициализации множества умений. Будем далее рассматривать ограниченное единичным кубом непрерывное пространство действий в виде $A = [0, 1]^p$. Тогда последовательность действий лежит в pn -мерном единичном кубе $A^n = [0, 1]^{pn}$.

Z_0 инициализируется выборкой из равномерного распределения U_{A^n} . Размер b множества Z_0 изначально устанавливается большим, чем требуется в итоге количество сформированных умений $k \leq b$. Это необходимо для финальной фильтрации на шаге 6.

2.2. Тренировочное множество D

Начальный этап каждой i -ой итерации алгоритма (где $i = 1 \dots N$) определяет тренировочное множество D_i размером d , которое состоит из двух частей:

- выборки h элементов без повторов из Z_{i-1} с равномерным дискретным распределением;
- выборки $(d - h)$ элементов из равномерного непрерывного распределения $U[0, 1]^m$, как при инициализации Z_0 (см. раздел 2.1).

При взятии выборки из Z_{i-1} оставшаяся часть Z_{i-1} задает основу для Z_i , чтобы получаемое на следующем шаге (см. раздел 2.3) распределение умений было согласовано с множеством Z_i . Наличие второй части тренировочного множества обеспечивает постоянное исследование всего пространства возможных последовательностей действий.

2.3. Монте-Карло интегрирование Юнга

Для вычисления аппроксимации $\tilde{C}_D(s) \approx C(s)$ пропускной способности (1.4) необходима модель переходов в среде $p(s_z | s, z = \alpha^n)$. Истинная модель динамики среды $T(s_a | s, a)$ доступна только в виде «черного ящика» среды, что обеспечивает агента выборкой из распределения $s \sim p(s_z | s, z)$, но не им самим. В настоящей работе для аппроксимации истинного распределения $p(s_z | s, z = \alpha^n)$ используется нормальное распределение $N(\tilde{\mu}_{s_z}, \tilde{\sigma}_{s_z}^2)$, где $\tilde{\mu}_{s_z}, \tilde{\sigma}_{s_z}$ – математическое ожидание и стандартное отклонение выборки N_{MC} элементов (для многомерных состояний каждая отдельная компонента считается независимой), полученной из N_{MC} исполнений со случайным шумом одной последовательности действий $z = \alpha^n$ в среде из начального состояния $s = s_0$. Шум добавляется к каждому исполняемому действию и определяется нормальным распределением $N(0, \sigma^2 E)$ с диагональной ковариационной матрицей $\sigma^2 E$. Добавление шума крайне важно для детерминированных сред, так как иначе невозможно определить выборочное $\tilde{\sigma}_{s_z}$.

Определение распределений будущих состояний происходит для каждого умения из тренировочного множества $z = \alpha^n \in D$. В результате получаем множество $S_{D_i} = \{s_{\delta,\beta} \mid s_{\delta,\beta} \sim p(s_{z_\beta} \mid s, z_\beta), \delta \in \{1 \dots N_{MC}\}, z_\beta \in D_i\}$ из $d \times N_{MC}$ будущих состояний для всех умений из D_i , а также d нормальных распределений $N(\tilde{\mu}_{s_z}, \tilde{\sigma}_{s_z}^2)$ для каждого умения. Аппроксимационная плотность вероятности $p_{\delta\beta}^v = \tilde{p}(s_{\delta,\beta} \mid s, z_v)$ достичь $s_{\delta,\beta} \in S_{D_i}$ при условии выполнения умения $z_v \in D_i$ определяется плотностью нормального распределения, полученного ранее:

$$P_{\delta\beta}^v = P_{N(\tilde{\mu}_{s_{z_v}}, \tilde{\sigma}_{s_{z_v}}^2)}(s_{\delta,\beta} \mid s, z_v). \quad (2.1)$$

Перед началом итераций по алгоритму Монте-Карло интегрирования Юнга задается начальное приближение ω^0 для распределения умений $\omega_\beta = \tilde{p}(z_\beta \mid s)$, которое определяется равномерным дискретным распределением $\omega_\beta^0 = 1/d$.

Итерации происходят N_{JMCI} шагов. Для каждого шага $j=1, \dots, N_{JMCI}$ и каждого умения $z_v \in D_i$ обновление происходит с использованием уже посчитанных вероятностей (2.1) следующим способом:

$$c_v^j = \frac{1}{N_{MC}} \sum_{\delta=1}^{N_{MC}} \ln \left[\frac{P_{\delta v}^v}{\sum_{\beta=1}^d P_{\delta\beta}^v \omega_\beta^{j-1}} \right], \quad (2.2)$$

$$\omega_v^j = \frac{\omega_v^{j-1} \exp(c_v^j)}{\sum_{\beta=1}^d \omega_\beta^{j-1} \exp(c_\beta^j)}.$$

Итоговому распределению умений $\tilde{\omega}_i = \omega^{N_{JMCI}}$ соответствует пропускная способность:

$$\tilde{C}_{D_i}(s) = \tilde{C}_i = \sum_{v=1}^d \omega_v^{N_{JMCI}} c_v^{N_{JMCI}}. \quad (2.3)$$

2.4. Фильтрация умений

На предыдущем шаге получено распределение умений $\tilde{\omega}_i$ (2.2), на основе которого и производится фильтрация. Среди всех умений из тренировочного множества D_i выбирается h элементов с наибольшей вероятностью $\tilde{\omega}_i$. Эти топ- h умений добавляются к множеству умений Z_{i-1} , образуя новое множество Z_i . Причем размер Z_i остается b , так как до этого Z_{i-1} уменьшалось на h элементов (шаг 2, раздел 2.2).

В результате Монте-Карло интегрирования Юнга мы получаем помимо распределения умений ещё и пропускную способность. Стоит отметить, что когда изначальный алгоритм Блахут-Аримото применяется к дискретным распределениям состояний и действий, то пропускная способность определяет логарифм среднего числа достижимых агентом состояний. В адаптированном варианте для непрерывных переменных эта величина позволяет уже говорить о количестве областей пространства состояний, в которых оказывается агент. Что в свою очередь дает ориентир на то, сколько умений из всего множества D_i действительно приведут к различным исходам. Таким образом, при подборе параметров алгоритма, а именно размера h подвыборки (для шага 2, раздел 2.2) из Z_{i-1} , необходимо ориентироваться на $h \cong \exp(\tilde{C})$.

2.5. Критерий останова и финальная фильтрация

Алгоритм останавливается по выполнении N итераций. Для определения требуемого количества итераций для множества умений Z_i вычисляется пропускная способность \tilde{C}_{Z_i} аналогично шагу 3 (см. раздел 2.3). Установление пропускной способности на постоянном среднем значении означает достижение предела алгоритмом.

После того, как значение пропускной способности установилось, в Z_N остались наиболее удачные умения. Мы выбираем среди них топ- k (как на предыдущем шаге, см. раздел 2.4), которые и образуют финальный набор умений Z , используемый для решения других задач.

Такой итеративный способ получения последовательностей действий позволяет найти приближенно топ- k последовательностей из истинного $\omega(\alpha^{|s|})$ распределения, максимизирующего взаимную информацию (1.4).

3. Результаты формирования умений

Для полного представления о сформированных алгоритмом умениях были проведены вычислительные эксперименты, в которых рассмотрены получаемые траектории движения агента, а также изучена возможность практического применения умений в задаче обучения с подкреплением – достижение целевого состояния.

В качестве тестовых были выбраны среды из окружения OpenAI Gym [Brockman et al., 2016] по управлению агентом на плоскости с использованием физического симулятора MuJoCo [Todorov et al., 2012]. В этих средах агент управляет физической оболочкой на основе силового воздействия на актуаторы. В работе рассмотрены агенты Point и Ant. У первого $A \in [-1, 1]^2$ – две силы в плоскости пола xy . У второго $A \in [-1, 1]^8$ – 8 управляющих воздействий на сочленения. Также были проведены экспе-

рименты на игровой среде Toy (собственная реализация) – упрощенный вариант Point, где действия определяют, насколько сместится агент за один шаг в пределах окна $[-1,1]^2$ с центром в текущем положении агента.

3.1. Анализ умений

Для каждой из трех тестовых сред пространство состояний для формирования умений является пространство декартовых координат агента на плоскости. Начальное состояние – начало координат $s_0 = (0,0)$. Длина умения – последовательности действий его составляющих – $n = 10$, а множество умений состоит из $k = 20$ элементов. Остальные параметры алгоритма представлены в табл. 1.

Получаемые траектории сформированных умений (рис. 1, нижняя строка) отличаются друг от друга и приводят к широкому покрытию конечных состояний. Набор из случайных умений (определенных, как при инициализации множества Z) показывает траектории хуже по разнообразию и покрытию (рис. 1, верхняя строка). Можно также отметить, что алгоритм позволяет формировать умения для всех трех сред, в которых совершенно различны как динамика переходов, так и управляющие воздействия.

Таблица 1

Среда	σ	N_{MCI}	N_{MC}	N	d	h	b	$\exp(\tilde{C}_Z)$
Toy	0,3	10	30	100	20	10	30	10,8
Point	0,01	20	30	100	30	15	25	12,5
Ant	0,01	20	30	100	30	15	25	13,8

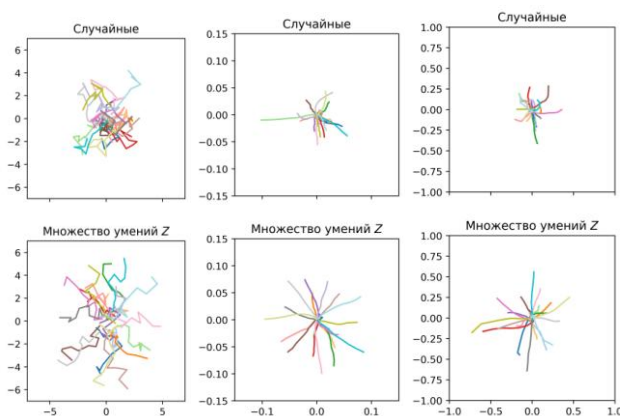


Рис. 1. Траектории умений для сформированного множества Z и случайного. По столбцам траектории в средах – Toy, Point, Ant соответственно

Одним из успешных алгоритмов формирования умений является DADS [Sharma et al., 2020], в котором на основе модели и методов глубокого обучения определяется пространство дискретных умений. На рис. 2 представлены траектории агента DADS. Стоит заметить, что умения DADS определены как стратегии $\pi(s_{t+1} | s_t, z)$ и не имеют конечной длины, как рассматриваемые в настоящей работе умения. Однако на начальных шагах наши умения не уступают по покрытию всех возможных направлений алгоритму DADS.

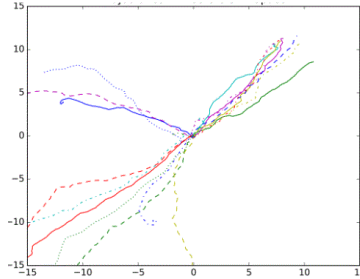


Рис. 2. Траектории умений из дискретного пространства в алгоритме DADS (рисунок из оригинальной работы [Sharma et al., 2020]) для среды Ant

3.2. Умения для достижения целей

Практическое применение сформированных умений проверяется на задаче достижения целей. Агент получает информацию о цели через её координаты, которые дополняют пространство состояний, а также через сигнал вознаграждения:

$$R(s_t) = - \frac{\|s_t^{xy} - g\|_2}{L \|s_0^{xy} - g\|_2}, \quad (3.1)$$

где s_t^{xy} – координаты положения агента на плоскости, $L = 200$ – фиксированная длина эпизода, g – координаты цели. Цели выбираются равномерно случайно из квадрата $[-15, 15]^2$ в начале каждого эпизода.

Для обучения стратегии выбора умений (1.3) использован алгоритм PPO [Schulman, 2017], реализация из библиотеки SB3 [Raffin et al., 2021]. Вознаграждение определяется по формуле (1.2) – суммированием по умению сигнала (3.1), где длина умения $n = 10$. Мы использовали два набора умений для сравнения. Первый – случайные последовательности действий. Второй множество умений, сформированных алгоритмом. В разделе 3.1 представлены эти умения для каждой из сред.

На рис. 3 представлено среднее расстояние между агентом и целью за эпизод по мере обучения:

$$\Delta = \frac{1}{L} \sum_{i=1}^L \frac{\|s_i^{xy} - g\|}{\|s_0^{xy} - g\|}, \quad (3.2)$$

обозначения как в (3.1). Отметим, что для среды Toy сформированные умения не дают изменения в производительности, но для Point и Ant изменение наблюдается, причем в обоих случаях сформированные умения улучшают качество работы агента. Однако для среды Ant это изменение не так существенно.

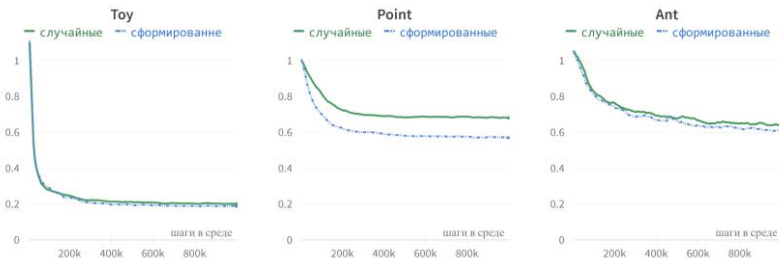


Рис. 3. Среднее относительное расстояние до цели за эпизод (3.2) для агента со случайными и сформированными умениями. Один шаг соответствует выполнению одного действия. Данные усреднены по 3 запускам с разными генераторами случайных чисел

Заключение

Предложенный алгоритм показал способность формировать дискретное множество умений агента, отличающихся большим разнообразием соответствующих траекторий агента. Эксперименты с тремя различными средами управления движением показали применимость нашего подхода как к совсем простому пространству действий (среда Toy), так и к сложному управлению сочленениями физической модели агента (среда Ant).

Практическая применимость сформированного множества умений была оценена в рамках достижения агентом целевой точки в пространстве состояний. Проведенные эксперименты обнаружили, что в случае простой задачи Toy умения, сформированные предложенным алгоритмом, не отличаются по качеству выполнения задачи от случайных. Однако для более сложных сред сформированные навыки обеспечивают более высокую скорость обучения. Было обнаружено, что для самой сложной задачи Ant этот прирост производительность не так существенен, как для задачи Point. Прояснение этого явления является задачей нашей текущей работы.

В настоящей работе пространство состояний агента ограничивалось декартовыми координатами на плоскости. Для применения предложенного алгоритма с многомерными пространствами состояний необходимо адаптировать процесс вычисления вероятностей предсказанных состояний, что влечет за собой внедрение обучаемой модели мира. Это направление деятельности является для нас приоритетом в дальнейших исследованиях.

Список литературы

- [Тарасов, 2002] Тарасов В.Б. От многоагентных систем к интеллектуальным организациям. – М.: Эдиториал УРСС, 2002.
- [Andrychowicz et al., 2017] Andrychowicz M., Wolski F., Ray A., Schneider J., Fong R., Welinder P., McGrew B., Tobin J., Pieter A. and Zaremba W. Hindsight Experience Replay // *Advances in Neural Information Processing Systems*. – 2017. – No. 30.
- [Brockman et al., 2016] Brockman G., Cheung V., Pettersson L., Schneider J., Schulman J., Tang J., and Zaremba W. Openai gym // *CoRR*. – 2016. – No. abs/1606.01540.
- [Colas et al., 2022] Colas C., Karch T., Sigaud O. and Oudeyer P.-Y. Autotelic agents with intrinsically motivated goal-conditioned reinforcement learning: a short survey // *Journal of Artificial Intelligence Research*. – 2022. – No. 74.
- [Eysenbach et al., 2018] Eysenbach B., Gupta A., Ibarz J., and Levine S. Diversity is all you need: Learning skills without a reward function // *CoRR*. – 2018. – No. abs/1802.06070.
- [Hafner et al., 2022] Hafner D., Lee K.-H., Fischer I. and Abbeel P. Deep Hierarchical Planning from Pixels // *Advances in Neural Information Processing Systems*. – 2022. – No. 35.
- [Jung et al., 2011] Jung T., Polani D. and Stone P. Empowerment for continuous agent environment systems // *Adaptive Behavior*. – 2011. – No. 19 (1).
- [Klyubin et al., 2005] Klyubin A. S., Polani D. and Nehaniv C. L. All else being equal be empowered // *Advances in Artificial Life. ECAL*. – Berlin: Springer-Heidelberg, 2005. – Lecture Notes in Computer Science, 3630.
- [Mendonca et al., 2021] Mendonca R., Rybkin O., Daniilidis K., Hafner D., and Pathak D. Discovering and achieving goals via world models // *Advances in Neural Information Processing Systems*. – 2021. – No. 34.
- [Raffin et al., 2021] Raffin A., Hill A., Gleave A., Kanervisto A., Ernestus M. and Dormann N. Stable-Baselines3: Reliable Reinforcement Learning Implementations // *Journal of Machine Learning Research*. – 2021. – No. 22 (268).
- [Ryan et al., 2000] Ryan R. M. and Deci E. L. Intrinsic and extrinsic motivations: Classic definitions and new directions // *Contemporary educational psychology*. – 2000. – No. 25 (1).
- [Salge et al., 2014] Salge C., Glackin C. and Polani D. Empowerment—an introduction // *Guided Self-Organization: Inception*. – Berlin: Springer-Heidelberg, 2014.
- [Schulman et al., 2017] Schulman J., Wolski F., Dhariwal P., Radford A. and Klimov O. Proximal policy optimization algorithms // *CoRR*. – 2017. – No. abs/1707.06347.
- [Sharma et al., 2020] Sharma A., Gu S., Levine S., Kumar V. and Hausman, K. Dynamics-aware unsupervised discovery of skills // *International Conference on Learning Representations*, Addis Ababa, Ethiopia. 2020.
- [Todorov et al., 2012] Todorov E., Erez T. and Tassa Y. Mujoco: A physics engine for model-based control. // *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. – 2012.
- [Zadaianchuk et al., 2022] Zadaianchuk A., Martius G. and Yang, F. Self-supervised reinforcement learning with independently controllable subgoals // *Conference on Robot Learning*. Auckland, NZ. 2022.

УДК 004.896

СИНТЕЗ ЭВРИСТИЧЕСКИХ РЕШЕНИЙ В РАМКАХ АДАПТИВНОГО УПРАВЛЕНИЯ МОБИЛЬНЫМ РОБОТОМ В ДИНАМИЧЕСКОЙ СРЕДЕ*

М.В. Макаров (*nauka-murom@yandex.ru*)

И.А. Семенов (*79601712352@yandex.ru*)

Н.С. Пиджакова (*trantinanatalya@yandex.ru*)

Муромский институт Владимирского государственного
университета, Муром

В работе представлены теоретические сведения, раскрывающие сущность явления синтеза эвристического решения нового типа в рамках адаптивного управления мобильным роботом. Описано экспериментальное исследование, обосновывающее техническую реализуемость подобного процесса, а также указывающее на то, что его осуществление приводит к повышению адаптационных способностей робота в динамической среде существования.

Ключевые слова: искусственный интеллект, мобильные роботы, адаптивное управление, принятие решения.

Введение

Автономное функционирование мобильного робота (МР) в своей естественной среде эксплуатации связано с необходимостью принятия решений в условиях неопределенности, вызванной неполнотой получаемой информации, её недостоверностью или содержащимися в ней противоречиями. Применение строгих алгоритмических методов или существующих инструментов интеллектуальной обработки информации на основе регрессии множественного прецедентного опыта не приводит к эффективному осуществлению подобного процесса или принципиально невозможно. Однако решение данной проблемы может быть результатом формирования интеллектуальных свойств МР, которые являются следствием инкорпорации особых квази-когнитивных механизмов понятийного или абстрактного мышления [Мелехин и др., 2022], способных формировать

* Исследование выполнено за счет гранта Российского научного фонда № 22-21-20111, <https://rscf.ru/project/22-21-20111/>.

собственную рациональную стратегию поведения, в результате чего становится возможным перестраивать внутри системы управления функциональные зависимости операций исходя из особой оценки ситуации.

В связи с этим следует пересмотреть подход к организации процесса обработки информации и принятия решения так чтобы результатом интеграции принципов понятийного или абстрактного мышления в систему управления МР стало возникновение феномена проявления поведенческих функций (ПФ). Иначе говоря, должно происходить установление ассоциативной связи квази-когнитивного агента с внешней действительностью, формирующей внутри системы инструменты решения локальных задач, не входящих в генеральную совокупность задач, существующих для объекта в момент его обучения.

Исходя из этого, очевидна потребность в проведении исследования, результаты которых могут стать основой для перехода к управлению, включающему в себя синтез сложных эвристических решений [Макаров и др., 2022] и, как следствие, автономную адаптацию поведения МР под изменение внешних условий динамической среды его существования.

1. Теоретические аспекты исследования

Существуют различные представления о механизмах реализации ПФ у МР, где лидирующие позиции занимают биоинспирированные модели поведения интеллектуальных технических объектов [Мальшев и др., 2020]. Однако в последнее время стала выраженной тенденция ухода от чистого биоинспирированного подхода к созданию более абстрактных моделей поведения [Карпов, 2022]. Универсальным общим подходом к технической реализации ПФ является выработка условных рефлексов внутри компонентов принятия решений МР. В качестве примера можно выделить работу [Корсаков и др., 2021], где использование спайковых нейронных сетей обеспечивает установление ассоциативных связей между отдельными явлениями и возможными видами активностей МР.

В общем виде ПФ можно выразить через рациональную реакцию МР на импульсы, побуждающие его к осуществлению физических исполнительно-двигательных активностей в рамках своих компетенций. В большинстве случаев эти активности связаны с перемещением МР в оперативном пространстве или реализации иных механических действий исполнительными органами его сочленений.

Использование стандартных нейросетевых архитектур для поддержки адаптивных средств обработки информации имеет место в существующей научной практике. В качестве примера можно привести использование рекуррентных нейронных конструкций прогнозирования событий для роботов с непрерывным обучением [Осипов и др., 2020]. Однако для осуществле-

ния требующегося процесса синтеза эвристических решений нового типа они не состоятельны. Это связано с тем, что искусственные нейронные сети не обладают структурной пластичностью и не могут в достаточной мере обеспечить формирование внутреннего контекста для формирования стратегии рационального поведения. Таким образом, требуется организация дополнительного информационного субстрата внутри вычислительных архитектур систем управления МР.

Для этого может использоваться метод, основанный на идее особых принципов активации нейронов. В привычном случае активация нейронов это нелинейная функция, которая генерирует соответствующий входу выход. В таком случае мы получаем скалярное значение эквивалентное активации нейрона. В нашем случае предлагается активацией считать именно функцию. Таким образом, получаем модификацию полносвязной нейросетевой ячейки и специальной функции активации нейронов. На выходе данного слоя ожидается нечёткое высказывание, интерпретируемое специальной системой нечеткого вывода. Под высказыванием подразумевается выбранный вариант следующего шага системы с обоснованием в виде одной из конечного множества стратегий: обход препятствия слева, обход препятствия справа, ожидание и разведка, возврат и т.п.

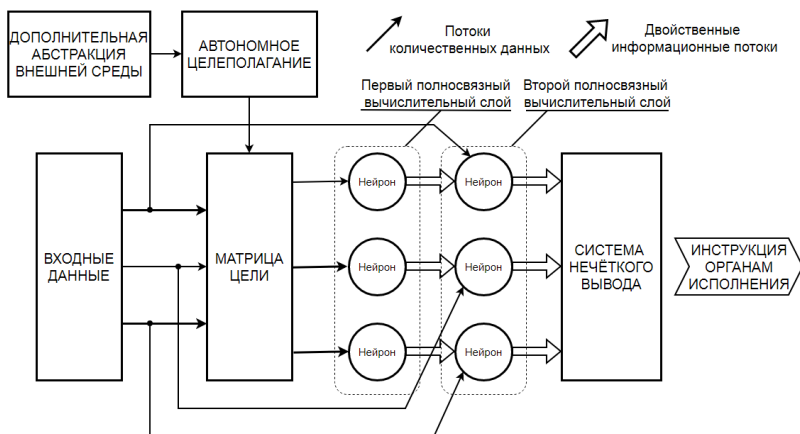


Рис. 1. Структурная схема процесса обработки информации и синтеза эвристического решения нового типа

На рис. 1 представлена ячейка с подобной структурной организацией, которая предполагает как минимум два вычислительных слоя. Первый слой получает используемую для принятия решения входную информацию, но не в чистом виде, а на основе производимых матрицей целеполагания пре-

образований. Результатом активации первого слоя является не вектор скалярных значений степени активации каждого нейрона, а множество функций. Эти функции будут использоваться в качестве активационных для нейронов последующего слоя. Аналогичная количественная форма входной информации подается и на второй вычислительный слой, где происходит другой этап преобразования образной информации для формирования нового пространства признаков. Такая архитектура ячейки позволит воссоздать структурную вариативность и способствует устранению строгой зависимости синтезируемого решения от входных сенсорных данных.

Отличительной особенностью процесса активации будет тот факт, что при использовании привычных базовых нелинейных функций предполагается, что значение выхода нейрона является количественным эквивалентом степени его участия в общем результате решения задачи. Такой подход позволяет описывать степень участия в количественном измерении, но не в качественном. Для демонстрации качественных характеристик процесса обработки информации активацией нейрона должно быть не единичное скалярное значение, а функция. Это способствует формированию внутреннего состояния системы, которое включает в себя отражение внешних условий решения задачи. Такое состояние оказывает действие (способ описания процесса) на активацию, что в свою очередь и определяет контекст синтезируемых решений.

2. Методология исследования

Методология проведенного экспериментального исследования была направлена на обоснование реализуемости процесса синтеза эвристических решений нового типа, а также на оценку влияния данного механизма обработки информации на формирование адаптационных способностей системы управления МР при его функционировании в условиях динамической среды существования.

В качестве объекта исследования использовалась компьютерная модель абстрактного разведывательного МР. С целью её создания была задействована виртуальная среда моделирования роботов Robot Operating System (ROS) в сочетании с динамическим 3D симулятором с открытым исходным кодом Gazebo и языком программирования Python. Предусмотрено воссоздание виртуальных сред существования модели за счёт использования находящихся в открытом доступе готовых ROS-моделей замкнутых пространств внутри различных помещений. Сбор сенсорной информации для осуществления реалистичного поведения модели при взаимодействии со статическими и динамическими компонентами внешней среды выполнялся средствами лидарного сканирования и визуальной одометрии.

Проведен ряд экспериментов внутри 20 различных по конфигурации виртуальных пространств, где в каждом случае симулировалось перемещение модели МР в директивно задаваемую точку. Предусмотрено, что в первую очередь будет осуществлен сбор информации о поведении объекта исследования при решении задачи навигации одним из существующих алгоритмических методов с элементами интеллектуальной обработки информации. Для этого использовался наиболее эффективный алгоритм исполнения процедуры локализации и картографирования TinySLAM с интеллектуальным построением динамической сетки занятости DOG, как это было выполнено в работе [Макаров и др., 2022]. В результате получено 20 траекторий, в которых для сравнительного анализа использовались по 200 шагов. Анализ состоял в оценивании каждого принятого решения о дальнейшем шаге модели МР. Это стало возможно благодаря переходу к бинарной классификации принятых решений: верно принятое решение, то есть способствующее получению оптимальной с точки зрения затрачиваемого времени на перемещение траектории при учёте ограничений, накладываемых динамическими и статическими объектами, а также неверно принятое решение, которое вызывает обратный эффект. Количественная оценка верности или неверности принятых решений в контексте затрачиваемого времени на перемещение модели МР осуществлялась с помощью рекурсивных методов динамического программирования.

Затем в систему были инкорпорированы предлагаемые механизмы обработки информации и принятия решения с последующим прохождением аналогичных маршрутов. Вследствие этого появилась возможность сравнить полученные траектории и оценить изменения принимаемых решений объектом исследования, а также как эти изменения повлияли на время достижения цели.

Кроме того, эксперименты повторялись в различных условиях, характеризующихся уровнем динамических свойств оперативного пространства объекта исследования. Следовательно, можно выделить несколько этапов, на первом из которых модель функционировала без взаимодействия с динамическими объектами, а далее в рамках используемых оперативных пространств последовательно осуществлялось добавление четырех типов динамических дестабилизирующих воздействий: появляющихся и исчезающих объектов, равномерно движущихся объектов, движущихся объектов с ускорением, искривляющих пространство объектов.

3. Результаты исследования и их обсуждение

Для демонстрации полученных результатов исследования произведена количественная оценка эффекта инкорпорации предлагаемых механизмов обработки информации и принятия решения в состав разработанной ком-

пьютерной модели МР. Симулировалась её работа, в ходе которой на вход системы управления подавались результаты исполнения процедуры локализации и картографирования (динамическая сетка занятости), а выдавался вектор, указывающий на одну из возможных активностей для текущего шага: движение вперед согласно принятой траектории, отклонение от траектории принятой вправо, отклонение от принятой траектории влево, остановка и ожидание или разворот.

Первым критерием, интерпретирующим результаты проведенного исследования, является установление факта отсутствия полной зависимости принимаемых решений от вектора входной информации. Усовершенствованная система управления меняла решения при аналогичных сенсорных воздействиях, что указывает на формирование внутреннего субстрата необходимого для дополнительной оценки текущей ситуации исходя из воздействия динамических дестабилизирующих факторов.

Согласно реализуемой методологии по всем проведенным экспериментам собирались статистические данные, представляющие собой сравнительные характеристики принимаемых решений исходной модели МР и усовершенствованной. В результате получена представленная на рис. 2 ROC-диаграмма, показывающая отношение доли верно измененных решений от их общего числа (True Positive Rate) к неверно измененным решениям от их общего числа (False Positive Rate).

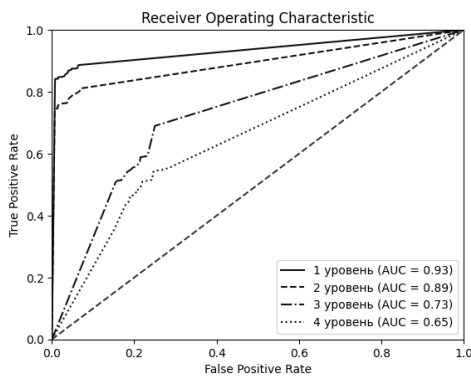


Рис. 2. Количественные результаты экспериментально исследования

Из ROC-диаграммы можно увидеть, что при различных по дестабилизирующему воздействию четырех уровнях динамических свойств оперативного пространства объекта исследования присутствовала существенная доля верных изменений решений вызванных инкорпорацией механизмов синтеза эвристических решений нового типа. Можно констатировать,

что на всех уровнях удалось добиться результатов, позволяющих говорить о существенном повышении адаптационных способностей модели МР. Данный процесс стабилен и не обладает аномалиями, а также стационарен для всех 20 помещений, в которых осуществлялись эксперименты.

Заключение

Полученные результаты позволяют сделать вывод, что существует возможность формирования механизмов обработки информации, которые могли бы использоваться в качестве субстрата для синтеза нового типа эвристических решений внутри системы управления МР в динамической среде существования. В рассматриваемом случае внесение структурных изменений в компонент принятия решения привело к возникновению отклика в адаптационных способностях объекта исследования при дестабилизации внешних условий. Подобный информационный субстрат для нового способа описания внешней среды привел к расширению поведенческого потенциала МР.

Список литературы

- [Мелехин и др., 2020] Мелехин В.Б., Хачумов М.В. Виды мышления автономных интеллектуальных мобильных систем и особенности их организации // Материалы 15-й мультиконференции по проблемам управления. – 2022.
- [Макаров и др., 2022] Макаров М.В., Семенов И.А., Демидов А.А., Грантина Н.С. Исследование нового типа эвристических решений для адаптивного управления мобильным роботом в динамической среде // Сборник научных трудов XXIV Международной научно-технической конференции «Нейроинформатика-2022», Россия. 2022.
- [Мальшев и др., 2020] Мальшев А.А., Бугров Е.В. К вопросу о параметрах биоинспирированных моделей поведения при моделировании групповой фуражировки // Труды СПИИРАН. – 2020. – № 19 (1).
- [Карпов, 2022] Карпов В.Э. Биоинспирированные модели поведения. Системный кризис или новые горизонты? // Сборник научных трудов XI Международной научно-практической конференции «Интегрированные модели и мягкие вычисления в искусственном интеллекте (ИММВ-2022)», Россия. 2022.
- [Корсаков и др., 2021] Корсаков А.М., Бахшиев Л.А., Астапова Л.А., Станкевич Л.А. Реализация поведенческих функций на спайковых нейронных сетях // Информатика и автоматизация. – 2021. – № 20 (3).
- [Осипов и др., 2020] Осипов В.Ю., Милосердов Д.И. Нейросетевое прогнозирование событий для роботов с непрерывным обучением // Информационно-управляющие системы. – 2020. – № 5 (108).
- [Макаров и др., 2022] Макаров М.В., Астафьев А.В., Семенов И.А. Исследование интеллектуальных элементов управления мобильным роботом и обеспечение информационной безопасности процесса его функционирования в динамической среде // Известия Юго-Западного государственного университета. – 2022. – № 2 (26).

МНОГОАГЕНТНАЯ ТЕХНОЛОГИЯ ОРГАНИЗАЦИИ ПРОИЗВОДСТВЕННЫХ И БИЗНЕС-ПРОЦЕССОВ ЦИФРОВЫХ ПРЕДПРИЯТИЙ*

Ю.Ф. Тельнов (*Telnov.YUF@rea.ru*)

В.А. Казаков (*Kazakov.VA@rea.ru*)

А.В. Данилов (*Danilov.AV@rea.ru*)

Российский экономический университет им. Г.В. Плеханова, Москва

Применение цифровых платформ индустрии четвертого поколения позволяет интегрировать производственные и бизнес-процессы цифровых предприятий. Для повышения гибкости формирования цепочек создания стоимости взаимодействующих предприятий на основе цифровых платформ предлагается применение многоагентной технологии. Для организации многоагентной производственной системы, обеспечивающей реализацию компонентной структуры продукции и услуг цифрового предприятия, обосновывается набор микросервисов в активной части административных оболочек активов. В статье предлагается ролевая модель взаимодействия агентов (административных оболочек) цифрового предприятия на основе онтологии предметной области и бизнес-модели, определяемой в зависимости от стадии жизненного цикла продукции и услуг.

Ключевые слова: цифровая платформа, цифровой двойник, компоненты цифрового предприятия, административная оболочка актива, многоагентная система, производственные и бизнес-процессы, цепочка создания стоимости

Введение

Разработка вариантов и механизмов цифровой трансформации производственных систем является предметом исследования многих национальных проектов и стратегических инициатив, в том числе проекта создания индустрии 4-го поколения [RAMI, 2018], концепции Промышленного интернета от ИС [PIRA, 2022], проектов создания умного производ-

* Работа выполнена при финансовой поддержке РФФ (проект № 22-11-00282 «Методы и модели цифровой трансформации предприятий на основе интеллектуальных технологий»).

ства [NIST, 2021], проекта ТехНет НТИ «Фабрики будущего» [Технет, 2017], [Технет, 2021] и др. Цифровая трансформация промышленного производства в настоящее время осуществляется по направлению к децентрализации принятия управленческих решений и динамическому построению гибких цепочек создания стоимости, включающих производственные и бизнес-процессы, что позволяет ускорить выпуск и повысить качество продукции и услуг.

Основой цифровой трансформации предприятий является использование цифровых платформ, двойников, потоков, базирующихся в свою очередь на применении технологий промышленного интернета, облачных вычислений, обработки больших объемов данных и искусственного интеллекта. Цифровые платформы обеспечивают высокий уровень кооперации и взаимодействия предприятий, образующих бизнес-экосистемы, в рамках динамически создаваемых цепочек поставок и производства [Цифровая трансформация, 2021], [Тельнов и др., 2022а]. Цифровые двойники, под которыми понимается динамические модели различных типов активов: продукции/услуг и их составных частей, оборудования, производственных линий, цехов и фабрик, осуществляют непрерывный сбор информации об активах на протяжении всего жизненного цикла функционирования для принятия своевременных решений на различных уровнях управления [Минаев и др., 2019], [Технет, 2019]. Цифровые потоки (нити) связывают цифровые двойники в управляемые производственные и бизнес-цепочки в соответствии с конкретными потребностями потребителей (заказчиков) [Вајај et al., 2018]. Цифровые двойники обслуживаются цифровой платформой, которая обеспечивает формирование цифровых потоков [Тельнов и др., 2022б].

Современные исследования показывают, что затраты на создание и эксплуатацию цифрового двойника непрозрачны и, как правило, оцениваются как слишком высокие. Отсюда вытекает, что технология цифровых двойников еще не нашла широкого применения, особенно на малых и средних предприятиях [Абрамов и др., 2022].

Для повышения эффективности применения и снижения затрат на создание и эксплуатацию цифровых двойников в настоящее время ведутся разработки в таких проектах, как создание и использование административных оболочек активов (проект цифровой платформы индустрии четвертого поколения), развитие и применение языка определения цифровых двойников (DTDЛ – Microsoft – язык описания моделей для устройств IoT Plug and Play, цифровых двойников устройств и логических цифровых двойников), разработка цифровых двойников с использованием эталонной модели облачных киберфизических систем (C2PS) [Vogel-Heuser et al., 2021]. Среди перечисленных подходов наиболее развитым подходом яв-

ляется концепция применения административных оболочек активов (административных оболочек) [BMWi, 2016], которая закреплена в ГОСТ Р 59799-2021¹ и будет рассматриваться нами в данной работе.

В общем виде все подходы связаны с созданием в качестве цифрового двойника динамической модели физических объектов или производственных систем в течение всего жизненного цикла [Минаев и др., 2019], [Технет, 2019]. Цифровые двойники позволяют не только отображать в информационном плане состояние активов, но также и осуществлять с помощью сервисов реагирование на возникающие события в работе производственной системы или проводить моделирование поведения активов на перспективу с помощью инструментария прогнозирования на реальных данных или данных имитационного моделирования. Общим недостатком многих промышленных технологий создания цифровых двойников является неспособность к самообучению и самоорганизации для оптимизации производственных и бизнес-процессов по ходу функционирования производственной системы и отсутствие развитого механизма принятия решений для автономной работы цифровых двойников.

Более ранним подходом к построению производственных систем по сравнению с цифровыми двойниками является подход на основе применения многоагентных технологий [Wooldridge, 2009], [Городецкий и др., 2017], который подразумевает создание автономных программных агентов для отдельных активов, осуществляющих по сути функции цифровых двойников, но благодаря наличию интеллектуального механизма принятия решений и реализации взаимодействий агентов обладает способностью динамического формирования цифровых потоков (нитей). В [Городецкий и др., 2017] приводится аналитический обзор многоагентных технологий и их применения, в том числе в производственной и бизнес-сфере. Авторы отмечают сложности в формализации концептуальной модели многоагентной системы, описании языка общения агентов и построении архитектуры программно-коммуникационной среды, в рамках которой взаимодействуют агенты, что не позволило в свое время создать работоспособные программные среды для реализации МАС и уступили место более простым сервисно-ориентированным и микросервисным архитектурам.

В настоящее время интерес к многоагентным системам вернулся с позиции возможности интеграции интеллектуальной многоагентной технологии и сервисной реализации административных оболочек активов. Наиболее сложные аспекты построения многоагентных систем связаны с организаци-

¹ ГОСТ Р 59799-2021 Национальный стандарт Российской Федерации. Умное производство. Модель эталонной архитектуры Индустрии 4.0 (RAMI 4.0) (IEC PAS 63088:2017, NEQ). – Москва, Российский институт стандартизации, 2021.

ей активного и проактивного взаимодействия компонентов производственных цепочек, реализуемых с помощью многоагентной технологии. В этом случае централизованное управление выполнением производственных и бизнес-процессов заменяется на распределенное принятие решений агентами, что позволяет производственной системе автономно адаптироваться к изменяющимся условиям и задачам [Vogel-Heuser et al., 2021].

В статье предлагается подход к развитию ролевой модели взаимодействия программных агентов, который определяет построение эффективных механизмов поиска решений в производственных системах, построенных на основе цифровых платформ.

1. Организация многоагентного взаимодействия на основе цифровых двойников - административных оболочек активов

В современных условиях значительная часть разработок административных оболочек активов в большей степени сосредоточена на информационном отображении состояния активов на протяжении всего жизненного цикла и не связана с аспектами принятия решений и совместной работы активов в производственных или бизнес-процессах. В этом случае административная оболочка актива (AS – Asset Administration Shell) представляет собой файл цифровых данных в одном из широко распространенных форматов XML или JSON, AS вместе с активом составляют i4.0-компонент (компонент) [BMWi, 2016]. Обмен информацией между компонентами осуществляется опосредовано путем пересылки файлов о состоянии активов между участниками производственных или бизнес-процессов, отвечающих за выполнение операций с этими активами с позиции различных ролей.

Многоагентная реализация взаимодействия компонентов предполагает применение активного и проактивного режимов использования административных оболочек [Komesker et al., 2022], [Sakurada et al., 2022]. Активный режим реализует доступ к административным оболочкам через API интерфейс, а проактивный режим осуществляет автоматическое взаимодействие между административными оболочками с помощью обмена сообщениями на языке I4.0 Language (стандарт взаимодействия VDI/VDE 2193) [Belyaev et al., 2019], основанного на спецификациях языка коммуникации агентов ACL FIPA.

Наиболее простой способ реализации многоагентных систем заключается во включении набора микросервисов, обеспечивающих принятие решений, в форме реестра используемых сервисов непосредственно в состав административной оболочки [Sakurada et al., 2022], [Komesker et al., 2022]. В качестве обязательных микросервисов выделяются микросервисы, обеспечивающие взаимодействие компонентов в производственных и

бизнес-процессах, и микросервисы, обеспечивающие интеграцию административных оболочек с физическими активами. В качестве микросервисов, составляющих основу агентской организации оболочек, включаются микросервисы сбора данных о работе актива, мониторинга и диагностики состояний, принятия решений и их исполнения. Микросервисы сбора данных предполагают организацию и ведение базы данных. Микросервисы мониторинга и диагностики состояний могут быть основаны на реализации обработки больших объемов данных. Микросервисы принятия оперативных решений по устранению отклонений в процессах могут базироваться на базе производционных правил. Микросервисы принятия тактических решений могут вызывать подключение программных модулей MES и ERP систем для планирования действий на определенном временном горизонте. Микросервисы принятия стратегических решений о реорганизации производственных и бизнес-процессов для выпуска новой или модифицированной продукции или оказания услуг могут использовать более сложные математические и эвристические модели.

В структуре административных оболочек активов с позиции организации доступа к данным и взаимодействия между ними с использованием языка I4.0 Language [Belyaev et al., 2019] выделяются пассивная часть, связанная с отображением свойств актива в каждый момент времени, и активная часть, связанная с набором микросервисов, обрабатывающих состояние актива и управляющая его поведением и взаимодействием с другими компонентами системы.

В структуре административной оболочки актива интерес представляет выделение двух режимов обращения к компоненту, которые могут исполнять агенты в форме этих оболочек: либо запрос выполнения микросервиса, либо предоставление микросервиса на исполнение. Эти режимы соответствуют двум основным ролям, которые могут исполнять программные агенты на предприятиях: создание продукции (агент продукта) и производство (обеспечение) ресурсов (агент ресурса) [Vogel-Heuser et al., 2021], [Belyaev et al., 2019].

Механизмы взаимодействия агентов системы посредством асинхронного обмена языковыми сообщениями (коммуникативные акты) и их обработки на основе предоставления сервисов хорошо специфицированы в стандартах FIPA (Foundation for Intelligent Physical Agents) [FIPA, 2002]. На базе этих спецификаций были разработаны ряд программных платформ, среди которых наиболее известной является платформа JADE [Bellifemine et al., 2007], [JADE, 2018]. В настоящее время разработана платформа на языке Python – PADE [PADE, 2023], [Melo et al., 2019]. При построении архитектуры многоагентной производственной системы с использованием административных оболочек предлагается использовать

платформу, основанную на спецификациях FIPA, как достаточно хорошо апробированную в практических приложениях [Salazar et al., 2019], [Salazar et al., 2022], которые демонстрируют свою эффективность с позиции достижения критериев автономности, реактивности, проактивности, прогнозирования и человеко-машинного взаимодействия. При этом в реализации микросервисов принятия решения предусмотрено включение механизма правил обработки динамических событий и процедур многокритериальной оценки на основе нечеткой логики.

2. Ролевая модель многоагентного взаимодействия компонентов цифрового предприятия

Решение задачи многоагентного взаимодействия компонентов производственной системы является простым в случае, когда изготовление продукции осуществляется в рамках одного цифрового предприятия, на котором спецификации продукции понятны всем исполнителям и все участники производственного процесса находятся в отношении строгого подчинения. В этом случае используемые агенты имеют реактивный характер, то есть все сообщения между агентами однозначно понимаются и однозначно исполняются.

В случае создания цифровых сетевых предприятий (сетевых предприятий) с переменным участием различных исполнителей возникают потребности в создании интеллектуальных агентов, способных к интерпретации запросов, их оценке и поиску наилучших решений. При этом сетевое предприятие поддерживается множеством агентов, каждый из которых отвечает за производство отдельного компонента или всего продукта (услуги). Для формирования сетевого предприятия используется облачная инфраструктура, в свою очередь поддерживаемая агентом облачной цифровой платформы. Структура многоагентной системы цифрового сетевого предприятия представлена на рис. 1.

Каждый агент включает в свой состав набор функциональных сервисов, который определяется моделью организации производственных и бизнес-процессов в зависимости от стадии жизненного цикла продукции и типа реализуемой бизнес-модели. Так для производственных и бизнес-процессов на стадии разработки конструкции продукции и планирования производства один функциональный сервис использует модель интеграции участников цепочки создания стоимости (Один головной агент-продукта – N агентов-ресурсов), а на стадии собственно производства и сопровождения другой функциональный сервис использует модель сервисного обслуживания потребителей продуктов и услуг (Один агент-продукции – Один агента-ресурса) [Тельнов и др., 2022a]. Агент платформы осуществляет реализацию функций инфраструктуры бизнес-экосистемы, в которой формируются цифровые сетевые предприятия [Тельнов и др., 2022б].

Каждый функциональный сервис как композитный сервис в своем составе может содержать, как отмечалось выше, микросервисы взаимодействия с другими компонентами, интеграции с физическими объектами, сбора данных о работе актива, мониторинга и диагностики состояний, принятия решений и их исполнения.

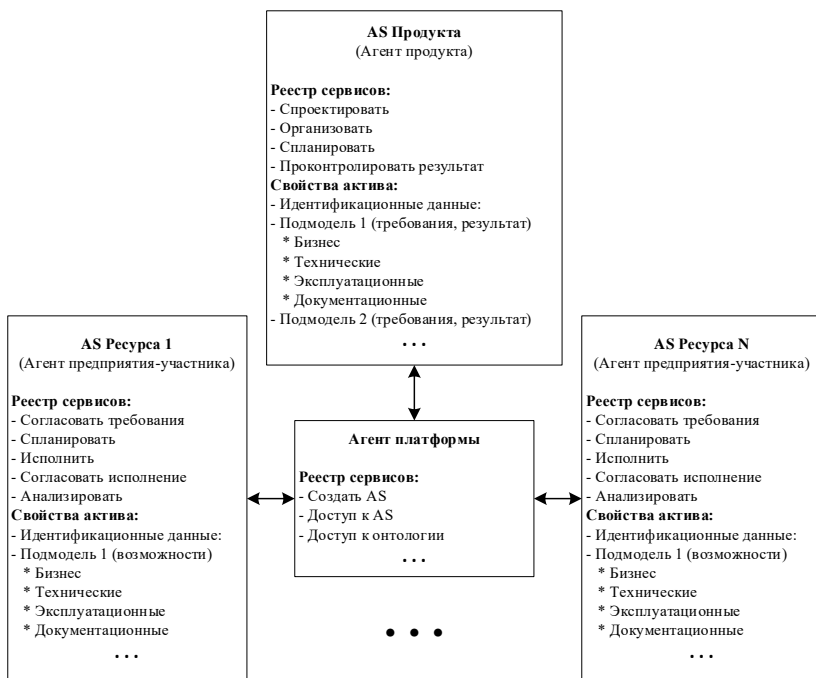


Рис. 1. Структура многоагентной системы цифрового сетевого предприятия

Наиболее сложными с интеллектуальной точки зрения микросервисами являются микросервисы взаимодействия между компонентами (агентами) и принятия решений, которые вызывают необходимость решения двух следующих задач:

1. Понимание сообщений, посылаемых от одного агента к другому с учетом возможной гетерогенности многоагентной системы, когда каждый агент является по сути цифровым двойником предприятия, а предприятия, производящие компоненты, используют специфические знания об изготовляемом продукте.

2. Решение задачи выбора наилучшего поставщика ресурса для изготовления изделия со стороны производителя продукта и соответственно решение задачи оценки возможности выполнения поставленной задачи со стороны предприятия, создателя ресурса.

Процесс формирования запроса передающим агентом, его кодирования по общей онтологии, передачи фрагмента онтологии в качестве объекта запроса, декодирование и преобразования в формат принимающего агента на платформе JADE описан в [Тельнов и др., 2018]. Определенные трудности в реализации описанного процесса могут возникать, когда онтологии передающего и принимающего сообщения агентов полностью не совпадают, как правило, в процессах проектирования или совершенствования конструкции продукции. В этом случае должен быть организован переговорный процесс между агентами по поводу содержания онтологии с подключением экспертов для выравнивания используемой терминологии, в результате которого исходная онтология может быть расширена или уточнена [Тельнов и др., 2022б]. Для упрощения преобразования свойств и установления соотношений между запросами и возможностями разных административных оболочек возможным подходом может быть онтологическое представление пассивной части оболочек в RDF или OWL формате, что требует проведения дополнительного исследования.

Основное содержание базы знаний для работы микросервисов принятия решений должно соответствовать принципам модели BDI (believe, desire, intention) [Городецкий и др., 2017], что соответствует набору критериев оценки требований или решений; целевой функции, интегрирующей оценки критериев; набору действий, соответствующих намерениям достижения целевой функции. Для агентов-ресурсов база знаний нацелена на оценку возможностей реализации запросов агентов-продуктов, а для агентов-продуктов база знаний должна позволять генерацию требований к выпускаемой продукции и её составным частям, а также оценку предложений со стороны агентов-ресурсов.

Рассмотрим ролевую модель формирования цепочки создания стоимости, используемая на стадии определения требований и проектирования конструкции изделий, по которой формируется состав участников цифрового предприятия, а также при организации цепочек поставок, когда по некоторым причинам состав участников цепочки поставок должен формироваться по ходу выполнения процесса поставок (на лету). На рис. 2 представлена диаграмма последовательности, отражающая последовательность взаимодействия агента продукта и N агентов-ресурсов с использованием протоколов обмена сообщениями FIPA.

В рамках процесса подбора партнеров по созданию компонентов продукта в рамках общей цепочки создания стоимости на первом этапе агент продукта подбирает потенциальных партнеров среди агентов ресурсов, зарегистрированных на платформе. В качестве агентов ресурсов выступают цифровые двойники участников платформы, обладающих необходимыми производственными ресурсами для выполнения работ по созданию компонентов продукта. Агент-менеджер платформы (MAC-платформа i4.0) подбирает необходимых партнеров по запросу агента продукта на основе внутреннего каталога.

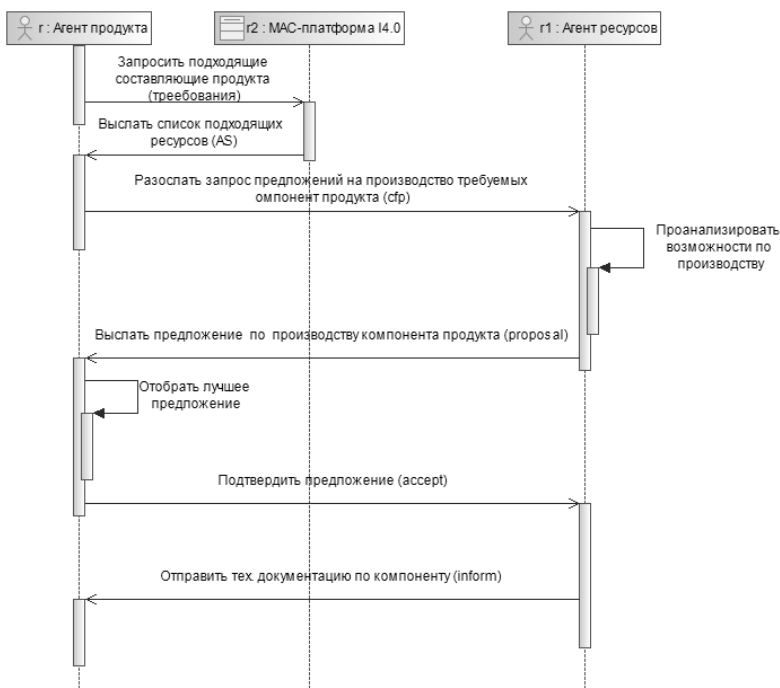


Рис. 2. Диаграмма последовательностей модели формирования цепочки создания стоимости

Далее агент продукта рассылает запрос предложений по производству необходимых компонентов продукта отобранным ранее агентам ресурсов. Агент ресурса, получивший предложение, на основе анализа имеющихся свободных ресурсов и их возможностей формирует предложение на производство компонента и отправляет его агенту продукта.

Агент продукта по истечении времени сбора предложений проводит качественный критериальный анализ поступивших предложений и выбирает наилучшего партнера для каждого компонента продукта. При необходимости может быть запрошена дополнительная информация о возможностях производства у агента ресурса.

Далее агент продукта высылает отобранным агентам ресурсов подтверждение их предложения в виде предварительного контракта – соглашения об уровне обслуживания (SLA) с конкретизированными параметрами производства. В ответ на подтверждение агент ресурса информирует агента продукта о согласии с SLA путем отсылки технической документации на производимые компоненты. В качестве формата обмена данными между агентами системы предлагается использовать формат aasx, являющийся стандартом для систем индустрии четвертого поколения.

Заключение

Проведенное исследование показало, что для организации производственных и бизнес-процессов цифровых предприятий с помощью технологий цифровых платформ и цифровых двойников наиболее рациональными методами является применение комбинации многоагентных и микросервисных архитектур в рамках расширения возможностей административных оболочек активов.

В статье предложена структура многоагентной системы цифрового сетевого предприятия, в которой взаимодействие агентов продуктов и агентов ресурсов осуществляется как напрямую, так и через агента-цифровой платформы с помощью микросервисов, обеспечивающих взаимодействие и доступ к свойствам компонентов. Для подготовки, формирования и оценки решений о траектории цепочки создания стоимости предлагается интеллектуальный микросервис, содержащий базу знаний и модель вывода, характер которых определяется требованиями композитного сервиса по реализации бизнес-модели. В статье разработана ролевая модель многоагентного взаимодействия компонентов цифрового предприятия, позволяющая интегрировать участников цепочки создания стоимости.

Исследование методов и средств многоагентной технологии для организации производственных и бизнес-процессов будет продолжено для различных типов бизнес-моделей.

Список литературы

[Абрамов и др., 2022] Абрамов В.И., Бобоев Д.С., Гильманов Т.Д., Семенов К.Ю. Теоретические и практические аспекты создания цифрового двойника компании // Вопросы инновационной экономики. – 2022. – Т. 12, № 2. – С. 967-980. – DOI: 10.18334/vinec.12.2.114890.

- [Городецкий и др., 2017] Городецкий В.И., Бухвалов О.Л., Скобелев П.О., Майоров И.В. Современное состояние и перспективы индустриальных применений многоагентных систем // Управление техническими системами и технологиями. Управление большими системами. – 2017. – № 66. – С. 94-157.
- [Минаев и др., 2019] Минаев В.А., Мазин А.В., Здирук К.Б., Куликов Л.С. Цифровые двойники объектов в решении задач управления // Радиопромышленность. – 2019. – Т. 29, № 3. – С. 68-78. – DOI: 10.21778/2413-9599-2019-29-3-68-78
- [Тельнов и др., 2018] Тельнов Ю.Ф., Данилов А.В., Дивеев Р.И., Казаков В.А., Ярошенко Е.В. Разработка прототипа многоагентной системы сетевого взаимодействия высших учебных заведений // Открытое образование. – 2018. – № 6. – С. 252-264.
- [Тельнов и др., 2022а] Тельнов Ю.Ф., Брызгалов А.А., Козырев П.А., Королева Д.С. Выбор типа бизнес-модели для реализации стратегии цифровой трансформации сетевого предприятия // Бизнес-информатика. – 2022. – № 4. – С. 50-67.
- [Тельнов и др., 2022б] Тельнов Ю.Ф., Казаков В.А., Данилов А.В., Денисов А.А. Требования к программной реализации системы Индустрии 4.0 для создания сетевых предприятий // Программные продукты и системы. – 2022. – Т. 35, № 4. – С. 557-571. – DOI: 10.15827/0236-235X.140.557-571.
- [Технет, 2017] Многоуровневая структура фабрик будущего // Ассоциация «Технет». Инфраструктурный центр «Технет» Фабрики будущего (technet-nti.ru). – URL: <https://technet-nti.ru/article/fabriki-buducshego> (дата обращения: 26.04.2023).
- [Технет, 2019] Цифровые двойники в высокотехнологичной промышленности. Экспертно-аналитический доклад // Инфраструктурный центр «Технет» НТИ. – М., 2019. [Электронный ресурс]. – URL: http://assets.fea.ru/uploads/fea/news/2019/12_december/28/cifrovoy_dvoinik.pdf (дата обращения: 26.04.2023).
- [Технет, 2021] План мероприятий («дорожная карта») «ТЕХНЕТ 4.0» (передовые производственные технологии) Национальной технологической инициативы. – URL: https://assets.fea.ru/uploads/fea/news/2021/06/24/2021_0624_DK_Technet.pdf (дата обращения: 26.04.2023).
- [Цифровая трансформация, 2021] Цифровая трансформация промышленных предприятий в условиях инновационной экономики: монография / под научной редакцией доктора экономических наук Веселовского М.Я. и кандидата экономических наук Хорошавиной Н.С. – М.: Мир науки, 2021. – 296 с. Сетевое издание. – URL: <https://izd-mn.com/PDF/06MNNPM21.pdf> (дата обращения 26.04.2023).
- [Bajaj et al., 2018] Bajaj M., Hedberg T. System lifecycle handler – Spinning a digital thread for manufacturing // 28th Annual INCOSE International Symposium. Washington, DC, 7–12 July 2018. – Vol. 28. – P. 1636-1650. – DOI: 10.1002/j.2334-5837.2018.00573.x.
- [Bellifemine et al., 2007] Fabio Bellifemine, Giovanni Caire, Dominic Greenwood Developing multi-agent systems with JADE – Wiley; 1 edition. – April 2, 2007.
- [Belyaev et al., 2019] Alexander Belyaev, Christian Diedrich. Specification "Demonstrator I4.0-Language" v3.0 // Technical Report IFAT-LIA 07/2019 09. July 2019.
- [BMW, 2016] Structure of the Administration Shell. Continuation of the Development of the Reference Model for the Industrie 4.0 Component, working paper Plattform Industrie 4.0, BMW, 2016.

- [**FIPA, 2002**] FIPA ACL Message Structure Specification, Foundation for Intelligent Physical Agents. – 2002. – URL: <http://www.fipa.org/specs/fipa00061/SC00061G.pdf> (дата обращения 28.07.2023).
- [**IIIRA, 2022**] The Industrial Internet Reference Architecture - Version 1.10. An Industry IoT Consortium Foundational Document. 2022-11-07. – URL: <https://www.iiconsortium.org/wp-content/uploads/sites/2/2022/11/IIIRA-v1.10.pdf> (дата обращения 28.07.2023).
- [**JADE, 2007**] JADE (Java Agent Development Network) [Электронный ресурс]. – URL: <https://jade.tilab.com> (дата обращения 28.07.2023).
- [**Komesker et al., 2022**] Simon Komesker, William Motsch, Jens Popper, Aleksandr Sidorenko, Achim Wagner, Martin Ruskowski Enabling a Multi-Agent System for Resilient Production Flow in Modular Production Systems // 55th CIRP Conference on Manufacturing Systems, Procedia CIRP 107 Published by Elsevier B.V. – 2022. – P. 991-998.
- [**Melo et al., 2019**] Melo L.S., Sampaio R.F., Leão R.P.S., Barroso G.C. and Bezerra J.R. Python-based multi-agent platform for application on power grids // International Transactions on Electrical Energy Systems. – 2019. – Vol. 29, No. 6. – P. 1-14.
- [**PADE, 2023**] Python Agent DEvelopment framework. Multi-agent Systems for Python Language! – URL: <https://pade.readthedocs.io/en/latest> (дата обращения 28.07.2023).
- [**RAMI, 2018**] Reference Architectural Model Industrie 4.0 (RAMI4.0) – An Introduction. – URL: <https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/rami40-an-introduction.htm> (дата обращения: 28.07.2023).
- [**Sakurada et al., 2022**] Lucas Sakurada, Paulo Leitao, Fernando De la Prieta. Agent-Based Asset Administration Shell Approach for Digitizing Industrial Assets // IFAC PapersOnLine. – 2022. – 55-2. – P. 193-198.
- [**Salazar et al., 2019**] Luis Alberto Cruz Salazar, Daria Ryashentseva, Arndt Lüder, Birgit Vogel-Heuser. Cyber-physical production systems architecture based on multi-agent's design pattern—comparison of selected approaches mapping four agent patterns // The International Journal of Advanced Manufacturing Technology. – 2019. – 105. – P. 4005-4034. – <https://doi.org/10.1007/s00170-019-03800-4> (дата обращения: 28.07.2023).
- [**Salazar et al., 2022**] Luis Alberto Cruz Salazar, Birgit Vogel-Heuser. A CPPS-architecture and workflow for bringing agent-based technologies as a form of artificial intelligence into practice // Automatisierungstechnik. – 2022. – 70 (6). – P. 580-598. – https://www.researchgate.net/publication/361152397_A_CPPS-architecture_and_workflow_for_bringing_agent-based_technologies_as_a_form_of_artificial_intelligence_into_practice (дата обращения: 28.07.2023).
- [**NIST, 2021**] Smart Manufacturing. Model-Based Enterprise Program. NIST, 2021. – URL: <https://www.nist.gov/programs-projects/model-based-enterprise-program> (дата обращения: 28.07.2023).
- [**Vogel-Heuser et al., 2021**] Birgit Vogel-Heuser, Felix Ocker, Tobias Scheuer. An approach for leveraging Digital Twins in agent-based production systems // Automatisierungs-technik, 2021 (дата обращения: 28.07.2023).
- [**Wooldridge, 2009**] Wooldridge M. An Introduction to MultiAgent Systems. – 2nd ed. – John Wiley and Sons, 2009.

АВТОМАТИЧЕСКАЯ РЕГИСТРАЦИЯ ПРИСУТСТВИЯ СТУДЕНТОВ НА УЧЕБНОМ ЗАНЯТИИ С ПОМОЩЬЮ КОМПЬЮТЕРНОГО ЗРЕНИЯ*

О.И. Федяев (*olegfedyayev@yahoo.com*)

И.А. Коломойцева (*bolatiger@mail.ru*)

Донецкий национальный технический университет, Донецк

Цель работы – устранить потери времени преподавателя, которые возникают при ручной регистрации присутствия большого количества учащихся в аудитории. Автоматическая регистрация осуществляется с помощью компьютерного зрения в режиме реального времени. Признаки лица человека формируются свёрточной нейронной сетью. Система контроля выполняет следующие функции: фиксирует лица студентов на видеокамеру при входе в аудиторию, сравнивает лица с базой данных студентов группы, отмечает присутствие на лекции (или опоздание) в случае успешной идентификации, сохраняет данные в электронном журнале. Систему можно рекомендовать преподавателям образовательных учреждений.

Ключевые слова: компьютерное зрение, регистрация студентов, свёрточная нейронная сеть.

Введение

В статье рассматривается научно-практическая проблема автоматического распознавания человека по его лицу, которая относится к области компьютерного зрения (Computer Vision – CV). В частности приводится постановка и решение конкретной задачи оперативного визуального контроля присутствия студентов на учебных занятиях с помощью компьютерного зрения. Актуальность данной задачи обусловлена, во-первых, потерями времени преподавателя, которые возникают при ручной регистрации присутствия большого количества учащихся в аудитории (в потоке из 70 человек потери на «переключку» составляют в среднем 85 мин в семестре и даже больше); во-вторых, «переключка» проходит шумно и не всегда достоверно.

* Данная работа выполнена по плану НИР ДонНТУ, тема № Н 2022-15.

Основные трудности компьютерного распознавания лиц в реальном времени связаны с быстрой изменчивостью изображений объектов в видеопотоке: положение, размер и ракурс лица в кадре, освещение и т.д. [Deep Learning, 2023].

В настоящее время большие перспективы в преодолении перечисленных проблем связывают с применением глубоких нейронных сетей. К этому классу относится многослойная свёрточная нейронная сеть (Convolutional Neural Network) [Гудфеллоу и др., 2018], которая является развитием идей таких архитектур нейронных сетей, как многослойные сети типа когнитрон и неокогнитрон.

На данный момент свёрточная нейронная сеть и её модификации считаются лучшими по точности и скорости распознавания объектов на изображении. Более того, распознавание лиц из видеопотока с камеры показывает, что нейронные сети с такой архитектурой способны работать в режиме реального времени даже на устройствах с ограниченными ресурсами.

В последнее время известными фирмами предложены мощные библиотеки, в которых реализованы различные модели глубоких нейронных сетей, позволяющие решать сложные задачи распознавания. Поэтому целью данной работы является оценка возможности реализации нейросетевого распознавания лиц из видеопотока на базе существующих инструментальных средств [Антонио и др., 2018], [Reading, 2023] и создание системы оперативного визуального контроля присутствия студентов на учебных занятиях, что подчёркивает её уникальность.

1. Процесс распознавания лиц на кадрах из видеопотока

Функциональная схема видеорегистрации студентов при входе в аудиторию с помощью компьютерного зрения показана на рис.1. Методами машинного обучения решаются задачи локализации лиц на снимках с видеокамеры и их распознавание, а также формирование электронного журнала группы.

Блок обнаружения лиц принимает изображения с веб-камеры в режиме реального времени, выделяет и локализует на них лица. Эту функцию выполняет алгоритм детектирования лиц на текущем кадре видеопотока. В результате формируется последовательность изображений лиц, захватываемых видеокамерой, для последующего их распознавания.

Каждое выделенное изображение лица передаётся в блок формирования вектора признаков, который реализует распознавательную функцию $f: X \rightarrow Y$, где X – множество входных изображений лиц; Y – множество векторов признаков для лиц из X . Таким образом, нейросетевая функция f каждому выделенному лицу $x \in X$ ставит в соответствие вектор признаков $y \in Y$ ($y=f(x)$), которыми характеризуется данное лицо.



Рис. 1. Основные процессы компьютерного распознавания лиц человека

Для выработки признаков лица применялась свёрточная нейронная сеть, которая предварительно была обучена на примерах фотографий 2622-х человек (по 1000 фотографий на человека) [Гудфеллоу и др., 2018], [Антонио и др., 2018]. Поэтому результатом работы сети является 2622-мерный вектор, каждый элемент которого представляет собой вероятность сходства лица с одним из обучающего множества. Считается, что два изображения лица относятся к одному человеку, если они в одинаковой мере похожи на лицо из обучающего множества. Для этого вектора признаков этих изображений в пространстве лиц из обучающего множества должны образовывать между собой достаточно острый угол.

При настройке системы необходимо предварительно сформировать базу данных лиц для всех распознаваемых людей, представленных конечным множеством соответствующих фамилий L . С этой целью для x с помощью нейросетевой распознавательной функции f определяется множество правильных пар

$$\{(y, l) \mid y = f(x), x \in \bar{X}, l \in L\},$$

где \bar{X} – множество подготовленных фотографий распознаваемых лиц, т.е. эталоны изображений распознаваемых лиц; y – вектор признаков изображения лица $x \in \bar{X}$; l – фамилия человека, фотография которого изображена на снимке x . Всё множество пар (y, l) заносится в базу данных векторов признаков лиц.

В штатном режиме работы системы, т.е. при распознавании, в блоке сравнения вектор признаков распознаваемого лица, полученный с выхода свёрточной нейронной сети, сравнивается со всеми векторами базы данных. Процедура сравнения основывается на методе вычисления косинусного сходства вектора распознаваемого лица с каждым вектором-эталоном из базы данных по следующей формуле

$$\text{Сходство} = \frac{Y \cdot \bar{Y}}{\|Y\| \cdot \|\bar{Y}\|} = \frac{\sum_{i=1}^n y_i \cdot \bar{y}_i}{\sqrt{\sum_{i=1}^n y_i^2} \cdot \sqrt{\sum_{i=1}^n \bar{y}_i^2}},$$

где Y и \bar{Y} – вектора признаков соответственно распознаваемого лица и лица-эталона из базы данных; $n = 2622$.

Расознаваемое лицо считается соответствующим эталону, если полученный коэффициент сходства выше определённого значения (в работе использовалось значение 0.7).

2. Алгоритм локализации изображения лица

Ввод видеопотока в систему распознавания осуществлялся с помощью программных средств библиотеки OpenCV [Reading, 2023]. Для решения задачи автоматической локализации лиц на кадрах видеопотока применялся метод Виола-Джонса [Paul et al., 2001], который рассматривает лицо как изображение, состоящее из набора пикселей. Этот метод базируется на разнице оттенков пикселей в прямоугольнике с двумя областями: тёмной и светлой. Каждая такая область – признак (или функция) Хаара. Признаки можно по-разному применять к анализу изображения. В настоящее время существует эффективный алгоритм бустинга, ориентированный на машинное обучение различных классификаторов Хаара выделять заданные объекты на изображениях [Paul et al., 2001].

Поэтому в работе были проанализированы четыре стандартных классификатора, построенных на методе Виола-Джонса и реализованных в библиотеке OpenCV, а именно: `default`, `alt`, `alt2` и `alt_tree`. Для оценки качества их работы были составлены тестирующие наборы, отражающие реальные ситуации при распознавании лиц.



Рис. 2. Влияние угла поворота лица на эффективность обнаружения

В первом эксперименте оценивалось влияние угла поворота лица на эффективность обнаружения. Для эксперимента взят один снимок лица, на котором легко (без влияния фона) выделяется лицо всеми четырьмя классификаторами. Далее были произведены повороты исходного изображения лица на 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 градусов. Полученные изображения были обработаны всеми классификаторами. Результат тестирования показаны на рис. 2.

Во втором эксперименте оценивалось влияние поворота головы вправо/влево на эффективность обнаружения лица. Для эксперимента взяты лица двух человек, сфотографированных при разных углах поворота головы: от -90 градусов (человек смотрит влево) до $+90$ градусов (человек смотрит вправо). На рис. 3 показаны примеры снимков и полученный результат тестирования.

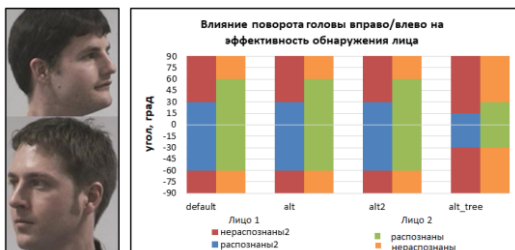


Рис. 3. Влияние поворота головы вправо/влево на эффективность обнаружения лица

Были аналогично проведены эксперименты по влиянию наклона головы (вверх/вниз) и наличие очков, на эффективность обнаружения лица разного типа. Во всех экспериментах результаты тестирования показали, что классификатор alt лучше справляется с задачей обнаружения лица. Поэтому этот каскадный классификатор использовался для обработки кадров из видеопотока. На выходе классификатор определяет местоположение выделенного лица на снимке набором параметров, включающим координаты и размеры прямоугольной рамки, которой ограничивается изображение лица человека.

Выделенное лицо на снимке, которое распознаётся, возможно будет более информативным для нейросети, если с помощью 3D-моделирования создать 3D-лицо и его подавать на вход нейросети. Эта интересная концепция может в определённой степени лучше решить проблему ориентации головы, но она будет проверена авторами позже.

На рис.4 показана регистрация лица, прямо смотрящего в камеру. С помощью меню этого окна можно формировать базу данных с информацией о фамилиях студентов $I \in L$ и фотографиях $x \in \bar{X}$. Все изображения выделенных лиц нормируются к стандартному размеру 224×224 пикселей, каждый из которых представлен в виде трёх составляющих в формате RGB.

Тестирование системы проводилось с использованием веб-камеры ASUS. Камера выдаёт изображение в формате VGA (640×480 , 18-битный RGB) с частотой 5 кадров в секунду.

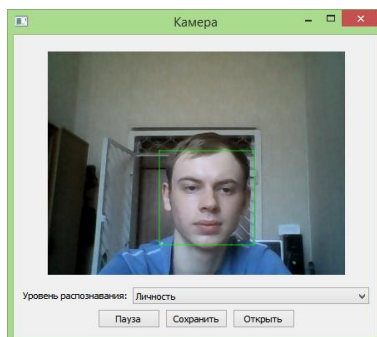


Рис. 4. Регистрация лица с видеокамеры

Угол обзора ориентировочно равен 30° . Нейросетевые алгоритмы исполнялись на процессоре Intel Core i7–5500U (2 физических ядра тактовой частотой 3,0 ГГц). Локализация лиц на изображении происходит за 15–45 мс, создание вектора признаков одного лица свёрточной нейросетью – за 550–590 мс. Реальная производительность обработки видеок кадров на данной конфигурации составляет приблизительно 1,7 кадров в секунду.

3. Архитектура свёрточной нейронной сети

Свёрточная нейронная сеть была разработана для эффективного распознавания различных объектов на изображении [Reading, 2023]. Её многослойная архитектура состоит из свёрточных слоёв (convolution layers) и субдискретизирующих слоёв (subsampling layers или pooling layers, слоёв подвыборки), которые чередуются друг с другом (рис. 5).

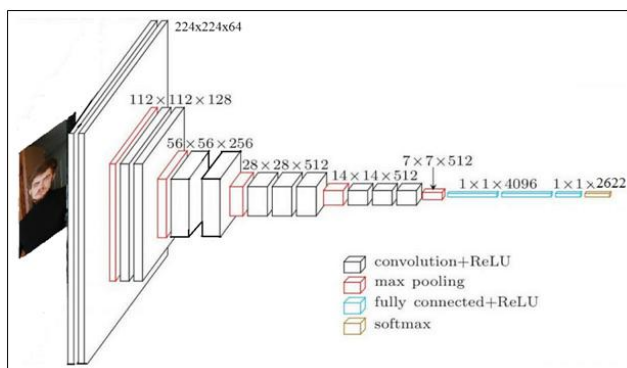


Рис. 5. Архитектура модели многослойной свёрточной нейронной сети VGGFace

В каждом слое имеется набор из нескольких плоскостей признаков. Причём нейроны одной плоскости имеют одинаковые веса, ведущие к соответствующим локальным участкам предыдущего слоя. Изображение предыдущего слоя как бы сканируется небольшим окном, т.е. пропускается сквозь набор весов (ядро свёртки), и результат сканирования отображается на соответствующем нейроне текущего слоя. Ядро свёртки интерпретируют как графическое кодирование какого-либо признака, например, наличие горизонтальной или вертикальной линии. Таким образом, набор плоскостей представляет собой карты признаков (feature map), что позволяет каждой плоскости находить «свои» участки изображения в любом месте предыдущего слоя.

4. Видеоконтроль присутствия студентов в аудитории

Программная реализация системы присутствия студентов в учебном классе выполнена с помощью ряда технологий программной инженерии. При конструировании системы были использованы язык программирования Python, фреймворк Qt5, СУБД SQLite, библиотеки NumPy, OpenCV, Keras, Hickle.

Задача автоматического учёта присутствия студентов на учебных занятиях включает следующие подзадачи:

- съёмка входящих в аудиторию студентов на видеокамеру;
- анализ кадров видеопотока на наличие лиц и их локализация;
- сравнение выделенных лиц с базой данных учащихся;
- отметка о присутствии на занятии (или опоздании) в случае успешной идентификации;
- сохранение данных в электронном журнале посещений.

Во время работы системы могут возникнуть различные проблемы, которые усложняют процесс распознавания, а именно: угол поворота головы студента был недостаточно прямым по отношению к камере, слабое освещение, студент прошёл мимо камеры очень быстро, слишком большое количество лиц на изображении, из-за чего некоторые кадры не были захвачены и т. д. Преподаватель также может допустить ошибку при использовании системы, например, включив режим регистрации опозданий слишком рано. Поэтому в системе предусмотрен режим ручного редактирования электронного журнала.

С помощью главного окна системы преподаватель может визуально наблюдать за процессом видеорегистрации и управлять информацией о студентах (рис. 6). В левой части окна высвечивается таблица со списком студентов группы, в которой преподаватель проводит занятие. Во время входа студента в аудиторию в таблице фиксируется время прибытия и его состояние (не отмечен, прибыл, опоздал, отсутствует). Для удобства ви-

деоконтроля цвет строки студента меняется в зависимости от его состояния. На данном рисунке показан факт регистрации студента в таблице текущего занятия по учебной дисциплине «Распознавание образов».

В окне справа находятся элемент выбора дисциплины, кнопки остановить/продолжить распознавание, перехода в режим опоздания, сброса занятия и его завершения. По умолчанию все студенты в списке имеют статус «Не отмечен», а при распознавании обозначаются как «Присутствует». При переходе в режим опоздания (кнопка «Опоздание») статус всех неотмеченных студентов изменяется на «Отсутствует», а те, кто будет распознан в этом режиме, помечаются статусом «Опоздание».

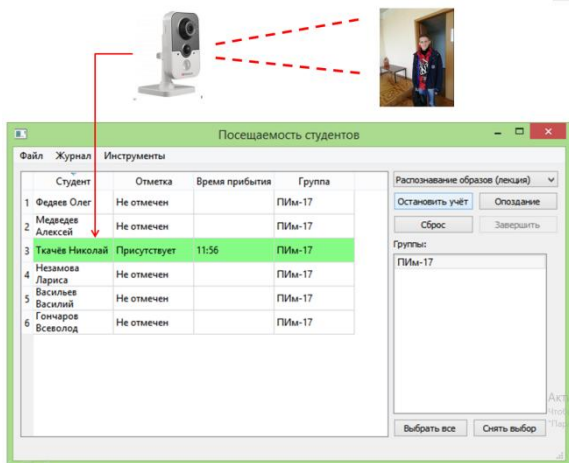


Рис. 6. Главное окно системы регистрации студентов

При нажатии кнопки «Завершить» прекращается визуальная регистрация студентов, а данные записываются из таблицы текущего занятия в общий электронный журнал. Эта кнопка активна только в режиме опоздания (таким образом, в журнале не должно оставаться записей со статусом «Не отмечен»).

Ячейки таблицы, отображающие отметки о присутствии, меняют свой цвет в зависимости от содержимого и поддерживают возможность редактирования своего значения при помощи выпадающего списка.

Пункты главного меню предоставляют преподавателю дополнительный сервис управления студентами:

- добавление, изменение имён студентов и удаление групп;
- отображение студентов по группам, добавление, редактирование и удаление студентов;

- изменение группы студентов;
- добавление и изменение фотографии (фотография может быть выбрана из файла либо с камеры; при этом на фотографии может быть обнаружено несколько лиц, в этом случае верное лицо необходимо выбрать).

Электронный журнал представляет собой реляционную таблицу, содержащую следующие данные: фамилии студентов, дату проведения занятия, предмет, группу, отметки о присутствии. В системе предусмотрено визуализация журнала в двух формах: по определённому предмету в группе или по конкретному студенту. На рис. 7 показано окно представления журнала по группе ПИМ-17 и дисциплине «Распознавание образов».

Студент	2019-05-28	2019-05-29	2019-05-31
1 Федеев Олег	Отсутствует	Отсутствует	Отсутствует
2 Медведев Алексей	Присутствует	Присутствует	Опоздание
3 Ткачев Николай	Присутствует	Отсутствует	Отсутствует
4 Незамова Лариса	Отсутствует	Присутствует	Отсутствует

Рис. 7. Просмотр посещаемости занятий студентами группы

Заключение

Данная система позволяет благодаря компьютерному зрению устранить потери времени преподавателя при ручной регистрации присутствия учащихся в аудитории и облегчить ему ведение журнала учёта. Правовые нормы, связанные с видеорегистрацией учащихся, будут соблюдены путём письменного согласия участников учебного процесса.

При проведении лекций с большим количеством студентов система автоматизирует учёт присутствия студентов в аудитории и накапливает в электронном журнале статистические данные об учебной дисциплине и каждом студенте в течение семестра.

Система может быть установлена на ноутбуке преподавателя и выполнять не только оперативный учёт присутствия студентов на занятиях, но и стать платформой для реализации других функций контроля проведения учебных занятий. Поэтому данную систему можно рекомендовать преподавателям университетов, учителям общеобразовательных школ и средних профессиональных образовательных учреждений.

Список литературы

- [**Deep Learning, 2023**] Deep Learning – Wikipedia [Электронный ресурс] – Режим доступа: https://en.wikipedia.org/wiki/Deep_learning.
- [**Гудфеллоу и др., 2018**] Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А.А.Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.
- [**Антонио и др., 2018**] Антонио Джулли, Суджит Пал. Библиотека Keras – инструмент глубокого обучения. Реализация нейронных сетей с помощью библиотек Theano и TensorFlow / пер. с англ. Слинкин А.А. – М.: ДМК Пресс, 2018. – 294 с.
- [**Reading, 2023**] Reading and Writing Images and Video - OpenCV 2.4.13.7 documentation [Электронный ресурс] // [opencv.org](https://docs.opencv.org/2.4/modules/highgui/doc/reading_and_writing_images_and_video.html). – Режим доступа: https://docs.opencv.org/2.4/modules/highgui/doc/reading_and_writing_images_and_video.html.
- [**Paul et al., 2001**] Paul Viola, Michael Jones. Rapid Object Detection using a Boosted Cascade of Simple Features, 2001. – P. 1-8.
- [**Le Cunn et al., 1995**] Le Cunn Y, Bengio Y. Convolutional neural networks for Images, Speech and Time Series, AT&T Laboratories, 1995. – P. 1 – 14.
- [**VGG Face, 2023**] VGG Face Descriptor [Электронный ресурс] // [robots.ox.ac.uk](http://www.robots.ox.ac.uk/~vgg/software/vgg_face/). – Режим доступа: http://www.robots.ox.ac.uk/~vgg/software/vgg_face/.
- [**SQLite, 2023**] SQLite Home Page [Электронный ресурс] // sqlite.org. – Режим доступа: <https://sqlite.org/index.html>.

УДК 004.822

МЕТОД СЦЕНАРНО-ИНФОРМАЦИОННОГО АНАЛИЗА ДЛЯ ИНТЕЛЛЕКТУАЛЬНОГО УПРАВЛЕНИЯ АДАПТИВНОЙ ПОДГОТОВКОЙ ГРУПП СПЕЦИАЛИСТОВ

И.А. Фролов (*igor-frolov-81@mail.ru*)

Военная академия войсковой противовоздушной обороны ВС РФ,
Смоленск

В работе обоснована необходимость интеллектуализации управления адаптивной подготовкой групп специалистов для выполнения сложных задач, согласованных по времени, месту и целям в организационно-технических системах различного назначения. Показан подход к адаптивной подготовке, а также задачи, решение которых востребовано для управления этим процессом. Предложен метод сценарно-информационного анализа для интеллектуального управления адаптивной подготовкой групп специалистов. Показаны результаты экспериментальных исследований по повышению уровня подготовленности специалистов групп за счет интеллектуального управления их адаптивной подготовкой.

Ключевые слова: адаптивная подготовка, метод сценарно-информационного анализа, группы специалистов, интеллектуальное управление, единое виртуальное пространство.

Введение

В настоящее время существует проблема повышения эффективности, организации и управления процессом подготовки специалистов в составе групп, выполняющих сложные задачи, согласованные по целям, месту и времени в организационно-технических системах (ОТС) различного назначения, обусловленная трудностями моделирования условий выполнения задач, приближенных к реальным, из-за больших материальных затрат и риска для жизни обучаемых специалистов. Одним из путей решения этой проблемы может являться адаптивная подготовка групп специалистов для отработки специальных задач и взаимодействия между ними с применением современных информационных технологий, технологий моделирования проблемных ситуаций и современных технических средств обучения в информационно-технологической среде (ИТС).

Под адаптивной подготовкой понимается процесс подготовки групп специалистов с учетом динамично изменяющихся условий выполнения задач в ОТС различного назначения и индивидуальных особенностей обучаемых специалистов.

1. Задачи интеллектуального управления адаптивной подготовки групп специалистов

Одновременная подготовка групп специалистов, имеющих разные задачи, объединенные одной целью и согласованные по времени и месту, в ИТС значительно повысит интенсивность подготовки, что, в конечном итоге, приведет к повышению ее эффективности за счет отработки большего количества решаемых задач.

Однако, это повлияет на качество и оперативность принятия управленческих решений руководителем занятия и выработки управляющих воздействий на специалистов в зависимости от полученных результатов. При повышении интенсивности подготовки групп специалистов руководителю занятия необходимо будет оценить, провести анализ качества выполненных операций алгоритма решения задачи, выявить совершаемые при этом ошибки, оперативно принять решение на изменение и выработать управляющие воздействия для каждого специалиста группы в реальном масштабе времени, практически не имея возможности для детального анализа всех альтернативных вариантов и возникающих последствий их реализации, что будет практически невыполнимо ввиду большого объема информации, которую необходимо будет обработать руководителю занятия, особенно на этапах отработки взаимодействия между различными группами.

Анализ процесса подготовки групп специалистов в ИТС позволил определить задачи, которые необходимо будет решать руководителю занятия для повышения качества управления при их адаптивной подготовке (табл. 1).

Таблица 1

Этапы управления адаптивной подготовкой групп специалистов ОТС	Задачи интеллектуального управления адаптивной подготовкой групп специалистов ОТС
Получение задачи на подготовку	• уяснение задачи руководителем занятия.
Поиск, сбор, анализ и обработка информации в ИТС	• определение руководителем занятия информационной потребности относительно поставленной задачи на подготовку;

		<ul style="list-style-type: none"> • анализ и оценка руководителем занятия качества найденной и полученной информации (полнота, точность, актуальность, своевременность и т.д.); • определение pertinентности полученной информации.
	Определение целей подготовки	<ul style="list-style-type: none"> • анализ динамического изменения условий выполнения задач по накопленному опыту; • определение и адаптация целей каждого этапа подготовки в соответствии с направленностью и спецификой решаемых задач; • -определение показателей и критериев достижимости каждого этапа подготовки.
Планирование подготовки и прогнозирование ее результатов	Анализ исходных данных для подготовки	<ul style="list-style-type: none"> • построение эталонных профилей специалистов; • определение и анализ текущего состояния специалистов; • актуализация профиля специалистов; • анализ эталонного и актуального профилей специалистов и определение степени несоответствия между требуемым и текущим уровнем подготовленности; • определение имеющихся ресурсов для подготовки (время, учебные объекты, ТСО и т.д.); • анализ прецедентных траекторий подготовки специалистов; • построение структуры процесса адаптивной подготовки; • определение необходимых информационных ресурсов для адаптивной подготовки; • построение сценарно-информационной модели адаптивной подготовки.
	Моделирование	<ul style="list-style-type: none"> • построение моделей эталонных процессов адаптивной подготовки; • построение дискретно-событийной модели процесса адаптивной подготовки;

		<ul style="list-style-type: none"> • моделирование процесса адаптивной подготовки; • анализ результатов моделирования процесса адаптивной подготовки и выявление рисков недостижимости целей каждого этапа адаптивной подготовки; • корректировка (при необходимости) модели процесса адаптивной подготовки и его моделирование в изменившихся условиях; • учет сформированной траектории в базе данных прецедентов.
	Принятие решения на подготовку	<ul style="list-style-type: none"> • постановка задач на подготовку; • реализация принятого решения.
	Контроль подготовки	<ul style="list-style-type: none"> • контроль результатов этапа подготовки; • контроль порядка выполнения алгоритмов решения задач; • фиксация ошибок, совершаемых при выполнении операций алгоритма решения задач; • контроль времени выполнения операций алгоритма решения задач.
	Оценивание результатов	<ul style="list-style-type: none"> • оценивание степени несоответствия заданных результатов с текущими; • оценивание выполнения правильной последовательности алгоритмов алгоритма решения задач; • оценивание слаженности групп ОТС различного назначения; • формирование комплексной оценки за этап подготовки (упражнение).
	Анализ результатов подготовки	<ul style="list-style-type: none"> • определение типов ошибок, пробелов в знаниях; • определение причин совершения ошибок; • анализ результата выполнения этапа подготовки (упражнения); • вывод о достижимости / не достижимости этапа подготовки;

	<ul style="list-style-type: none"> • анализ чувствительности специалистов к управляющим воздействиям; • анализ динамики изменения знаний, умений, навыков; • сопоставление результатов выполнения этапа подготовки специалиста с прецедентным этапом подготовки; • выявление причин в случае не достижения заданных результатов этапа подготовки; • учет результата этапа подготовки и пополнение базы данных прецедентов.
<p>Принятие решения на изменение и формирование управляющих воздействий</p>	<ul style="list-style-type: none"> • принятие решения (при необходимости) на изменение управляющих воздействий; • корректировка модели процесса адаптивной подготовки и его моделирование в изменившихся условиях; • анализ результатов моделирования; • формирование управляющих воздействий.

В условиях напряженности выполнения специальных задач и отработки взаимодействия между группами специалистов руководителю занятия необходимо будет контролировать правильность выполнения алгоритма решения всей совокупности задач, выполняемых этими группами, что, очевидно, будет невозможно вследствие большого количества личного состава в группах и мероприятий (операций), которые необходимо будет выполнить.

Психофизиологические возможности руководителя занятия будут снижаться с увеличением плотности потока поступающей информации, на основании которой необходимо будет принимать управляющее решение, особенно на этапах отработки взаимодействия между группами специалистов в ИТС (рис. 1).

Для оценки времени, затрачиваемого на принятие решения на подготовку групп специалистов и формирование для них управляющих воздействий проведены экспериментальные исследования, которые показали снижение оперативности принятия решения и выработки управляющих воздействий на обучаемых специалистов в зависимости от количества решаемых руководителем занятия задач в процессе адаптивной подготовки групп специалистов в ИТС (рис. 2). Кроме того, выявлено, что оперативность принятия решения также зависит от уровня подготовленности самого руководителя занятия.

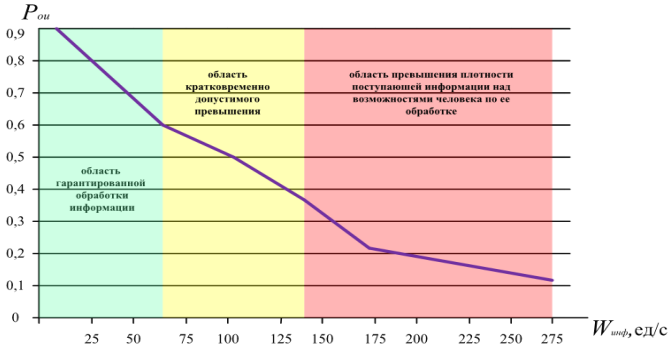


Рис. 1. Соответствие плотности потока информации, подлежащей обработке, психофизическим возможностям руководителя занятия

Все это обуславливает необходимость применения интеллектуальных методов для реализации интеллектуального управления адаптивной подготовкой групп специалистов в целях автоматизации процесса объективного контроля выполнения алгоритма решения задач специалистами, оценки и анализа полученных результатов, принятия решения на изменение и выработку управляющих воздействий для каждого специалиста из состава группы на основе результатов моделирования процесса их адаптивной подготовки [Бобряков и др., 2021]. Процесс интеллектуального управления адаптивной подготовкой групп специалистов должен быть конвергентным процессу моделирования подготовки для повышения качества управления, прогнозирования достижимости целей и задач подготовки.

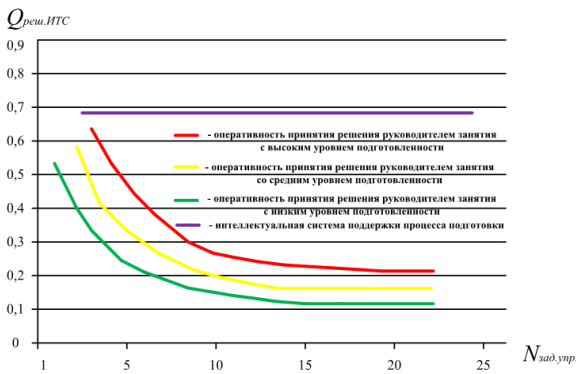


Рис. 2. Оперативность принятия решения руководителем занятия на изменение и формирование управляющих воздействий в ИТС

2. Метод сценарно-информационного анализа адаптивной подготовки групп специалистов

Для реализации интеллектуального управления процессом адаптивной подготовки групп специалистов разработан метод, позволяющий реализовать предлагаемый подход к построению процесса адаптивной подготовки групп специалистов и его анализу.

Метод сценарно-информационного анализа адаптивной подготовки групп специалистов включает следующие этапы:

Этап 1. Формализация целей, показателей и критериев адаптивной подготовки.

На основании руководящих документов по организации подготовки групп специалистов и регламентов решения задач осуществляется декомпозиция процесса адаптивной подготовки групп специалистов на этапы, для каждого этапа определяются цели, показатели и задаются критерии достижения целей каждого этапа с последующей их формализацией [Борисов и др., 2021]. На основании критериев настраиваются иерархические модели оценки. Первый уровень иерархии моделей предназначен для оценивания результатов выполнения задач отдельными специалистами из состава группы, второй уровень – для оценивания слаженности группы и достижения ее общей цели.

Этап 2. Построение эталонных профилей и моделей обучаемых специалистов.

Эталонный профиль отдельного специалиста описывает множество атрибутов (компетенций), которыми должен обладать определенный специалист группы. Так как приобретение компетенций носит динамический характер, то в эталонном профиле каждый набор компетенций соответствует определенному этапу подготовки.

Исходя из максимально достижимого уровня подготовленности специалистами с различными индивидуальными особенностями, определяемые на входном контроле с помощью различных тестов, все они классифицируются на «отличников», «хорошистов» и «троечников».

Эталонные профили отдельных специалистов строятся отдельно для каждого класса специалистов со своим набором атрибутов.

Эталонный профиль группы содержит набор компетенций и свойств, которыми должна обладать группа в целом на определенном этапе подготовки.

В процессе подготовки происходит сравнение эталонных профилей и моделей как отдельных специалистов, так и группы в целом.

Модель группы $M_{гп}$ представляет собой фиксирующую модель, содержащую информацию о количестве специалистов в группе и результаты отработки учебных задач на определенных этапах подготовки и соответствующий этим результатам уровень компетенций.

$$M_{sp} = \langle M_1, M_2, \dots, M_n \rangle,$$

где M_1, M_2, \dots, M_n – модель n -го специалиста группы.

Модель специалиста M_n представляет собой сетевую модель [Рыбина, 2023], содержащую личную информацию об обучаемом специалисте, уровень его знаний и умений, приобретенные на определенных этапах подготовки, индивидуальные особенности (характеристики обучаемости и заботы).

Этап 3. Построение моделей эталонных процессов адаптивной подготовки групп специалистов.

В целях организации и качественного планирования процесса адаптивной подготовки групп специалистов осуществляется построение моделей ее эталонных процессов, которые определяют последовательность этапов подготовки и необходимые ресурсы для их осуществления (время, технические средства обучения, реальные ОТС и т.д.). Модели эталонных процессов адаптивной подготовки строятся для каждого класса специалистов и различаются количеством информационных ресурсов, необходимых для прохождения этапов подготовки. Так, например, эталонный процесс «троечника» будет направлен, в основном, на приобретение практических умений с минимально необходимым набором знаний и будет предполагать большее количество повторений, а эталонный процесс «хорошиста» будет направлен на овладение большим количеством знаний для более глубокого понимания процессов, протекающих в ОТС.

Для моделирования процессов адаптивной подготовки групп специалистов выбраны методы процессного управления, позволяющие моделировать коллективную деятельность в сложных организационно-технических системах, основанные на понятии «бизнес-процесса» («делового» процесса), представляющего собой конечную совокупность взаимосвязанных, последовательных, регламентированных действий с распределением различных ресурсов [Сысков и др., 2015].

Для согласованного и формализованного представления «деловых» процессов адаптивной подготовки специалистов, состоящих из процессов индивидуальной и групповой подготовки, использованы усовершенствованные диаграммы нотации ARIS [Фролов, 2021].

Фрагмент процесса индивидуальной подготовки специалиста по различным дисциплинам в нотации ARIS представлен на рис. 3.



Рис. 3. Фрагмент процесса подготовки специалиста по различным дисциплинам

Очевидно, что при таком подходе представления процесса адаптивной подготовки специалистов «деловой» процесс можно интерпретировать как индивидуальную траекторию обучения специалистов. В зависимости от этапа подготовки индивидуальные траектории обучения должны отражать не только последовательность подготовки отдельных специалистов групп, но и объединяться в коллективную траекторию подготовки группы в целом.

Этап 4. Определение текущего уровня подготовленности специалистов группы и актуализация профиля группы.

Профиль группы включает профили отдельных специалистов ее из состава.

На данном этапе по результатам серии контрольных упражнений или по результатам предыдущих занятий с использованием нечетких моделей оценивания определяется текущий уровень подготовленности специалистов и согласованности их действий в составе группы.

Этап 5. Анализ актуализированных и эталонных профилей.

Для оценки степени рассогласования между целью определенного этапа подготовки и текущим состоянием группы специалистов, выработки на них управляющих воздействий при планировании процесса адаптивной подготовки групп специалистов осуществляется сравнение актуализированных и эталонных профилей. В результате для каждого специалиста определяется множество атрибутов, которые необходимо приобрести для достижения цели текущего этапа подготовки.

В случае, если «троечником» или «хорошистом» приобретены все атрибуты текущего этапа подготовки, для него формируются управляющие воздействия из более высокого класса. По результатам индивидуальной подготовки специалист может быть перемещен в более высокий класс и на этапе групповой подготовки получать задачи более высокого уровня.

Этап 6. Анализ «цифровых следов подготовки» и прецедентных сценариев адаптивной подготовки.

Результаты, предыдущих занятий групп специалистов («цифровые следы подготовки»), которые фиксируются в ИТС:

- динамика подготовки в случае, если специалист перешел из группы ОТС другого назначения;
- результаты прохождения по индивидуальной траектории подготовки (количество освоенных этапов подготовки, результаты контрольных занятий);
- не освоенные на этапе индивидуальной подготовки из-за ограничений временного ресурса компетенции, которые возможно освоить на этапе групповой подготовки и т.д.).

Анализ «цифровых следов подготовки» позволяет более качественно строить индивидуальные траектории специалистов с учетом динамики освоения ими программы подготовки.

Кроме того, для сокращения времени анализа и построения (корректировки) структуры процесса адаптивной подготовки группы специалистов происходит анализ его прецедентных сценариев.

Этап 7. Формирование сценария адаптивной подготовки.

С учетом всех имеющихся ресурсов для подготовки, особенностями текущего этапа адаптивной подготовки, индивидуальных особенностей специалистов из состава группы, и их уровня подготовленности происходит определение последовательности получения знаний и овладения умениями, необходимыми для достижения цели данного этапа адаптивной подготовки. Для этого на основе моделей эталонных процессов адаптивной подготовки строятся индивидуальные траектории подготовки каждого специалиста и группы в целом.

Этап 8. Построение сценарно-информационной модели адаптивной подготовки.

На основе сформированного сценария адаптивной подготовки из онтологии информационных ресурсов [Борисов и др., 2020] осуществляется выбор информации, необходимой для обеспечения получения знаний и умений специалистами при прохождении текущего этапа подготовки.

Сценарий адаптивной подготовки специалистов, наполненный необходимыми информационными ресурсами, представляет собой сценарно-информационную модель адаптивной подготовки групп специалистов.

3. Результаты экспериментальных исследований

Для оценки эффективности разработанного метода сценарно-информационного анализа для интеллектуального управления адаптивной подготовкой групп специалистов были проведены экспериментальные исследования.

Одной из основных задач исследований являлась оценка повышения уровня подготовленности специалистов, в том числе в составе группы, при интеллектуальном управлении их адаптивной подготовкой. Исследования проводились на тренажере, позволяющем обеспечивать одновременную подготовку специалистов в составе группы ОТС специального назначения, для выполнения сложных задач, согласованных по времени, месту и целям.

Повышение уровня подготовленности специалистов оценивалось относительно контрольной группы, управление подготовкой которой проводилось с использованием существующих методических подходов как на этапе индивидуальной, так и на этапе групповой подготовки.

Результаты оценки выполнения контрольных упражнений показывают повышение уровня подготовленности группы специалистов ОТС специального назначения, управление подготовкой которой проводилось с использованием разработанного метода, на этапе индивидуальной подготовки в среднем на 31,9%, на этапе групповой подготовки в среднем на 29,4%.

Заключение

Таким образом, повышение эффективности адаптивной подготовки групп специалистов за счет увеличения интенсивности этого процесса требует решение множества задач управления. Решение указанных задач руководителем занятия «вручную» снижает оперативность принятия решения и выработку управляющих воздействий на специалистов группы. Это обусловило разработку метода сценарно-информационного анализа для интеллектуального управления адаптивной подготовкой групп специалистов. Применение разработанного метода при проведении экспериментальных исследований позволило повысить оперативность принятия решений и формирования управляющих воздействий, что повлияло на повышение уровня подготовленности как отдельных специалистов, так и группы в целом.

Список литературы

- [Бобряков и др., 2021] Бобряков А.В., Мисник А.Е., Прокопенко С.А. Управление производственно-техническими процессами в сложных системах на основе нейро-нечётких сетей Петри // Интегрированные модели и мягкие вычисления в искусственном интеллекте: Сборник научных трудов. Т. 1. – М: Универсум, 2021.
- [Борисов и др., 2021] Борисов В.В., Захарченков К.В., Кутузов В.В., Мисник А.Е., Прокопенко С.А. Моделирование образовательных процессов на основе нейро-нечетких темпоральных сетей Петри // Прикладная информатика. – 2021. – № 4..
- [Рыбина, 2023] Рыбина Г.В. Интеллектуальные обучающие системы на основе интегрированных экспертных систем: учебное пособие. – М.: Директ-Медиа, 2023.

- [Сысков и др., 2015]** Сысков В.В., Борисов В.В. Постановка задачи процессного управления в сложных организационно-технических системах // Известия Смоленского государственного университета. – № 2/1. – Смоленск: СмолГУ, 2015.
- [Фролов, 2021]** Фролов И.А. Процессный подход к организации индивидуальной и коллективной подготовки специалистов на основе нечетких сетей Петри // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник научных трудов. Т. 1. – М.: Универсум, 2021.
- [Борисов и др., 2020]** Борисов В.В., Котов Д.В., Молявко А.А. Метод интеллектуального информационного поиска на основе композиционного онтологического подхода // Восемнадцатая Национальная конференция по искусственному интеллекту с международным участием, КИИ-2020. Труды конференции. – М.: Изд-во МЭИ, 2020.

Секция 4 | **МАШИННОЕ ОБУЧЕНИЕ,
НЕЙРОСЕТЕВЫЕ МЕТОДЫ**

UDC 004.056

**ENHANCING IOT SECURITY THROUGH
CONVOLUTIONAL VARIATIONAL AUTOENCODERS
AND RESAMPLING TECHNIQUES WITH GAN***

Huiyao Dong (*hydong@itmo.ru*)

ITMO University, Saint Petersburg, Russia

Igor Kotenko (*ivkote@comsec.spb.ru*)

St. Petersburg Federal Research Center of the Russian Academy
of Sciences, Saint Petersburg

The Internet of Things (IoT) is a pivotal constituent of the contemporary technological revolution and has experienced expeditious expansion in recent times. The proliferation of IoT devices has led to enhanced convenience and automation. However, the extensive deployment of IoT devices has also engendered concerns regarding data privacy and security. Among various detection and prevention methodologies, deep learning is emerging as a prominent trend. This paper utilises convolutional variational autoencoders (CNN-VAE) and resampling techniques for network attacks detection. The proposed methodology employs a hybrid data resampling technique to tackle the issue of imbalanced classes, followed by the implementation of a CNN-VAE classification model with a weighted loss function. The experiments demonstrate that the light-

* This research is being supported by the grant of RSF #21-71-20078 in SPC RAS.

weighted CNN-VAE outperforms the baseline models. Therefore, it possesses the capability to effectively detect intrusive activities in real-world settings and strengthen the IoT security.

Keywords: Internet of Things, cyber attacks, variational autoencoder.

1. Introduction

IoT is regarded as one of the pillars of the current technological revolution. In the past decade, IoT, a comprehensive, intelligent network-based system with reliable data transmission and sophisticated data processing between intra-network devices, has become increasingly prevalent. IoT network provides access to external computation and data and equips devices with computational intelligence to facilitate more convenient applications. IoT devices are ubiquitous; they can utilise sophisticated servers and databases, listen to and respond to human commands, and provide automated services. Moreover, the high reliance on IoT raises concerns regarding data privacy and security, as most of the information is stored in the cloud and is vulnerable to hacking. In addition to network disruption and blackout, network intrusions are a significant IoT risk. According to [Vega, 2020], approximately 31 billion IoT devices have already been connected, with 75 billion expected by 2025. In the meantime, widespread IoT adoption introduces IoT security risks. Extensive interconnections between vulnerable devices can result in scalability issues; meanwhile, the network is always required to access data or computational power outside of the specific device, and this interconnectivity and continuous use of IoT devices create the risk of widespread cyber attacks.

IoT security is crucial because full user acceptance cannot be accomplished without dependable artificial intelligence (AI) based IoT-appropriate security models. Numerous published works demonstrate the efficacy of AI approaches, including cyber security monitoring and management systems [Kotenko et al., 2012; Kotenko et al., 2014]. Taking into account current trends in machine learning-based network intrusion detection, the most promising methods considering the combination as traditional methods [Kotenko et al., 2018] as well as the application of deep learning (DL).

For instance, DoS attacks against IoT devices have been identified as significant hazards, and various machine learning security monitoring and management systems have been proposed. A hybrid framework utilising Long Short-Term Memory Neural Network (LSTM) and Gated Recurrent Unit Neural Network (GRU) could deliver effective learning ability and model generalizability for the long- and short-term trends in DDoS attack sequences [Yuan et al., 2017]. Besides, LSTM-based DDoS mitigating technology could be deployed at the Software Defined Network (SDN) controllers and applied to both cloud and fog environments [Priyadarshini Ret al., 2022]. Take Botnet as an

illustration; a DL-based method named Botnet Traffic Shark (BoTShark) utilised autoencoder and CNN to analyse network transactions and detect malicious traffic [Homayoun et al., 2018], whereas the bidirectional LSTM-RNN-based hybrid model also demonstrated its capability of detecting various types of botnets [McDermott et al., 2018]. Thus, DL appears to be an optimistic alternative to traditional machine learning and basic statistical technologies for protecting IoT devices from cyber attacks.

This study concentrates on the convolutional neural network (CNN)-based variational autoencoder (CNN-VAE) and resampling techniques for detecting traffic attacks, even though numerous previous works have confirmed autoencoders' intrusion detection capability.

The significant contribution of this paper is as follows:

(1) In order to detect malicious behaviour, we address the problem of extremely imbalanced classes in network traffic data and devise a hybrid data resampling technique. First, the edited nearest neighbour (ENN) algorithm is used to eliminate redundancy in major classes and strengthen the borders between distinct classes; then, the generative adversarial network (GAN) is used to generate synthetic samples that are close to the origins of minority classes.

(2) We deploy a VAE with CNN layers as the classification model and a weighted loss function to balance the performance against infrequent attacks.

(3) In the experiment, we compare the performance of CNN-AE with a traditional simple structure and VAE built with only fully connected networks against a baseline of one-dimensional deep convolutional model VGG-19 [Simonyan et al., 2015]. CNN-VAE's structure has demonstrated greater capability than other structures.

This paper's remainder is structured as follows: The second section provides a summary of related work in VAEs and their application to attack detection; the third section demonstrates the proposed methodology; the fourth section presents the experiment and its results; and the final section concludes the paper.

2. Related Work

Generative models are statistical models aiming to accurately represent the probability distribution derived from the available observations [Cinelli et al., 2021]. From the perspective of algorithms, An et al. [An et al., 2015] proposed a method for detecting anomalous data points using construction probability, which is a probabilistic measure of the variability of variables' distribution. This approach is based on algorithms. Alvis et al. [Alvis et al., 2021] conducted research on using variational autoencoders and stochastic gradient-descent scheme to address the complex nonlinear relationship between the latent and original representations. Research dedicated to semi-supervised learning [Kingma et al., 2015] also emphasised the importance of utilising deep generative models in conjunction with an approximate Bayesian-based variational method.

Nowozin [Nowozin, 2018] utilised the importance-weighted autoencoder (IWAE) methodology to create a set of estimators that are free from bias and rely on the IWAE bounds. These estimators exhibit enhanced performance in image classification and optimal computational efficiency. Rosca et al. [Rosca et al., 2017] investigated the limitations of the learning process in a generative model and proposed a combined model of both VAE and GAN with a unified objective-based optimisation approach.

These models have the potential to be utilised in IoT for anomaly detection, as they possess the inherent ability to differentiate objects, events, or behaviours that deviate from their normal characteristics. Wei [Wei, 2022] suggested a variational fuzzy autoencoder approach for machine vision anomaly detection in Industry 4.0, in which fuzzy entropy and Euclidean fuzzy similarity measurements are utilised for assess and identify anomalies. Chen et al. [Chen et al., 2020] offered a sliding-window convolutional VAE for the multivariate time series data analysis and the concurrent anomaly detection for industry IoT.

3. Methodology

AEs are a type of neural networks designed to acquire knowledge of data representation through encoding and decoding procedures. The encoder is responsible for reducing the dimensionality of the input data, while the decoder is for expanding the reduced representation back to its original input space. The goal is to reduce the reconstruction error, which compels the network to acquire the most significant features of the input.

Our previous research has evaluated the efficacy of conventional AE constructed using CNN architecture [Dong et al., 2023]. Meanwhile, rather than mapping original high dimensional data into bottlenecks, the variational autoencoder uses the probabilistic approach to learn the origins and map them into distribution-based space, namely the mean and variance. Additionally, it utilises the variational inference algorithm to perform data sampling and compute the latent variable distribution during the decoding process. The latent-variable model offers increased flexibility during the learning and parameter estimation process.

The proposed workflow involves a resampling stage for data, which utilises the ENN undersampling algorithm and GAN oversampling. This is followed by a classification stage for data, which applies VAE and weighted loss optimisation techniques. The objective is to create a VAE-based intrusion detection model that can detect anomalies and label various network traffic. To achieve this, the output layers are adjusted to evaluate the model's binary classification and multi-label classification capabilities.

3.1. Resampling with ENN and GAN

The significant challenge for intrusion detection is the presence of highly imbalanced data, where the frequency of network intrusions is negligible compared to the vast number of normal behaviours. The hybrid resampling technique is employed to tackle this problem. After performing data cleaning and normalisation, the edited nearest neighbours (ENN) algorithm is utilised for undersampling. This algorithm reduces the sample volumes of major classes and improves the border differentiation between different classes by removing fuzziness. The algorithm used in our method is demonstrated in Fig. 1.

Then, GAN was employed for generating synthetic samples close to the original samples of rare attacks. GANs can learn deep data representations even when the sufficient data is not enough. This is achieved by the generation of simulated samples. Additionally, GANs can enhance data classification performance by leveraging their generative capabilities. The workflow of oversampling with GAN is presented in Fig. 2.

Algorithm 1 ENN

```
1:  $X \leftarrow x$ 
2: int  $K = 5$ ;
3: int  $n = \text{len}(X)$ 
4: for  $\text{iteration} = 1, 2, \dots, n$  do
5:    $x_{knn} = K\text{NearestNeighbours}(k = K, \text{algorithms} = \text{ballTree})$ 
6:    $\text{classes} = \text{unique}(x_{knn}, \text{returnCounts} = \text{True})$ 
7:   for  $i$  in  $x_{knn}$  do
8:     Compute major and minor classes in  $x_{knn}$  then remove those in
       minor class.
9:   end for
10: end for
```

Fig. 1. Pseudocode – ENN for Undersampling

Algorithm 2 GAN

```
1:  $G = \text{generator}()$ 
2:  $D = \text{discriminator}()$ 
3:  $\text{GAN} = \text{defineGAN}(G, D)$ 
4: Train the discriminator on the original data
5: for  $\text{iteration} = 1, 2, \dots, k \dots$  do
6:   Generate  $G$  a batch of samples
7:   Evaluate the discriminator  $D$  on the generated samples
8:   Store the generated samples  $x_1, \dots, x_G$  into new set  $X'$ 
9: end for
10: Optimize loss binarycrossentropy with  $K$  epochs and minibatch size  $M$ 
11: Combine original and new set  $X_{\text{combine}} = X' + X$ 
```

Fig. 2. Pseudocode – GAN for oversampling

The autoencoder generator, constructed with a fully connected network, acquires the fundamental features of the data in the minor classes and generates new samples through simulation. Then, the discriminator subsequently performs binary classification tasks to assess the authenticity of generated samples. The optimisation utilises the binary cross entropy loss function. Upon completion of the training batches, the pre-trained model produces new samples that are subsequently combined with the original datasets and randomised.

3.2. CNN-VAE Functioning

This paper proposes to adopt the architecture of CNN-VAE, as the VAE can generate a latent space representation of original traffic data and CNN architecture can capture complex patterns and important features. The capabilities of this hybrid architecture have been validated in the image generation and reconstruction tasks. The model adheres to the conventional "encoder-decoder" structure, and its notation is as follows:

$$X' = g(f(X)), f \sim \text{encoder}, g \sim \text{decoder}. \quad (3.1)$$

The key idea behind the VAE is to use variational inference to capture the comprehensive probability distribution $p(x)$ from the observations $D = \{x\}_n$. Under the assumption of independence and identicalness of all samples, the Bayesian network can represent the probability of all observations by multiplying the probabilities of individual observations. Equation (3.2) can denote the log of the probabilistic representation:

$$\log p(D) = \sum_{x \in D} \log p(x). \quad (3.2)$$

The training of VAE is to optimise the lower bound on the log-likelihood of the data, namely Evidence Lower Bound (ELBO) [Kingma et al., 2019]. The ELBO comprises two constituents: the reconstruction term, which gauges the model's proficiency in reconstructing the input data from the latent representation, and the regularisation term, commonly referred to as Kullback-Leibler divergence, which coerces the latent distribution to be proximate to the prior distribution. With the joint probabilistic distribution of original observation x and latent representation z generated by encoder $p(\mathbf{x}, \mathbf{z})$, the ELBO can be expressed as below:

$$\log p(x) = E[\log p(x|z)] - KL[q(z|x) || p(z)]. \quad (3.3)$$

In (3.3), q and p represent probability distributions. The KL divergence term $KL[q(z|x) || p(z)]$ measures the difference between the approximate posterior distribution of observation over generated presentation $q(z|x)$ and the prior distribution of generated $p(z)$. E denotes the expected value, the notation

represents the reconstruction error from the original observation x and its latent representation z . Specifically, the KL divergence quantifies the information loss between the latent patterns generated and the original distribution and is used to enforce regularisation in VAE. The function can be expanded as below:

$$\begin{aligned}
 KL[q(z|x) || p(z)] &= \int \{q(z|x) \log[q(z|x)/p(z)]\} d(z/x) & (3.4) \\
 &= E[\log q(z|x) - \log p(z)] + E[\log p(z) - \log q(z|x)] .
 \end{aligned}$$

(3.4) illustrates that when the approximate posterior distribution is extremely close to the true prior distribution, the KL value is close to zero. The VAE utilises the reparameterization technique to optimise the ELBO. This technique involves a sampling step that occurs between the encoding and decoding processes. The reparameterization method eliminates the random noise variable from the model, resulting in improved gradient computation efficiency. With a Gaussian distribution $q(z|x)$ and its corresponding mean μ and log variance \log_var , the sampling algorithm generates a random vector ϵ from the standard normal distribution, then computes the latent vector z as the sum of the mean and a scaled version of the ϵ with the variance as (3.5):

$$z = \mu + \exp(0.5 * \log_var) * \epsilon. \quad (3.5)$$

3.3. Model Architecture and Weighted Loss Optimization

The model comprises two components: a CNN-AE and a classifier model constructed using fully connected layers.

Table 1 presents the architecture and activation function utilised in the Keras package. The encoder is comprised of three convolutional layers in one dimension, which are subsequently followed by max pooling layers. The flattened output is processed by two parallel dense layers to calculate the mean and log variance of the latent space. The reparameterization trick is utilised at this stage to generate samples from the latent space. The decoder is composed of three 1D convolutional layers that are succeeded by upsampling layers to restore the initial input shape. The VAE employs a loss function that comprises two components, namely the reconstruction loss and Kullback-Leibler divergence loss. The input to the classifier is the reconstructed representation. The data is then processed using three fully connected layers, and then output as the binary prediction.

Table 1

Model Architecture in the Keras

	CNN-VAE		Classifier
Structure	encoder_model = [Conv1D(32, 3), MaxPooling1D(2), Conv1D(64, 3), MaxPooling1D(2), Conv1D(128, 3), MaxPooling1D(2), Flatten(), z_mean = Dense(16), z_log_var = Dense(16) z = Lamb- da(sampling)([z_mean, z_log_var])]	decoder_model= [Dense(8*128), Reshape((8, 128), Conv1D(128, 3), UpSampling1D(2), Conv1D(64, 3), UpSampling1D(2), Conv1D(32, 3), UpSampling1D(2), Dense(input_dim)]	clf = [Dense(64), Dense(128), Dense(256), Dense(512), Dense(1)]
Activation	ReLU		ReLU & Sigmoid (output layer only)
Loss	Avg(mean square error + KL error)		Binary cross entropy

Although VAEs are capable of anomaly detection, their performance is usually shifted towards the majorities in a multiple label classification task. A MTL-based classifier is constructed for multi-classification to ensure balanced performance, in which each type of network traffic is defined as a binary classification task. Then, we design a weighted binary cross entropy (BCE) loss function for binary classification. Finally, we compute task-weighted vector based on the weight of the different classes in original dataset, to grant important tasks (namely, rare attacks) bigger weights.

For the loss of binary classification, we use class weighting techniques to assign higher weights to the minorities in training. With target variable consists of only positive and negative labels, their weights can be computed as (3.6):

$$w_{pos} = \frac{\text{count}(Y_{pos})}{\text{count}(Y)}, w_{neg} = \frac{\text{count}(Y_{neg})}{\text{count}(Y)}. \quad (3.6)$$

Then, we calculate the loss for original label y_i and predicted ones y_{pred} .

$$Loss = -\sum_{i=1}^N [y_i * \log(y_{pred}) + (1 - y_i) * \log(1 - y_{pred})]. \quad (3.7)$$

Lastly, we compute weighting vector for the two classes and final loss:

$$w_i = y_i * w_{pos} + (1 - y_i) * w_{neg}, w_i \rightarrow W, \quad (3.8)$$

$$\textit{Weighted_loss} = \textit{mean}(\textit{Loss} * \textit{W}). \quad (3.9)$$

This weighted loss function will assign higher weights to the positive minorities, so that models are forced to learn better representations of the minority classes.

4. Experiments

UNSW-NB15, a public IoT dataset [Nour et al., 2015], is utilised for experiments. It contains 49 protocol-based features, including benign behaviours and nine attacks. Except for regular behaviours, generic traffic, and exploits, all other traffic types are minor with less than 10% share.

We analyse the effectiveness and efficiency of deep and simple models, as well as the significance of CNN-based variational models and MTL structures when choosing baselines. VGG19 is selected for the former purpose, which is a deep convolutional network. The model consists of five convolutional and pooling modules that progressively increase the output units of each convolutional layer. Despite its complexity and lengthy training time, the model achieved outstanding image classification results. The objective is to compare the performance of the deep model with a simple structured AE and analyse the trade-off between training time and prediction performance. As the network data is one-dimensional, we modify the convolutional layers to fit the data structure. For the latter, we implement normal CNN-AE, fully connected layer-based VAE with the same layer settings as the CNN-VAE model. For the experiments, both binary and multi-label classification are performed; the major evaluation metrics are accuracy, precision, recall and F-1 score.

Table 2 displays the binary classification's performance. The four models exhibit comparable performance in terms of accuracy and f1-score. The CNN-VAE model exhibits superior performance compared to the other three models, as evidenced by its recall and precision scores of 0.991 and 0.855, respectively. Moreover, this model exhibits the minimum loss score, which implies superior overall performance. VGG exhibits superior performance compared to CNN-AE and DNN-VAE, demonstrating that substantial depth of layers can enhance predictive capabilities. Nonetheless, the process of training this model is exceedingly time-consuming. The time required to complete 30 epochs for the four models with a batch size of 128 are as follows: 34280s, 2491s, 1912s, and 2556s. The application of a deep model such as VGG in real-time detection is impeded by the temporal expense. The experiment also demonstrates that the utilisation of a combination of CNN-VAE can yield superior predictive performance compared to CNN-AE and DNN-VAE.

Table 2

Binary classification's performance

	Loss	Accuracy	Recall	Precision	F1-score
VGG_1D	0.324	0.878	0.987	0.857	0.916
CNN-AE	0.362	0.866	0.983	0.845	0.909
DNN-VAE	0.323	0.868	0.956	0.865	0.908
CNN-VAE	0.320	0.879	0.991	0.855	0.917

Table 3 presents the multi-classification performance. The empirical evidence indicates that STL models exhibit a high degree of proficiency in accurately classifying the majority class of football, while displaying suboptimal performance in identifying infrequent attacks. Despite the notable performance of VGG in reconnaissance detection, its training time per batch amounts to 30:06:22. In contrast, autoencoder requires only 3 to 10 minutes per batch.

Simultaneously, the MTL models that were improved through the utilisation of the weighted BCE loss function exhibit equitable performance across majorities and can detect infrequent attacks. Taking generic attacks as an example, CNN-AE-MTL model exhibits superior performance in all evaluation metrics compared to other models. Despite the low incidence rate of worm attacks, at only 0.07%, all three MTL models demonstrated the ability to detect such attacks. Notably, the DNN-VAE model exhibited the highest detection rate of 86.2%.

Table 3

Multi-classification performance

		Normal	Generic	Exploits	Fuzzers	DoS
VGG-1D-STL	precision	0.610	0.910	0.440	0.170	0.400
	recall	0.850	0.980	0.380	0.000	0.000
	f1-score	0.710	0.940	0.400	0.000	0.000
CNN-AE-STL	precision	0.550	0.490	0.450	0.000	0.000
	recall	0.940	0.990	0.000	0.000	0.000
	f1-score	0.700	0.660	0.000	0.000	0.000
CNN-AE-MTL	precision	0.970	0.999	0.482	0.442	0.209
	recall	0.639	0.978	0.211	0.202	0.031
	f1-score	0.771	0.988	0.294	0.278	0.054
DNNVAE-MTL	precision	0.869	0.997	0.665	0.484	0.534
	recall	0.551	0.978	0.572	0.166	0.700
	f1-score	0.675	0.987	0.615	0.247	0.606

CNNVAE-MTL	precision	0.904	0.993	0.825	0.180	0.872
	recall	0.668	0.978	0.788	0.988	0.690
	f1-score	0.768	0.985	0.806	0.304	0.770
		Reconnaissance	Analysis	Backdoor	Shellcode	Worms
VGG-1D	precision	0.860	0.000	0.000	0.000	0.000
	recall	0.340	0.000	0.000	0.000	0.250
	f1-score	0.490	0.000	0.000	0.000	0.000
CNN-AE-STL	precision	0.000	0.000	0.000	0.000	0.000
	recall	0.000	0.000	0.000	0.000	0.000
	f1-score	0.000	0.000	0.000	0.000	0.000
CNN-AE-MTL	precision	0.677	0.053	0.062	0.124	0.002
	recall	0.367	0.669	0.042	0.454	0.323
	f1-score	0.476	0.099	0.050	0.194	0.005
DNNVAE-MTL	precision	0.237	0.192	0.176	0.071	0.029
	recall	0.639	0.687	0.761	0.327	0.862
	f1-score	0.345	0.300	0.286	0.116	0.057
CNNVAE-MTL	precision	0.312	0.344	0.536	0.235	0.026
	recall	0.804	0.688	0.797	0.784	0.646
	f1-score	0.450	0.459	0.641	0.361	0.050

The CNN-VAE model exhibits superior performance compared to other models across various attack types, including exploits, analysis, backdoors, and shellcode. It is noteworthy to highlight that the DNN-VAE model achieved optimal detection performance for all genuine attacks, thereby substantiating the efficacy of variational models.

Conclusion

This study proposes a hybrid model of CNN-VAE classification and GAN-based resampling techniques for network intrusion detection. The experimental results show that the CNN-VAE model outperforms the traditional simple structured CNN-AE, DNN-VAE, and one-dimensional VGG-19 models in terms of overall performance, and it exhibits superior performance across various rare attacks. In summary, CNN-VAE demonstrates its potential for real-world application in network intrusion detection systems.

Our next step is to examine and optimize the model structure on massive datasets and time series-involved network data. Besides, we would like to explore the possibility of combining CNN and recurrent network layers for better detection performance.

References

- [Alvis et al., 2021] Alvis J.L., Laloy E., Nguyen F., Hermans T. Deep generative models in inversion: The impact of the generator's nonlinearity and development of a new approach based on a variational autoencoder // *Computers & Geosciences*. – 2021. – No. 152.
- [An et al., 2015] An J., Cho S. Variational Autoencoder based Anomaly Detection using Reconstruction Probability // *Special Lecture on IE*. – 2015. – No. 32.
- [Chen et al., 2020] Chen T., Liu X., Xia B., Wang W., Lai Y. Unsupervised Anomaly Detection of Industrial Robots Using Sliding-Window Convolutional Variational Autoencoder // *IEEE Access*, 2020. № 8.
- [Cinelli et al., 2021] Cinelli P., Marins A., Barros S., Netto L. S. Variational Autoencoder // *Variational Methods for Machine Learning with Applications to Deep Networks*. – 2021.
- [Dong et al., 2023] Dong H., Kotenko I.V. An Autoencoder-based Multi-task Learning for Intrusion Detection in IoT Networks // *2023 IEEE Ural-Siberian Conference on Biomedical Engineering, Radio electronics and Information Technology, Russia*. – 2023.
- [Homayoun et al., 2018] Homayoun S., Ahmadzadeh M., Hashemi S., Dehghantanha A., Khayami R. BoTShark: A Deep Learning Approach for Botnet Traffic Detection // *Cyber Threat Intelligence. Advances in Information Security*. – 2018. – No. 70.
- [Kingma et al., 2019] Kingma D.P., Welling M. An Introduction to Variational Autoencoders // *Foundations and Trends in Machine Learning*. – 2019. – No. 12 (4).
- [Kotenko et al., 2014] Kotenko I.V., Saenko I.B. Creating New-Generation Cybersecurity Monitoring and Management Systems // *Herald of the Russian Academy of Sciences*. – 2014. – Vol. 84, No. 6. – P. 424-431.
- [Kotenko et al., 2018] Igor Kotenko, Igor Saenko, Alexander Branitskiy. Framework for Mobile Internet of Things Security Monitoring based on Big Data Processing and Machine Learning // *IEEE Access*. – 2018. – Vol. 6. – P. 72714-72723. – DOI: 10.1109/ACCESS.2018.2881998.
- [Kotenko et al., 2012] Kotenko I.V., Polubelova O.V., Saenko I.V., Chechulin A.A. Applying Ontologies and Logical Inference for Security Information and event Management // *Highly Available Systems*. – 2012. – Vol. 8, No. 2. – P. 100-108.
- [McDermott et al., 2018] McDermott C. D., Majdani F., Petrovski A.V. Botnet Detection in the Internet of Things using Deep Learning Approaches // *2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 2018*.
- [Nour et al., 2015] Nour M., Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set) // *Military Communications and Information Systems Conference (MilCIS)*. – 2015.
- [Nowozin, 2018] Nowozin S. Debiasing evidence approximations: On importance-weighted autoencoders and jackknife variational inference // *Proceedings of the international conference on learning representations, Vancouver, Canada, 2018*.

- [Rosca et al., 2017] Rosca M, Lakshminarayanan B, Warde-Farley D, Mohamed S. Variational approaches for auto-encoding generative adversarial networks // arXiv e-prints 1706.04987. – 2017.
- [Priyadarshini et al., 2022] Priyadarshini R., Barik R. K. A deep learning based intelligent framework to mitigate DDoS attack in fog environment // Journal of King Saud University – Computer and Information Sciences. – 2022. – No. 34 (3).
- [Simonyan et al., 2015] Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition // ICLR. – 2015.
- [Vega, 2020] Vega, M. Internet of things statistics facts & predictions [Electronic resource]. 30, November, 2020. URL: <https://review42.com/internet-of-things-stats/>
- [Wei, 2022] Wei J. A Machine Vision Anomaly Detection System to Industry 4.0 Based on Variational Fuzzy Autoencoder // Computational Intelligence and Neuroscience. – 2022. – No. 3.
- [Yuan et al., 2017] Yuan X., Li C., Li X. DeepDefense: Identifying DDoS Attack via Deep Learning // 2017 IEEE International Conference on Smart Computing (SMARTCOMP), Hong Kong, China, 2017.

РАЗРАБОТКА НЕЙРОСЕТЕВОГО МОДУЛЯ ДЛЯ ПРОГНОЗИРОВАНИЯ СПРОСА НА ЭНЕРГОПОТРЕБЛЕНИЕ ОБЪЕКТАМИ МАССОВОГО СТРОИТЕЛЬСТВА*

М.С. Башлыков (*BashlykovMatS@mpei.ru*)

П.Р. Варшавский (*VarshavskyPR@mpei.ru*)

С.В. Гужов (*GuzhovSV@mpei.ru*)

А.А. Сесин (*SesinAA@mpei.ru*)

Национальный исследовательский университет «МЭИ», Москва

В работе предлагается решение актуальной задачи прогнозирования спроса на энергопотребление объектами массового строительства с использованием аппарата искусственных нейронных сетей (ИНС). Выполнен анализ информации о моделях ИНС, широко применяемых для решения поставленной задачи. Разработана модель ИНС, ориентированная на прогнозирование спроса на энергопотребление объектами массового строительства, которая обучена и протестирована на наборах реальных данных о школах типовой застройки. Разработан программный нейросетевой модуль для создания моделей ИНС пользователями без специальных навыков программирования.

Ключевые слова: машинное обучение, искусственная нейронная сеть, прогнозирование, энергопотребление.

Введение

Проблемы совершенствования и адаптации энергетических систем и комплексов стали особенно актуальными в первые десятилетия XXI века с учётом изменения структуры энергобалансов потребителей, развития собственной генерации в том числе с нетрадиционными возобновляемыми источниками энергии, переходом к риск-ориентированному подходу

* Работа выполнена в рамках проекта «Разработка нейросетевого программного обеспечения по прогнозированию спроса на тепловую энергию объектами массового строительства города Москвы» при поддержке гранта НИУ «МЭИ» на реализацию программы научных исследований «Приоритет 2030: Технологии будущего» в 2022–2024 гг.

при эксплуатации оборудования. В области энергетики актуальна проблема прогнозирования спроса на энергоресурсы. Это связано с вопросом более точного определения объёмов использования топлива и необходимостью избегать ситуации, когда потребители получают количество тепловой энергии большее или меньшее, чем им требуется в действительности. Более точное прогнозирование позволяет повысить уровень энергосбережения, уменьшит экологическую нагрузку и экономить денежные средства [Барабанов и др., 2019].

Для решения поставленной задачи прогнозирования широко применяются статистические методы, позволяющие строить регрессионные модели. Также широко применяются методы искусственного интеллекта, преобладающим из которых является машинное обучение (ML), в особенности искусственные нейронные сети (ИНС). Методы, пригодные для автоматизации и независимые от экспертных оценок, могут опираться либо на использование метода наименьших квадратов (МНК), либо на ИНС. Применение МНК осложняется необходимостью качественной подготовки исходных данных, что является крайне трудоёмким процессом. Некоторые влияющие факторы крайне сложно учесть в детерминированных моделях: скорость физического износа; неравномерность теплотехнических характеристик ограждающих конструкций; частичная замена старого оборудования на новое; некачественная эксплуатация, всё, что касается переходных процессов, тепловой инерции здания и пр. ИНС хорошо показывают себя в условиях неизвестных, нелинейных зависимостей между параметрами и менее требовательны к характеристикам, включённым в исходные данные. Именно по этим причинам при проведении численного эксперимента принято решение использовать модель ИНС для решения указанной задачи.

Таким образом, задача настоящей работы состоит в частичной замене детерминированного расчёта прогнозного спроса на тепловую энергию на стохастические модели, а именно – ИНС. В качестве исходных данных используются фактические данные по исследуемому объекту – группе школ, построенным по типовому проекту 65-426/1 – полученные в процессе эксплуатации.

1. Описание нейросетевого модуля

Нейросетевой модуль представляет собой веб-приложение, которое позволяет выполнять задачи, изначально доступные специалистам без навыка программирования, экспертам в прикладной области. Таким образом, этот подход предполагает собой отхождение от использования программного кода напрямую. Модуль реализован с использованием клиент-серверной архитектуры (рис. 1).

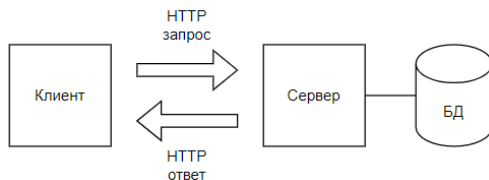


Рис. 1. Схема клиент-серверного взаимодействия

Общий список функциональных элементов веб-приложения.

- Клиентская часть:
 - загрузка наборов данных для последующего сохранения;
 - задание настроек новой модели ИНС;
 - управление наборами данных и моделями ИНС пользователя (переименование, удаление, просмотр информации о них);
 - выбор созданной модели ИНС для прогнозирования показателей на основе выбираемого набора данных;
 - загрузка информации о модели (показатели обучения, её структура).
- Серверная часть:
 - Обработка запросов:
 - для получения клиентом страниц и данных о сохранённых наборах данных и моделях ИНС;
 - для сохранения и обработки отправленных пользователем файлов (наборов данных), настроек модели ИНС;
 - для удаления неактуальных наборов данных и моделей ИНС по выбору пользователя.
 - Модуль работы с данными и моделями ИНС:
 - обработка набора данных и его сохранение в файловой системе сервера, а также в базе данных;
 - обучение модели ИНС и её сохранение в файловой системе сервера, а также в базе данных;
 - проведение прогнозирования при помощи сохранённой модели ИНС по сохранённому набору данных.

На данный момент клиентская часть создаётся при помощи стандартных средств HTML5, CSS и JavaScript, которые отражают в веб-браузере пользователя интерфейс для работы с программой, а также управляют логикой клиента. Серверная часть разработана на языке программирования Python с использованием фреймворка Flask для организации веб-части, библиотеки SQLAlchemy для организации взаимодействия с серверной БД, а также различными библиотеками машинного обучения и обработки данных (Tensorflow, Scikit-learn) и работы с наборами данных в различных форматах (NumPy, Pandas) [Raschka et al., 2019]. Интерфейс главной страницы приведён на рис. 2.

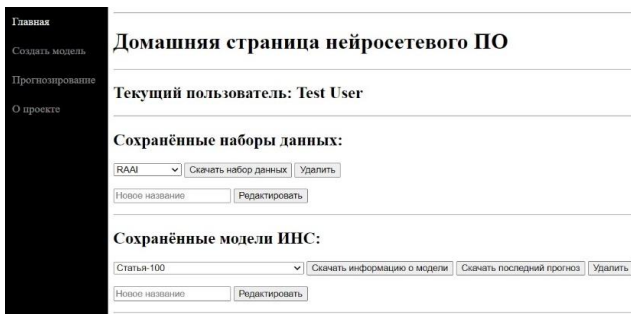


Рис. 2. Интерфейс главной страницы

При работе с модулем пользователь может загрузить файлы в формате .xlsx для дальнейшей работы с ними. При выборе набора данных из загруженных или загрузке нового файла пользователь устанавливает характеристики ИНС для обучения: число эпох обучения, используемую функцию потерь: относительную погрешность в процентах (англ. Mean Absolute Percentage Error, сокр. MAPE), среднюю абсолютная погрешность (англ. Mean Absolute Error, сокр. MAE) или среднее значение квадрата погрешности (англ. Mean Squared Error, сокр. MSE). Далее настраивается описание скрытых слоёв: тип слоя, число узлов в нём и используемая функция активации (рис. 3).

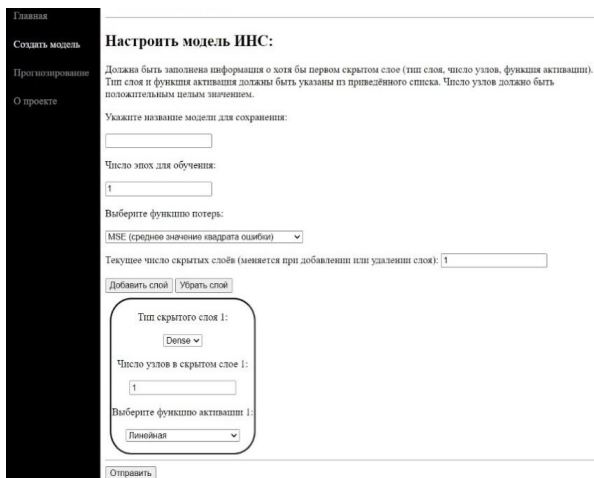


Рис. 3. Элемент интерфейса настройки модели

Обучение ИНС происходит на сервере, после чего пользователь может ознакомиться с описанием обучения модели в виде отчёта, а также производить прогнозирование при помощи сохранённой модели на основе файла .xlsx. Сам прогноз может быть скачан в виде файла .xlsx (рис. 4).

	A	B	C	D	E	F	G	H	I	J	K	L
1	ТЭ Потреблено, Гкал	ЭЭ Потреблено, кВт*ч	ХВС Потреблено, м3	ГВС Потреблено, м3	Т, С	Р, мм.рт.ст.	U, %	F, м/с	Nh, %	Год	Месяц	Прогноз
2	54,23	13530	112	150,41	4,7	766,7	81	1,5	90	2023	1	170,2282
3	190,72	12880	119	174,8	-5,4	755,9	84	1,4	90	2022	1	170,2545
4	73,15	13700	192	113,34	-0,9	757,4	77	1,2	75	2022	2	122,7322
5	130,995	11740	200	233,22	-0,7	767	60	1,2	40	2022	3	137,3988
6	86,36	13080	150	260,918	5,8	759,2	69	1,4	75	2022	4	95,05849
7	19,16	7420	109	131,36	10,5	758,6	56	1,4	50	2022	5	33,61943
8	0	2300	28	12,573	18,9	760,6	63	0,9	40	2022	6	1,925016
9	0	1380	18	5,858	20,8	739,5	66	0,9	50	2022	7	-1,40923
10	0	1620	23	12,629	22	765,8	58	0,7	25	2022	8	5,011336
11	34,7	12480	228,86	248,86	10,1	758,8	82	0,9	75	2022	9	29,40421
12	76,59	14900	151	235,842	7,5	762,7	82	0,8	75	2022	10	84,89262

Рис. 4. Пример файла с прогнозом

2. Применение моделей ИНС для прогнозирования спроса на энергопотребление объектами массового строительства

Имеющаяся задача регрессии попадает под категорию задач обучения с учителем. Для данного типа задач активно применяются различные модели ИНС: многослойные перцептроны (MLP), рекуррентные модели и их подтипы (RNN, LSTM, GRU и прочие), а также свёрточные модели (CNN) [Sarker, 2021].

В настоящее время наибольшее внимание уделяется прогнозированию спроса на электроэнергию, что может быть вызвано большей областью применения этого ресурса (питание нагревательных и охлаждающих приборов, бытовой техники), а также наличием большого объёма данных по потреблению. Среди моделей ИНС активно применяются в исследовательских работах различные типы рекуррентных моделей, а также их модификации [Kaur et al, 2022], [Shin, Woo, 2022], [Dong, 2022], [Sharma et al., 2022], [Mahjoub et al., 2022], [Zhao et al., 2023], [Bian et al., 2022], [Jiang et al., 2022]. Отдельно стоит выделить исследования, которые направлены не только на применение указанных моделей (например, в работе [González et al., 2022] исследуется уникальная модель SNN), а также развитие получают алгоритмы оптимизации, которые находят применение при разработке моделей ИНС [Liao, Jimenez, 2022].

Для разработки модели ИНС был подготовлен набор данных о школах по типовому проекту 65-426/1, собранный за период с января 2018 по январь 2023 г. Набор состоит из 1691 наблюдения, при этом 61 наблюдение экспертами по предметной области было отобрано в качестве тестового набора. Прогнозируемой переменной является потребление тепловой

энергии зданием школы (Гкал). Среди других данных по школам используются данные об объёмах потребляемой электроэнергии (кВт*ч), объём ХВС и ГВС (м³). Собранные климатические данные включают в себя среднемесячные значения показателей: температуры воздуха за месяц (°C), относительной влажности воздуха (%), атмосферного давления (мм.рт.ст.), скорости ветра (м/с), облачности небосвода (%). Для тестирования были выбраны различные варианты полносвязных ИНС (MLP) с числом скрытых слоёв от 1 до 3 с активационными функциями ReLU. Входной слой состоит из 10 нейронов (8 для указанных выше параметров, а также значение месяца и года), выходной слой содержит 1 нейрон. В качестве функции потерь использовалась MSE.

В процессе преобработки данных была выявлена их неполнота: в некоторой части наблюдений отсутствуют данные по отдельным показателям, что требует проведения очистки данных. Первоначальный подход предполагал удаление из рассмотрения некорректных данных (126 наблюдений) с последующей нормализацией значений для приведения к распределениям с нулевым средним и единичным стандартным отклонением: $x_{scaled} = (x - \mu) / \sigma$, где x – рассматриваемый показатель отдельного наблюдения, μ – среднее значение показателя во всей выборке, σ – среднеквадратичное отклонение показателя во всей выборке, x_{scaled} – новое значение показателя в наблюдении.

На основании проведённых численных экспериментов построены различные модели ИНС и выполнено их сравнение между собой. У большинства моделей лучшие результаты для них получены при обучении в течение порядка 50, 100 или 200 эпох при объёме пакетов (batch) 64: погрешности при большем числе эпох либо значительно не изменялись, либо ухудшались. Использование трёх скрытых слоёв и большего числа нейронов не дало значимого улучшения при сравнении с моделями с двумя слоями и средним числом нейронов. Так, одна модель с тремя слоями (512, 32 и 32 нейрона), обученная при объёме пакетов 64 и в течение 96 эпох, показала результат по MSE на тестовом наборе 672.94, когда другие модели с тремя слоями (64, 64 и 32 нейрона), объёмом пакетов 64 при 253 эпохах и двумя слоями (512 и 64 нейрона) и объёмом пакетов 32 при 55 эпохах дали близкие значения MSE 684.51 и 694.92, соответственно. Тем не менее, учитывая малое время обучения этих моделей, это не является недостатком. Таким образом, из рассмотренных моделей наилучшей оказалась ИНС с тремя скрытыми слоями, в первом из которых 512 нейрона, а во втором и третьем по 32. На рис. 5 приведены результаты работы на тестовой выборке модели с выбранными показателями.

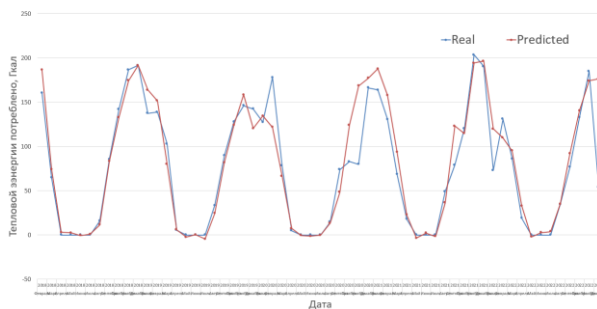


Рис. 5. Результат ИНС на тестовом наборе

В связи с большим числом наблюдений, в которых прогнозируемое значение принимает нулевое значение, применение метрики, вычисляющей относительную погрешность, не является информативным. Другие метрики показывают, что RMSE достигает 25.94, а MAE 14.90, однако если рассматривать отдельные абсолютные погрешности по каждому наблюдению, тогда медианное значение равно 8.38, что свидетельствует о завышении среднего отклонения из-за отдельных аномальных прогнозов. Из графика видно, что основные ошибки происходят в зимние месяцы, а также в марте, особенно в случае отдельных годов, когда происходили аномальные выбросы по теплопотреблению.

На основании полученных результатов можно сделать вывод о том, что следует проводить дальнейшие исследования по работе механизма ИНС в задачах прогнозирования спроса на энергоресурсы. В качестве рекомендации можно отметить необходимость сбора большего количества данных об энергопотреблении объектов типовой застройки, в частности, целесообразно для зимнего сезона использовать не среднемесячные показания, а более частые наблюдения (дневные колебания по энергопотреблению или учет пиковых часов потребления).

Заключение

В работе отмечена актуальность разработки методов и программных средств для решения задачи прогнозирования спроса на энергопотребление объектами массового строительства. Для решения указанной задачи предлагается использовать аппарат ИНС. Предложена модель ИНС, состоящая из входного слоя с 10 нейронами, трёх скрытых полносвязных слоёв из 512, 32 и 32 нейронов с функциями активации ReLU и выходным слоем с одним нейроном. Модель ИНС обучена в течение 96 эпох с использованием функции потерь MSE и объёмом пакетов 64. Модель протестирована на реальных наборах данных для школ по типовому проекту 65-426/1. Описаны особенности программной реализации нейросетевого модуля, разработанного при помощи

HTML5, CSS, JavaScript и языка Python с использованием фреймворка Flask и библиотек SQLAlchemy, Tensorflow, Scikit-learn, NumPy и Pandas. Разработка данного модуля ведется в рамках программы научных исследований «Приоритет 2030: Технологии будущего» в 2022–2024 гг.

Список литературы

- [**Барабанов и др., 2019**] Барабанов А.О., Гужов С.В. Прогнозирование тепловой нагрузки на отопление с использованием ИНС // С.О.К. – 2019. – № 11. – С. 28-30.
- [**Bian et al., 2022**] Haihong Bian, Qian Wang, Guozheng Xu, Xiu Zhao. Load forecasting of hybrid deep learning model considering accumulated temperature effect // *Energy Reports*. – 2022. – Vol. 8. – P. 205-215. – DOI: 10.1016/j.egy.2021.11.082.
- [**Dong et al., 2022**] Dong X., Deng S. & Wang D. A short-term power load forecasting method based on k-means and SVM // *J Ambient Intell Human Comput.* – 2022. – Vol. 13. – P. 5253-5267. – DOI: 10.1007/s12652-021-03444-x.
- [**González et al., 2022**] González Sopeña J.M., Pakrashi V., Ghosh B. A Spiking Neural Network Based Wind Power Forecasting Model for Neuromorphic Devices // *Energies*. – 2022. – Vol. 15. – DOI: 10.3390/en15197256.
- [**Jiang et al., 2022**] Feng Jiang, Qiannan Zhu, Jiawei Yang, Guici Chen, Tianhai Tian. Clustering-based interval prediction of electric load using multi-objective pathfinder algorithm and Elman neural network // *Applied Soft Computing*. – 2022. – Vol. 129. – DOI: 10.1016/j.asoc.2022.109602.
- [**Kaur et al., 2022**] Kaur D., Islam S.N., Mahmud M.A., Haque M.E., Dong Z.Y. Energy forecasting in smart grid systems: recent advancements in probabilistic deep learning // *IET Gener. Transm. Distrib.* – 2022. – Vol. 16. – P. 4461-4479. – DOI: 10.1049/gtd2.12603.
- [**Liao et al., 2022**] Liao S., Jimenez G. A new optimal prediction technique for energy demand based on CNN and improved water strider algorithm: a study on socio-economic-climatic parameters // *Evolving Systems*. – 2022. – Vol. 13. – P. 759-775. – DOI: 10.1007/s12530-021-09409-x.
- [**Mahjoub et al., 2022**] Mahjoub, S., Chrifi-Alaoui, L., Marhic, B., Delahoche, L. Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks // *Sensors*. – 2022. – Vol. 22. – DOI: 10.3390/s22114062.
- [**Raschka et al., 2019**] Sebastian Raschka, Vahid Mirjalili. Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2, 3rd Edition (ISBN 978-1-78995-575-0). – Packt Publishing, 2019.
- [**Sarker, 2021**] Sarker I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. // *SN COMPUT. SCI.* – 2021. – Vol. 2, 420. – <https://doi.org/10.1007/s42979-021-00815-1>.
- [**Sharma et al., 2022**] Sharma K., Dwivedi Y.K. & Metri B. Incorporating causality in energy consumption forecasting using deep neural networks. // *Ann Oper Res*. – 2022. – DOI: 10.1007/s10479-022-04857-3.
- [**Shin et al., 2022**] Shin S.-Y., Woo H.-G. Energy Consumption Forecasting in Korea Using Machine Learning Algorithms. // *Energies*. – 2022. – Vol. 15. – DOI: 10.3390/en15134880.
- [**Zhao et al., 2023**] Yubo Zhao, Ni Guo, Wei Chen, Hailan Zhang, Bochao Guo, Jia Shen, Zijian Tian. Multi-step ahead forecasting for electric power load using an ensemble model // *Expert Systems with Applications*. – 2023. – Vol. 211. – DOI: 10.1016/j.eswa.2022.118649.

НЕПОЛИНОМИАЛЬНОСТЬ ЧИСЛА СХОДСТВ ДЛЯ СЛУЧАЙНЫХ ОБУЧАЮЩИХ ВЫБОРОК

Д.В. Виноградов (*krrguest@yandex.ru*)
Федеральный исследовательский центр
«Информатика и управление» РАН, Москва

В этой статье мы приводим дополнительный аргумент в пользу вероятностного подхода к извлечению знаний с использованием операции сходства (ВКФ-метод) через неполиномиальность числа всех кандидатов. Более конкретно, мы доказываем два результата о такой неполиномиальности для случайных обучающих выборок, порождаемых последовательностями Бернулли. В случае плотной решетки будет доказано, что вероятность появления большой подрешетки, изоморфной булевой алгебре, в решетке кандидатов будет стремиться к единице по мере увеличения размера выборки. Для решетки средней плотности мы воспроизводим слегка измененный аргумент Т. Сакураи о неполиномиальности среднего числа кандидатов.

Ключевые слова: сходство, испытания Бернулли, решетка кандидатов, булева алгебра.

Введение

Парадигма извлечения знаний с помощью бинарной операции сходства была предложена В.К. Финном более 40 лет тому назад [Финн, 1983]. В ходе исследований этот подход развился в ДСМ-метод автоматического порождения гипотез [Финн и др., 2009]. Основной идеей, почерпнутой у Ф. Бэкона и Д.С. Милля, было использование множества общих признаков некоторой группы положительных примеров в качестве правдоподобной причины отнесения их к целевому классу. Затем, на наш взгляд, произошло «переобучение» – осмысленными стали считаться сходства всех групп положительных примеров. То, что это не так, мы смогли продемонстрировать [Виноградов, 2017] с помощью так называемых «фантомных» (или «случайных») гипотез. Не спасает даже «запрет контр-примеров» – метод фальсификации гипотез, предложенный В.К. Финном.

Среди исследователей ДСМ-метода с самого начала был известен результат о возможности «экспоненциального взрыва» – возникновение экспоненциального числа сходств (кандидатов в гипотезы о причинах)

относительно объема обучающей выборки. Наиболее резко это проявляется в случае n -мерной булевой алгебры, когда минимальная обучающая выборка занимает n^2 бит памяти, но кандидатов возникает 2^n , следовательно, для их хранения требуется $2^n \cdot n$ бит памяти. Для 32-мерной булевой алгебры выборка требует 128 байт, а памяти для сохранения всех кандидатов нужно уже $2^{37} = 16$ Гигабайт. Для 64-мерной булевой алгебры выборка занимает 512 байт, а памяти для сохранения всех кандидатов требуется $2^{67} \approx 10^{20}$ байт. Никакого разумного времени для записи такого объема памяти не хватает!

Интересным фактом о ДСМ-методе является то обстоятельство, что при его формальном изложении средствами многозначных логик [Финн, 1983] универсумом описаний объектов являлась как раз булева алгебра.

Ситуацию практической невозможности закончить вычисления сходств за реальное время игнорировали, полагая, что она носит единственный характер, однако это не так. Мы докажем, что для плотных случайных обучающих выборок, порождаемых последовательностями испытаний Бернулли с параметром $p = (1-1/m)$, вероятность возникновения подрешетки, изоморфной m -мерной булевой алгебре, будет стремиться к 1 при $m \rightarrow \infty$.

Возможным возражением к этому результату является аргумент, что в реальности обучающие выборки плотными не бывают, и, фактически, мы работаем с решеткой сходств, а не со всей булевой алгеброй.

Ситуация с ДСМ-вычислениями значительно улучшилась после того, как С.О. Кузнецов осознал [Кузнецов, 1999] возможность применения Анализа формальных понятий (АФП) – современного раздела теории решеток – к алгоритмизации ДСМ-метода.

Но даже с этими улучшениями проблема остается: для случая случайной обучающей выборки средней плотности (порождаемой последовательностью Бернулли с $p = 1/2$) мы воспроизведем слабую модификацию недавнего результата Т. Сакураи [Sakurai, 2021] о неполиномиальности среднего числа кандидатов.

Целью настоящей работы является окончательно похоронить заблуждения о возможности применения ДСМ-метода к обучающим выборкам достаточно большого объема. Надеемся, что доказанные результаты убедят любого непредвзятого читателя в этом.

1. Элементы теории решеток

Обучающей выборкой обозначим бинарное отношение $I \subseteq O \times F$ между элементами множества O , которые мы называем *обучающими объектами*, и элементами множества F , которые мы называем (бинарными) *признаками*.

Пару $\langle A, B \rangle$ назовем кандидатом, если $A = B' = \{o \in O: \forall f \in B [of]\} \subseteq O$ и $B = A' = \{f \in F: \forall o \in A [of]\} \subseteq F$. Первая компонента $A \subseteq O$ кандидата $\langle A, B \rangle$ называется *списком родителей*, а вторая $B \subseteq F$ – (общим) *фрагментом* кандидата.

Частичный *порядок* между кандидатами задается правилом: $\langle A, B \rangle \leq \langle C, D \rangle \Leftrightarrow B \subseteq D$.

Множество всех кандидатов (для фиксированной обучающей выборки) образует решетку с операциями:

$$\langle A, B \rangle \wedge \langle C, D \rangle = \langle (B \cap D)', B \cap D \rangle, \quad (1)$$

$$\langle A, B \rangle \vee \langle C, D \rangle = \langle A \cap C, (A \cap C)' \rangle. \quad (2)$$

Пусть $O = \{o_1, o_2, \dots, o_n\}$ будет множеством обучающих примеров (*коатомов*), каждый из которых описывается признаками из списка $F = \{f_1, f_2, \dots, f_n\}$, и $o_i I f_j \Leftrightarrow j \neq i$. Эта (минимальная) обучающая выборка представлена в табл. 1.

Таблица 1

Минимальная обучающая выборка для булевой алгебры

OF	f_1	f_2	\dots	f_{n-1}	f_n
o_1	0	1	\dots	1	1
o_2	1	0	\dots	1	1
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
o_{n-1}	1	1	\dots	0	1
o_n	1	1	\dots	1	0

Ясно, что $\{o_{j_1}, o_{j_2}, \dots, o_{j_s}\}' = F \setminus \{f_{j_1}, f_{j_2}, \dots, f_{j_s}\}$, так как добавление в сходство примера o_k с номером k удаляет из фрагмента признак f_k с тем же самым номером k . Очевидно, что даже минимальная обучающая выборка из табл. 1 занимает n^2 бит, а порождаемая из нее булева алгебра кандидатов требует $2^n \cdot n$ бит памяти, что препятствует использованию логико-комбинаторного ДСМ-метода [Финн, 1983] к извлечению знаний из обучающих выборок большого объема (Big Data), когда сначала порождаются все кандидаты, а затем некоторые из них отбрасываются с помощью проверок дополнительных фальсифицирующих условий (например, «запрет контр-примеров»).

С такой ситуацией можно было бы смириться, если бы случай возникновения булевой алгебры как подрешетки был уникальным. Но, как будет показано в следующем разделе, для случайных плотных обучающих выборок вероятность возникновения подрешетки, изоморфной m -мерной булевой алгебре, будет стремиться к 1 при $m \rightarrow \infty$.

Более того, как недавно показал японский математик Т. Сакураи [Sakurai, 2021], неполоynomialность среднего числа кандидатов имеет место и для случайных обучающих выборок средней плотности. Мы воспроизведем немного упрощенную версию его рассуждений в конце следующего раздела.

2. Основные результаты

2.1. Плотные случайные обучающие выборки

Для получения случайной плотной обучающей выборки из k примеров, описываемых n бинарными признаками, мы используем последовательность испытаний Бернулли $(\delta_{i,j})$, где $1 \leq i \leq k$ и $1 \leq j \leq n$. Вероятность успеха в этой серии положим равной $(1 - 1/m)$.

Для того, чтобы получить m -мерную булеву алгебру как подрешетку кандидатов для обучающей выборки из $k \geq m$ примеров нам потребуется найти m строк (перестановкой можно считать их первыми строками) и m столбцов, которые содержат единственный ноль и $m - 1$ единиц, так, чтобы для каждой строки в отобранных столбцах встречался ровно один ноль. Например, в табл. 2 в столбце f_j единственный ноль встречается в строке, соответствующей обучающему примеру o_i .

Таблица 2

Случайная обучающая выборка для булевой алгебры

OF	f_1	...	f_j	...	f_n
o_1	$\delta_{1,1}$...	1	...	$\delta_{1,n}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
o_{i-1}	$\delta_{i-1,1}$...	1	...	$\delta_{i-1,n}$
o_i	$\delta_{i-1,1}$...	0	...	$\delta_{i,n}$
o_{i+1}	$\delta_{i+1,1}$...	1	...	$\delta_{i+1,n}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
o_m	$\delta_{m,1}$...	1	...	$\delta_{m,n}$

Обозначим через c_i столбец высоты m , где $m-1$ позиция заполнены единицами, а единственный ноль встречается на i -ой позиции. Тогда ситуацию с j -м столбцом табл. 2 можно описать равенством $f_j = c_i$.

Рассмотрим теперь серию испытаний Бернулли $\delta_{1,1}, \dots, \delta_{m,n}$, заполняющую табл. 2 с вероятностью успеха $p = (1 - 1/m)$. Значение для параметра p определяется как оценка максимального правдоподобия, чтобы

обеспечить в среднем ровно один ноль в столбце высоты m . Тогда вероятность $\mathbf{P}[f_j = c_i]$ получить столбец e_i равна $s = p^{m-1}(1-p)$ для любых i и j , причем события $[f_j = c_i]$ и $[f_{j'} = c_{i'}]$ независимы для разных i, i', j и j' .

Теорема 1. При $m \rightarrow \infty$ и $n \geq e \cdot m \cdot \ln m$ для вероятности $\mathbf{P}[\forall j \exists i (f_j = c_i)]$ порождения m -мерной булевой алгебры как подрешетки кандидатов имеем $\lim_{m \rightarrow \infty} \mathbf{P}[\forall j \exists i (f_j = c_i)] \geq 1 - m^{(1-e \cdot m \cdot s)} \rightarrow 1$.

Доказательство. Рассмотрим вероятность дополнительного события $\mathbf{P} = \mathbf{P}[\exists j \forall i (f_j \neq c_i)]$. Неравенство Буля влечет $\mathbf{P} \leq m \cdot \prod_{i=1}^{e \cdot m \cdot \ln m} (1 - \mathbf{P}[f_j = c_i])$ из-за независимости событий $[f_j = c_i]$ и $[f_{j'} = c_{i'}]$ для разных i, i', j и j' . Используя неравенство $(1-s) \leq e^{-s}$, имеем оценку $\mathbf{P} \leq e^{(\ln m \cdot s \cdot e \cdot m \cdot \ln m)} = m^{(1-s \cdot e \cdot m)}$, где s – общая вероятность событий $[f_j = c_i]$ для разных i и j .

Оценим теперь $e \cdot m \cdot s = e \cdot m \cdot (1-1/m)^{m-1} \cdot (1/m) = e \cdot (m-1)^{m-1} / m^{m-1}$. Из известного неравенства $(1+1/(m-1))^{m-1} < e$ следует, что $e \cdot m \cdot s > 1$, откуда получаем $m^{(1-e \cdot m \cdot s)} \rightarrow 0$ при $m \rightarrow \infty$, что завершает доказательство теоремы.

Рассуждение, приведенное в этом доказательстве, аналогично случаю оценки хвоста вероятности в задаче «о собирателе купонов» (см., например, [Феллер, 1984]).

2.2. Обучающие выборки средней плотности

Таро Сакураи доказал неполиномиальность среднего числа кандидатов в случайной обучающей выборке, порождаемой серией испытаний Бернулли с вероятностью успеха $p = 1/2$. Мы немного упростим его рассуждения, рассмотрев квадратную выборку $k = n$, хотя в [Sakurai, 2021] размер обучающей выборки $k \times n$ выбирался случайно с условием, что $k \cdot n$ равно постоянной длине серии испытаний Бернулли.

Рассмотрим серию испытаний Бернулли $(\delta_{i,j})$, где $1 \leq i \leq n$ и $1 \leq j \leq n$, с вероятностью успеха $p = 1/2$, заполняющую обучающую выборку $I \subseteq O \times F$, где $O = \{o_1, \dots, o_n\}$ и $F = \{f_1, \dots, f_n\}$.

Каждый кандидат $\langle A, B \rangle$, $A \subseteq O$, $B \subseteq F$ с $|A| = m$ и $|B| = l$ может перестановками строк и столбцов быть преобразован к виду, показанному в табл. 3. Здесь каждый столбец из « $\langle \rangle$ » и каждая строка из « \leftarrow » должны содержать хотя бы один 0, а прямоугольник из « \leftarrow » может содержать любые биты.

Вероятность появления кандидата $\langle A, B \rangle$, $A \subseteq O$, $B \subseteq F$ с $|A| = m$, $|B| = l$, $|O| = |F| = n$ равна $\mathbf{P}[\langle A, B \rangle] = C_n^m \cdot C_n^l \cdot 2^{-m \cdot l} \cdot (1-2^{-m})^{n-l} \cdot (1-2^{-l})^{n-m}$. Сомножитель $C_n^m \cdot C_n^l$ соответствует условиям $|A| = m$, $|B| = l$, а $2^{-m \cdot l} \cdot (1-2^{-m})^{n-l} \cdot (1-2^{-l})^{n-m}$ – что верхний левый $m \times l$ прямоугольник состоит из 1, а каждый столбец из « $\langle \rangle$ » и каждая строка из « \leftarrow » должны содержать хотя бы один 0.

Случайная обучающая выборка средней плотности

$O \setminus F$	f_1	...	f_l	f_{l+1}	...	f_n
o_1	1	1	1			
\vdots	1	1	1			
o_m	1	1	1			
o_{m+1}	-	-	-	+	+	+
\vdots	-	-	-	+	+	+
o_n	-	-	-	+	+	+

Для кандидата $\langle A, B \rangle$, $A \subseteq O$, $B \subseteq F$ с $|A| = m$, $|B| = l$ введем индикатор $\langle \langle A, B \rangle | I \subseteq O \times F \rangle$, принимающий значение 1, если для обучающей выборки $I \subseteq O \times F$ пара $\langle A, B \rangle$ является кандидатом, и 0 в противном случае.

Тогда число кандидатов для $I \subseteq O \times F$ вычисляется по формуле $|H(I \subseteq O \times F)| = \sum_{\langle A, B \rangle} \langle \langle A, B \rangle | I \subseteq O \times F \rangle$. Ясно, что $|H(I \subseteq O \times F)|$ является целочисленной случайной величиной.

Так как среднее суммы равно сумме средних, то среднее число кандидатов вычисляем по формуле $\mathbf{E}|H| = \sum_{\langle A, B \rangle} \mathbf{P}[\langle \langle A, B \rangle | I \subseteq O \times F \rangle]$, где суммирование проходит по всем $\langle A, B \rangle$, $A \subseteq O$, $B \subseteq F$ с $|A| = m$, $|B| = l$.

Неполиномиальная нижняя оценка будет получена при замене суммы на единственное слагаемое с $m = l = \log_2 n = \text{lb } n$, т.е. $\mathbf{E}|H| \geq C_n^{\text{lb } n} \cdot C_n^{\text{lb } n} \cdot n^{-\text{lb } n} \cdot (1-1/n)^{n-\text{lb } n} \cdot (1-1/n)^{n-\text{lb } n}$.

Лемма 1. $(1 - 1/n)^{n-\text{lb } n} = e^{-1} \cdot (1 + o(1))$ при $n \rightarrow \infty$.

Доказательство. $(1 - 1/n)^{n-\text{lb } n} = \exp\{(n - \log_2 n) \cdot \ln(1 - 1/n)\} = \exp\{-(n - \log_2 n) \cdot \sum_k 1/(k \cdot n^k)\} = \exp\{-(n - \log_2 n) \cdot 1/n\} \cdot (1 + o(1)) = e^{-1} \cdot (1 + o(1))$.

Лемма 2. $(C_n^{\text{lb } n})^2 = \exp\{2 \cdot \log_2 n \cdot \ln n - 2 \cdot \log_2 n \cdot \ln \log_2 n + 2 \cdot \log_2 n \cdot \ln \log_2 n\} \cdot (1 + o(1))/2\pi$ при $n \rightarrow \infty$.

Доказательство представляет собой простые вычисления с использованием формулы Стирлинга.

Теорема 2. При $n \rightarrow \infty$ среднее число кандидатов в случайной средней плотности обучающей выборке ограничено снизу $\mathbf{E}|H| \geq n^{\text{lb } n}$.

Доказательство использует предыдущие рассуждения, обе леммы и тривильное преобразование $\exp\{\log_2 n \cdot \ln n\} = n^{\text{lb } n}$.

Закключение

Основные технические результаты работы – доказательства того факта, что в случайных обучающих выборках (средней и большой плотности) число кандидатов (сходств обучающих примеров) с большой вероятностью не является полиномиальным.

Однако главное значение работы – установление того факта, что ДСМ-метод автоматического порождения гипотез [Финн и др., 2009] не может быть применим к обучающим выборкам большого размера (Big Data), а требуется переход к ВКФ-методу¹ [Виноградов, 2022] алгебраического машинного обучения, который использует вероятностные алгоритмы, основанные на цепях Маркова.

Благодарности. Автор благодарит своих коллег по ВЦ им. А.А. Дородницына ФИЦ ИУ РАН за поддержку и полезные дискуссии. Особая признательность выражается Л.А. Якимовой за совместную работу, обсуждения и поддержку.

Список литературы

- [Виноградов, 2017] Виноградов Д.В. Предельная вероятность порождения случайного сходства при наличии контр-примеров // Научная и техническая информация, Сер. 2. – 2017. – № 2. – С. 17-19.
- [Виноградов, 2022] Виноградов Д.В. Алгебраическое машинное обучение: упор на эффективность // Автоматика и телемеханика. – 2022. – № 6. – С. 5-23.
- [Кузнецов, 1999] Кузнецов С.О. О некоторых вопросах анализа понятий // Научная и техническая информация. Сер. 2. – 1999. – № 1–2. – С. 57-61.
- [Феллер, 1984] Феллер В. Введение в теорию вероятностей и ее приложения. В 2-х т. Т. 1: пер. с англ. – М.: Мир, 1984.
- [Финн, 1983] Финн В.К. О машинно-ориентированной формализации правдоподобных рассуждений в стиле Ф. Бэкона–Д.С. Милля // Семиотика и информатика. – 1983. – Вып. 20. – С. 35-101
- [Финн и др., 2009] Финн В.К., Аншаков О.М. (ред.) ДСМ-метод автоматического порождения гипотез: Логические и эпистемологические основания. – М.: URSS, 2009.
- [Sakurai, 2021] Sakurai Taro. On formal concepts of random formal contexts // Information Sciences. – 2021. – Vol. 578. – P. 615-620.

¹ назван так в честь проф. В.К. Финна.

УДК 004.93

СИСТЕМА ДИСТАНЦИОННОГО ИССЛЕДОВАНИЯ АРХЕОЛОГИЧЕСКИХ ПАМЯТНИКОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ГЛУБОКОГО МАШИННОГО ОБУЧЕНИЯ*

А.В. Вохминцев (*vav@csu.ru*)

Челябинский государственный университет, Челябинск

А.В. Мельников (*mav@csu.ru*)

Югорский НИИ информационных технологий, Ханты-Мансийск

М.А. Романов (*std.romanov.ma@gmail.com*)

Южно-Уральский государственный университет, Челябинск

А.В. Шицелов (*a_shicelov@ugrasu.ru*)

Югорский государственный университет, Ханты-Мансийск

Данная работа посвящена проблеме обнаружения и исследования археологических памятников на основе методов машинного обучения, картографирования и геофизики. В статье рассмотрена методология для картирования и съемки археологических объектов в виде 2D и 3D моделей, предложены методы машинного обучения на основе CNN для обнаружения и классификации и археологических объектов.

Ключевые слова: остаточные нейронные сети, классификация и сегментация данных, археологический памятник.

Введение

В настоящее время в археологической науке приобретает все большее значение применение методов точных и естественных наук. В этой связи актуальными становятся методы дистанционного зондирования больших участков Земли, создания 2D и 3D моделей больших объектов – древних поселений, могильников, культовых памятников. Для Южного Зауралья одной из ключевых исторических эпох, в которые происходило освоение этой территории, является бронзовый век, традиционно связывающийся с

* Работа выполнена при финансовой поддержке РФФ (проект 23-11-20007).

миграциями древних индоевропейцев, создавших здесь пласт археологических памятников, самым известным из которых является поселение Аркаим. В работе предложена методология дистанционного исследования археологических памятников с использованием методов картографирования и глубокого машинного обучения, которая апробирована на территории археологических памятников в районе п. Степное и Левобережное. Методы картографирования и геофизики отличаются разнообразием, например, часто используется сканирование поверхности памятников с использованием магнитометров и георадаров, применяется маловысотная аэромагнитная съёмка с БПЛА с камерами видимого диапазона, LiDAR и ToF [Lambers et al., 2019]. При решении задачи обнаружения и классификации археологических объектов и артефактов можно выделить три подхода: поиск специальных характеристик объектов для распознавания; статистические методы распознавания; с использованием обучаемых нейронных сетей. Методы, основанные на поиске характеристик объекта, выделяют из изображения объекта инвариантные к аффинным преобразованиям и яркостным изменениям характеристики объекта распознавания, такие как гистограммы ориентированных градиентов [Dalal et al., 2005], [Vokhmintcev et al., 2016], [Vokhmintcev et al., 2020b], инвариантные к масштабированию особые точки [Lowe, 1999], затем, размерность пространства характеристик максимально уменьшается и вектора характеристик сравниваются между собой. Процесс выделения и сравнения характеристик в лучших по качеству обнаружения алгоритмах этого типа обладает высокой вычислительной сложностью, алгоритмы плохо формализуются и не контролируют качество распознавания. Статистические методы распознавания объектов на основе линейной или нелинейной фильтрации [Manzurv et al., 2012] используют для распознавания объекта корреляционную функцию, которая позволяет установить корреляцию детектируемого объекта с анализируемым изображением на основе сравнения с заданным порогом. [Yaroslavsky et al., 1993]. Известно, что корреляционный детектор может произвести неправильное детектирование объекта и отождествить его с различными объектами фона изображения. Для преодоления этого недостатка было предложено множество эвристических методов и алгоритмов [Kumar et al., 2005]. В последнее время большой интерес при решении задачи обнаружения и классификации объектов связан с разработкой методов, использующих нейронные сети [Duda et al., 2000]. В настоящий момент глубокое машинное обучение применяется для анализа трехмерных данных, например, известны работы по классификации 3D объектов по точкам [Charles et al., 2017], [Komarichev et al., 2019] и графам [Smith et al., 2019]. В данной работе рассмотрены методы для автоматизации обнаружения археологических памятников на 2D и 3D моделях с помощью методов машинного обучения, картографирования и геофизики.

1. Признаки дешифрирования археологических памятников

В качестве источников данных в работе использованы:

Аэрофотоснимки, сделанные с 1954 по 1957 г. для целей лесного и сельского хозяйства и 60-х- 80-х годов для целей геологии и геодезии. Снимки сделаны в масштабе 1:14000 с большой разрешающей способностью для Кизильского района Челябинской области;

Результаты дистанционного зондирования Земли со спутников Sentinel-2 (с 2015 г.), Landsat 4-9 (с 1985 г.), Ресурс-П (с 2013 по 2021 г.), Канопус-В (с 2013 по 2023 г.);

Данные тахеометрической съемки, полученные с использованием тахеометра Trimble 3300 (Elta R55).

В дальнейшем планируется использовать:

Цифровые модели рельефа (Digital Terrain Model, DTM) и поверхности (Digital Surface Model DSM), созданные в результате поисковых экспедиций с 2006 г. по н.в;

Данные наземной магнитометрической съемки;

Данные сканеров глубины LiDAR, установленных на БПЛА.

До появления машинного обучения данные дистанционной разведки обрабатывались специалистами интерпретаторами, которые производили ручное выделение археологических памятников на снимках (рис. 1).

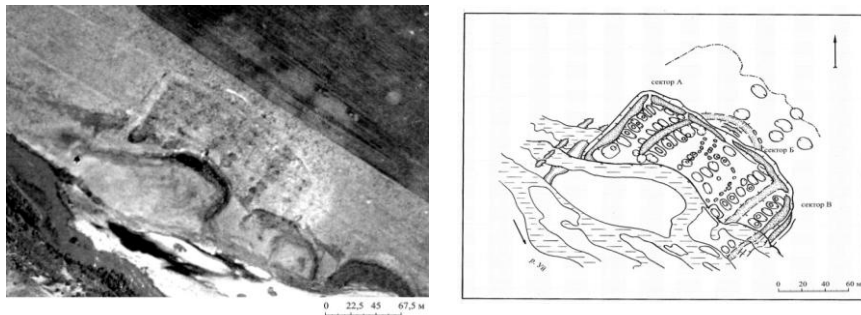


Рис. 1. Результаты дешифрирования археологического памятника Степное: (а) аэрофотоснимок (1986 г.), (б) схема поселения по результатам дешифрирования

Главные проблемы при автоматизированной обработке данных были связаны с ростом их объема и неоднозначностью интерпретации. Археолог допускал множество ошибок, пропускал различные особенности культурного слоя. Применение методов машинного обучения позволит автоматизировать этот процесс и преодолеть указанные недостатки. Нами бы-

ли формализованы признаки дешифрирования археологических объектов (табл. 1.), характерные для памятников бронзового века в Челябинской области [Зданович и др., 2003], [Носкевич, et al., 2003].

Таблица 1

Признаки дешифрирования археологических объектов

Название	Форма и размеры	Признаки дешифрирования
Грунтовый курган бронзового века (покрытый дерном) К1	Круглая или овальная. Диаметр от 7-8 до 25-30 м, высота в пределах 0,5-1 м.	Располагается на ровных участках береговых террас и коренного берега, в пределах 500-700 м от берега реки. Округлый или слегка овальный бугор, в ряде случаев окруженный ровиком. Перепад контраста между более светлым бугром и более темным ровиком вокруг него или окружающей поверхностью.
Грунтовый курган бронзового века (распаханная поверхность) К2	Круглая или овальная. Диаметр как правило от 7-8 до 25-30 м, высота меньше 0,5 м.	Располагается на ровных участках береговых террас и коренного берега, в пределах 500-700 м от берега реки. Осветленные округлы и овальные участки на фоне более темной пашни. Расположены группами или вытянутыми неровными линиями, ориентированными в широтном направлении.
Грунтовый курган железного века (покрытый дерном) К3	Круглая или овальная. Диаметр от 15-20 до 80-100 м, высота может составлять до 3-5 м.	Располагается на водораздельных возвышенностях, вершинах холмов и гор, на расстоянии 500-700 м от берега реки. Округлый (овальный) бугор, окруженный рвом, с провалом от осевшей погребальной камеры и/или грабительского вкопа в центре насыпи. Перепад контраста между светлым бугром и более темным провалом в центре насыпи, а также темным ровиком вокруг кургана или поверхностью.
Каменный курган средневековья / железного века К4	Круг или эллипс. Диаметр от 5-10 до 20-25 м, высота 1 м.	Располагается на водораздельных возвышенностях, вершинах холмов и гор, на расстоянии более 500-700 м от берега реки. Четкие контуры, более контрастное изображение, чем грунтовый курган.
Грунтовый или каменный курган раннего средневековья с "усами" К5	Головные курганы в количестве от 1 до 3.	Расположены на водоразделах, террасах рек и склонах. Грунтовые дорожки выглядят как белый равномерный тон. Каменные дорожки, дискретное мелкопятнистое изображение. Диаметр курганов до 30 м, длина "усов" до 200 м. "Усы" отходят на восток или юго-восток.

Погребальные культовые комплексы М1	Цепь круглых курганов, святилище в форме подковы или гантели, вал.	Расположены на пологих склонах водоразделов или в седловидных впадинах между вершинами сопок. Все элементы выделяются светлым тоном, по краям курганов присутствуют понижения темного тона. Размеры в пределах 30-40 м.
Могильники с каменными оградками эпохи средневековья М2	Овал, квадрат или прямоугольник, размеры от 2 x 1,5 м до 3-5 м x 3-5 м.	Расположены на пологих склонах водоразделов, имеют дискретное мелкопятнистое изображение. Выглядят как многочисленные решетки в виде пятен белого равномерного тона с более темным центром, Каменные оградки имеют более контрастное изображение.
Укрепленное поселение эпохи бронзы (с линейной или концентрической планировкой) П-1	Круг, эллипс, квадрат. Диаметр округлых от 120 до 200 м, размеры прямоугольных до 150-200 м.	Расположены на надпойменных террасах на расстоянии до 200 м от русла реки. Линейные п. расположены рядами параллельно руслу реки и окружены более светлым возвышением вала, а вокруг него – более темным по тону углублением рва. Концентрические п. образованы 1 или 2 вписанными окружностями, окруженными светлым валом с темной полосой от рва.
Неукрепленное поселение бронзового века П-2	Овал или прямоугольник, размеры до 35 x 30 м.	Одна или несколько цепочек жилищных впадин длиной до 200 м, вытянутых вдоль русла реки. Выделяются более темным тоном впадины и более светлым тоном окружающей ее обваловки или зольника.

1.2. Методология исследования археологических памятников

В работе предложена методология для картирования и съемки археологических объектов на основе методов машинного обучения, методов геофизики и картографирования (рис. 2). Входные данные можно разделить на две группы: 2D данные в виде изображений и 3D данные в виде облаков точек. Рассмотрим методы обработки этих данных. Для обнаружения и классификации объектов на двумерных ДТМ часто применяются различные классификаторы, такие как сопоставление шаблонов, объектно-ориентированный анализ изображений (OBIA) или сверточные нейронные сети (CNN). Остаточные нейронные сети ResNet [He et al., 2016] зарекомендовали себя как высоко точные инструменты обнаружения различных видов археологических объектов и артефактов, в данной работе использованы остаточные нейронные сети ResNet 50 и ResNet 152 для решения поставленной задачи, произведен сравнительный анализ сетей с помощью метрик F1 и IoU. Сеть ResNet принимает на вход квадратные изображения с набором признаков, производит анализ локальных особен-

ностей объектов с использованием множественных операций свертки, после чего выходные данные сети формируются с помощью полносвязного слоя, который представляет собой метрический классификатор на базе многослойной сети прямого распространения (MLP) и одной радиально-базисной сети (RBF). В дальнейшем планируется провести исследования с использованием архитектуры xResNet. Для повышения качества работы ReNet используется предварительная обработка изображений. Были использованы следующие методы обработки снимков из коллекции:

Шаг 1. Обрезка краев. Часть снимков имеют неровный край или рамку белого цвета с одной или нескольких сторон кадра. В работе убирается 1,5% сверху (снизу) снимка и 1% с левой (правой) стороны снимка.

Шаг 2. Унификация направления снимков. Для нормализации снимков важно, чтобы они были одинаково ориентированы в пространстве, все снимки были расположены по направлению сторон света.

Шаг 3. Восстановление изображения и удаление шума. Часть снимков содержит небольшую облачность или мелкие артефакты. Была использована быстрая билатеральная фильтрация снимков на основе параллельной декомпозиции на пространственные фильтры.

Шаг 4. Повышение контрастности. Для повышения контрастности снимков, содержащих полутоновые изображения, например, в виде пятен белого равномерного тона был использован алгоритм на основе многокомпонентного бета-распределения яркости изображения [Буланова et al., 2022]. Данный алгоритм помогает визуально улучшить детекцию микро-рельефа археологического объекта, его мелких деталей и текстуры, что позволит улучшить качество получаемых моделей обучения.

Шаг 5. Нормализация размера снимка. Так как на вход ResNet должны поступать только квадратные изображения, то вначале исходные снимки разрезаются на части в виде квадратов определенного размера с левого верхнего угла, при этом может сложиться ситуация, когда останутся области снизу и справа, которые не соответствуют квадрату. В этом случае производится сравнение размера снимка с порогом, если размер части снимка по ширине или высоте меньше этого порога, то он исключается из рассмотрения, в противном случае к нему добавляется некоторое количество темных пикселей по краю изображения.

Шаг 6. Нормализация входных данных. Исходные снимки сохраняются в различных форматах изображений, при этом значение каждого пикселя в отдельном канале расположено в диапазоне от 0 до 65535. Для повышения производительности алгоритма машинного обучения на основе сети ResNet в работе используется нормализация значений в диапазоне $[0,1]$.

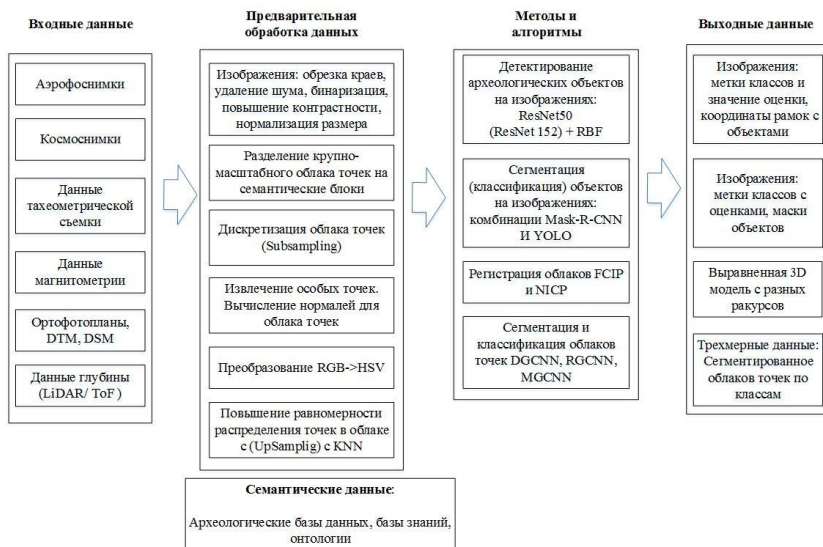


Рис. 2. Архитектура системы исследования археологических памятников

Для обнаружения, классификации и сегментации объектов на трехмерных DTM в форме облака точек используют множество методов, которые можно разделить на не прямые и прямые методы. Первая группа MVCNN [Su, 2015] и VoxNet [Maturana, 2015] методов использует последовательность двумерных изображений облака точек, полученных с разных углов обзора, которая обрабатывается CNN, далее ее выход в виде попиксельной семантической разметки проецируется обратно в трехмерное пространство. Основной недостаток методов данной группы связан с тем, такие сети позволяют хорошо обрабатывать только простые полигональные модели. Вторая группа методов PointNet [Charles et al., 2017], DGCNN [Wang et al., 2018], RGCNN [Te et al., 2018] использует множественные операции свертки на динамических графах, позволяет хорошо обрабатывать локальные и глобальные особенности анализируемых объектов. Данные методы имеют два недостатка, один из которых связан с зависимостью от способа получения облака точек, а другой связан с ограничением размерности анализируемого облака точек и формой анализируемых объектов – перечисленные ранее методы показывают высокую точность для объектов выпуклой формы. Вместе с тем в археологии, при дешифрировании трехмерных моделей DTM и ортофотопланов приходится иметь дело с облаками точек, которые были получены в результате реконструкции 3D модели с разных ракурсов [Вохминцев и др., 2020а], а

сами модели имеют большую размерность и содержат много плоских объектов, таких как погребальные культовые комплексы, могильники, жилища и более мелкие их детали, например, ограды и дорожки. Для преодоления этих недостатков нами разработана архитектура MGCNN (Multimodal Graph CNN, рис. 3) на основе DGCNN и RGCNN, которая в настоящее время проходит апробацию на различных коллекциях данных, получены результаты с преимуществом данной сети над известными сетями на основе GCNN для коллекции данных ModelNet 40 и 3D коллекции на основе тахеометрической съемки археологических памятников Челябинской области. Для повышения качества работа сети MGCNN используется предварительная обработка облаков точек.

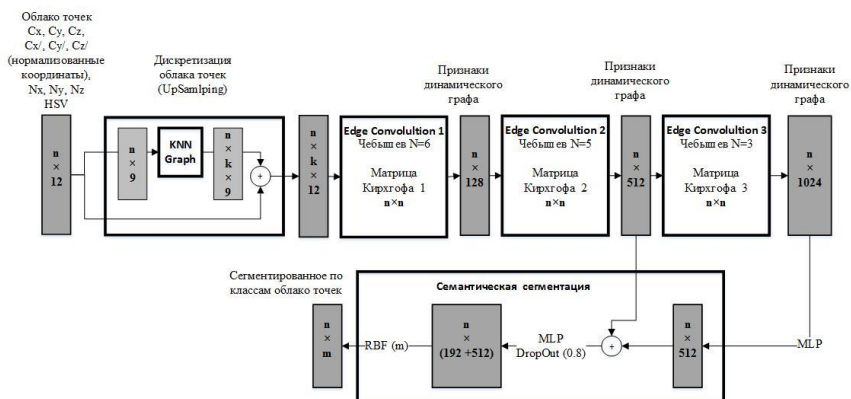


Рис. 3 Архитектура сверточной нейронной сети MGCNN

Шаг 1. Трехмерная DTM модель в автоматизированном режиме разделяется на семантические блоки, связанные с конкретным археологическим объектом или группой объектов.

Шаг 2. Дискретизация облаков точек (Down Sampling). Данные полученные с датчиков глубины (LiDAR или ToF) часто представляют из себя плотные облака точек, которые содержат избыточную информацию об объектах, в работе проводится дискретизация с шагом 0.01.

Шаг 3. Извлечение особых точек, вычислений нормалей в облаке. На основе полученных данных происходит формирование мультимодального вектора признаков, состоящего их координат объектов и их нормализованных координат, нормалей и данных о цвете точки в облаке. Дополнительно осуществляется преобразование каналов RGB в каналы HSV, так как последние являются независимыми в отличие от RGB.

Шаг 4. Повышение равномерности распределения в облаке точек (UpSampling) на основе метода kNN ($k=5$) позволяет детализировать локальную окрестность вокруг особых точек за счет увеличения количества информации об этой окрестности, что позволяет добиться лучшего качества сегментации трехмерного облака точек.

Перечисленные в разделе 1.1. источники данных использованы в качестве коллекции данных для обучения различных методов машинного обучения на основе CNN при обнаружении и классификации и археологических объектов на 2D и 3D моделях. В дальнейшем планируется расширить обучающую выборку с археологическими объектами в 5-6 раз с помощью геометрических преобразований и методов предобучения сверточных нейронных сетей Mask R-CNN.

1.3. Компьютерное моделирование

В данном разделе представлены и обсуждены результаты компьютерного моделирования. Оценим точность и сходимость предложенного метода детектирования на примере аннотированной коллекции данных об археологических памятниках бронзового века в Челябинской области. Для тестов применялся компьютер на базе Intel Core i7 с графическим процессором, обучение проводилось в течении 10 эпох, размер обучающей выборки составил – 274 кадра, тестовой выборки – 35 кадров.



Рис. 4 Результаты детектирования объектов при помощи Res Net 50 + RBF

На рис. 4 представлены результаты детектирования археологических объектов грунтовых курганов (К-1) и укрепленных поселений (П-1) в районе. п. Левобереженое. Из табл. 2. видно, что предложенный метод классификации археологических объектов позволяет хорошо детектировать на изображении укрепленные поселения (класс П1), каменные курганы раннего железного века или средневековья (класс К4), в то время как

грунтовые курганы (К1) идентифицируется с низким качеством, при их классификации возникает много ошибок первого и второго рода – искомый объект ложно классифицируется с фоном и наоборот в следствие неоднозначных признаков дешифрирования (табл. 1.) и наличия шумов различной природы на изображениях. Проведенные эксперименты показали, что архитектура ResNet 152 не имеет преимуществ по сравнению с ResNet 50, а использование радиально-базисной сети RBF в качестве метрического классификатора в связке с ResNet 50 позволяет улучшить точность классификации. На рис. 5 представлены оценки качества детектирования объектов на основе комбинации Res Net 50 + RBF в зависимости от числа эпох.

Таблица 2

Точность классификации с метриками F1 и IoU

Название сети	Класс объекта (метрика)					
	П-1 (F1)	П-1 (IoU)	К-1 (F1)	К-1 (IoU)	К-4 (F1)	К-4 (IoU)
ResNet 50	0.74	0.71	0.38	0.42	0.63	0.59
ResNet 152	0.73	0.69	0.39	0.42	0.61	0.58
ResNet-50 + RBF	0.83	0.76	0.44	0.47	0.76	0.74
ResNet-152 +RBF	0.78	0.74	0.43	0.42	0.75	0.73

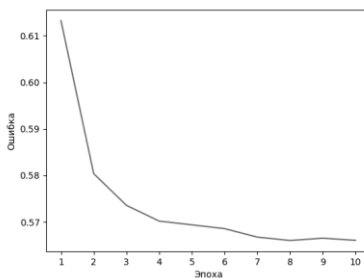
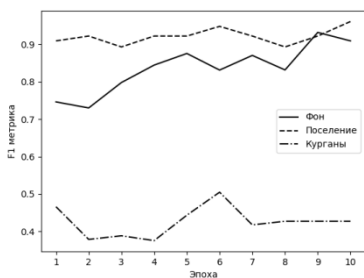


Рисунок 5 Качество метода детектирования: (а) точность метода по метрике F1 по группам классов для валидационной выборки, (б) оценка сходимости алгоритма

Заключение

В работе предложена методология для исследования археологических объектов на основе методов машинного обучения, методов геофизики и картографирования. Для детекции и классификации данных предложены

модифицированные архитектуры CNN: для 2D данных на основе ResNet, для 3D данных на основе GCNN. Были представлены алгоритмы для предварительной обработки двумерных и трехмерных данных, которые позволяют повысить качество работы CNN. В работе выделены классы археологических объектов бронзового века и признаки их дешифрирования, на основе которых создана обучающая выборка данных. Для аэрофотоснимков и космоснимков было проведено компьютерное моделирование с использованием известных архитектур остаточных сетей ResNet 50 и ResNet 152 и разных вариантов реализации полносвязного слоя, полученные результаты были обсуждены и обозначены преимущества предлагаемой архитектуры сети на базе ResNet 50 + RBF.

Список литературы

- [Буланова и др., 2022] Буланова Ю.А., Самандаров И.Р., Душатов Н.Т., Миратов З.М. Исследования методов повышения контраста мкм снимков // *Oriental Renaissance: Innovative, educational, natural and social sciences*. – 2022. – № 2 (10).
- [Вохминцев и др., 2020а] Вохминцев А.В., Мельников А.В. Метод навигации и составления карты в трехмерном пространстве на основе комбинированного решения вариационной подзадачи точка–точка *icp* для аффинных преобразований // *Информатика и ее применения*. – 2020. – Т. 14. – Вып. 1. – С. 101-112.
- [Зданович и др., 2003] Зданович Г.Б., Батанина И.М., Левит Н.В., Батанин С.А. Степь-лесостепь. Кизильский район // *Археологический атлас Челябинской области*. Вып. 1. – Челябинск: Южно-Уральское книжное изд-во, 2003.
- [Носкевич и др., 2003] Носкевич В.В., Федорова Н.В., Петров Ф.Н., Батанина Н.С. Реконструкция плана поселения эпохи бронзы Левобережное (Южный Урал, Россия) // *Волжская Археология*. – 2021. – № 3.
- [Charles et al., 2017] Charles R.Q., Su H., Kaichun M., Guibas L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation // *Proc. of Conference on Computer Vision and Pattern Recognition (CVPR), USA, 2017*.
- [Dalal et al., 2005] Dalal N., Triggs B. Histograms of Oriented Gradients for Human Detection // *Computer Vision and Pattern Recognition*. – 2005. – No. 1.
- [Duda et al., 2000] Duda R.O., Hart P.E., Stork D.G., *Pattern Classification*. – 2nd ed. Wiley— Interscience, 2000.
- [He et al., 2016] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), USA, 2016*.
- [Komarichev et al., 2019] Komarichev A., Zhong Z. A-CNN: Annularly Convolutional Neural Networks on Point Clouds [Электронный ресурс] // *Электрон. жур.* – 2019. – <http://arxiv.org/abs/1904.08017> (дата обращения: 01.04.2023).
- [Kumar et al., 2005] Kumar B.V.-K.V., Mahalanobis A., Juday R.D. *Correlation pattern recognition*. – UK: Cambridge Univ Press, 2005.
- [Lambers et al., 2019] Lambers K., Verschoof-van der Vaart W.V., Bourgeois Q.P.G. Integrating Remote Sensing, Machine Learning, and Citizen Science in Dutch Archaeological Prospection // *Remote Sensing*. – 2019. – No. 11 (7).

- [Lowe, 1999] Lowe D.G. Object recognition from local scale-invariant features // Proc. Int. Conference on Computer Vision, Greece. – 1999.
- [Manzurv et al., 2012] Manzurv T., Zeller J., Serati S. Optical correlator based target detection, recognition, classification, and tracking // Applied Optics. – 2012. – No. 51.
- [Maturana et al., 2015] Maturana D., Scherer S. VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition // Proc. of International Conference on Intelligent Robots and Systems (IROS), Germany, 2015.
- [Smith et al., 2019] Smith E.J., Fujimoto S., Romero A. GEOMETrics: Exploiting Geometric Structure for Graph-Encoded Objects [Электронный ресурс]. Электрон. жур. – 2019. – URL: <http://arxiv.org/abs/1901.11461> (дата обращения: 01.04.2023).
- [Su et al., 2015] Su H., Maji S., Kalogerakis E., Learned-Miller E. Multi-view Convolutional Neural Networks for 3D Shape Recognition // Proc. of International Conference on Computer Vision (ICCV). Chili, 2015.
- [Te et al., 2018] Te G., Hu W., Guo Z., Zheng A., Guo Z. RGCNN: Regularized Graph CNN for Point Cloud Segmentation // MM: Proc. of the 26th ACM international conference on Multimedia. South Korea, 2018.
- [Vokhmintcev et al., 2016] Vokhmintcev A.V., Sochenkov I.V., Kuznetsov V.V., Tikhonkikh D.V. Face recognition based on matching algorithm with recursive calculation of local oriented gradient histogram // Doklady Mathematics. – 2016. – No. 466 (3).
- [Vokhmintcev et al., 2020b] Vokhmintcev A., Timchenko M. The new combined method of the generation of a 3d dense map of environment based on history of camera positions and the robot's movements // Acta Polytechnica Hungarica. – 2020. – No. 17 (8).
- [Wang et al., 2018] Wang Y., Sun. Y., Liu Z., Sarma S.E., Bronstein M. M., Solomon J. M. Dynamic Graph CNN for Learning on Point Clouds // ACM Transactions on Graphics. – 2018. – No. 38 (5).
- [Yaroslavsky, 1993] Yaroslavsky L.P. The theory of optimal methods for localization of objects in pictures // Progress in Optics. – 1993. – No. 32.

УДК 004.896

ИНТЕЛЛЕКТУАЛЬНАЯ МОДЕЛЬ КОМПЛЕКСНОЙ ТЕХНОЛОГИЧЕСКОЙ СИСТЕМЫ ПЕРЕРАБОТКИ РУДНЫХ ОТХОДОВ*

М.И. Дли (*midli@mail.ru*)

А.Ю. Пучков (*putchkov63@mail.ru*)

Филиал Национального исследовательского университета «МЭИ»
в г. Смоленске, Смоленск

В работе предложена цифровая интеллектуальная модель сложной технологической системы переработки рудных отходов, состоящей из двух технологических линий. В основе модели лежит применение двух глубоких рекуррентных нейронных сетей, каждая из которых служит цифровой копией соответствующей линии, входящей в состав системы. Представлены результаты имитационного эксперимента, которые показали высокую точность прогноза, формируемого интеллектуальной моделью для переменных, характеризующих технологический процесс переработки рудных отходов.

Ключевые слова: глубокие нейронные сети, регрессионные модели.

Введение

Проблема переработки отходов руды на горно-обогатительных комбинатах является актуальной для многих стран, включая Россию. Значительные объемы рудных отходов захламляют большие, прилегающие к комбинатам, территории, подвергаются выветриванию с образованием мелкодисперсной пыли, содержащей различные токсичные вещества, которые могут нанести вред окружающей среде и здоровью людей. Поэтому одним из приоритетных направлений государственной политики Российской Федерации в области обращения с промышленными отходами, записанным в Федеральном Законе «Об отходах производства и потребления» (в редакции, введенной в действие с 01.01.2015), указывается использование наилучших доступных технологий их переработки и утилизации.

* Работа выполнена в рамках государственного задания, проект № FSWF-2023-0012.

В связи с этим актуальной исследовательской задачей является разработка новых технологий вторичного использования рудных отходов, а также создание для них современного информационного обеспечения, учитывающего последние достижения в области теории управления, искусственного интеллекта, вычислительной техники. Внедрение результатов таких исследований позволило бы не только уменьшить экологическую нагрузку для регионов, где расположены горно-обогатительные комбинаты, но и повысить экономическую эффективность производства, конкурентоспособность за счет более полного использования богатого химико-минералогического состава рудных отходов при их вторичной переработке [Meshalkin et al., 2017].

Объектом исследования представляемой работы выступала комплексная многостадийная химико-энерготехнологическая система переработки мелкодисперсных рудных отходов (integrated system for processing fine ore waste, ISPFOW). В ее состав входят две технологические линии (Process line, PL): одна использует процесс агломерации рудного сырья и содержит агломерационную машину (линия PL1), вторая – процесс окомкования рудного сырья в грануляторе с дальнейшим обжигом сырых окатышей в многокамерной обжиговой машине конвейерного типа (линия PL2). Выбор технологической линии зависит от результатов анализа гранулометрического состава отходов и в данном исследовании полагаем, что такой выбор уже осуществляется на этапе входного контроля [Puchkov et al., 2021].

В конце каждой линии находятся рудно-термические печи для плавления подготовленного из отходов рудного сырья материала. На выходе печи получают готовый продукт переработки, например, в случае отходов апатит-нефелиновых руд – это газообразный желтый фосфор, служащий ценным сырьем химической промышленности и производства удобрений для нужд сельского хозяйства [Dli et al., 2020].

Целью исследования являлась разработка структуры интеллектуальной модели ISPFOW, а также программного обеспечения, реализующего эту модель. В процессе исследования решались задачи обзора существующих концепций и подходов к построению программных моделей производственных технологических систем, разработка такой структуры, а также создание программного обеспечения для ее апробации.

Научную новизну результатов исследования составляет предложенная структура интеллектуальной модели ISPFOW, а их практическая значимость заключается в разработанном программном обеспечении, реализующем данную модель.

Представленные в работе результаты апробации созданного с учетом предложенной структуры интеллектуальной модели программного обеспечения показали высокую точность прогноза переменных, описывающих технологический процесс переработки рудных отходов.

1. Материалы и методы

При выборе концепции разработки интеллектуальной модели ISPFOW было выявлено, что современные производства характеризуются широким внедрением систем промышленного Интернета вещей (IIoT), следствием чего является значительное увеличение объемов поступающей информации о состоянии технологической системы, глубокий анализ которой способен обеспечить качественно новые возможности для ее управления [Boyes et al., 2018]. Также, значительное внимание в IIoT сейчас уделяется цифровым платформам, обеспечивающих информационной инфраструктурой подключенные к цифровым сетям промышленные устройства для сбора, обработки генерируемых данных, предоставления услуг на основе данных [Arnold et al., 2022], [Chernova et al., 2019]. В основе многих платформ лежит концепция «цифрового двойника», выступающего как центральная точка отсчета для получения информации о технологическом процессе и связанном с ним оборудовании, что может повысить эффективность обслуживания и сопровождения процессов в течение жизненного цикла технологической системы [Hickey et al., 2023]. Цифровой двойник, представляя собой высокоточную цифровую копию реальной сложной системы, позволяющую моделировать различные режимы функционирования и естественного износа («старения») оборудования, дает необходимую информацию для более высокого уровня иерархии управления сложной технической системой – ситуационному управлению, основанному на идеях теории искусственного интеллекта [Поспелов, 1986].

Большие объемы циркулирующей в IIoT информации способствуют тому, что значительное количество исследований и разработок цифровых двойников базируется на аппарате глубоких искусственных нейронных сетей [Gosavi et al., 2022], [VanDerHorn et al., 2021], [Chen et al., 2022]. Это мощный аппарат, обладающий высокими прогностическими способностями и, как все нейронные сети, позволяет дообучать уже созданные нейронные сети на новых ситуациях, которые возникают в процессе функционирования реальной системы, обеспечивая постоянное поддержание актуальных настроек сети и ее высокую точность, что особенно актуально для высокоответственных технических систем [Glukhikh et al., 2021]. Учитывая эти обстоятельства, было принято решение о создании интеллектуальной комплексной многостадийной химико-энерготехнологической системы переработки мелкодисперсного рудного сырья на основе аппарата глубоких искусственных нейронных сетей.

Информация, поступающая от первичных датчиков, установленных в контуре интеллектуального мониторинга ISPFOW, представлена в форме временных рядов, содержащих упорядоченные по времени показания контрольно-измерительной аппаратуры и первичных датчиков, расположенных

в различных зонах технологического процесса. Данные, являющиеся результатом видеоконтроля и акустического контроля, после преобразования специализированными алгоритмами, также приобретают форму числовых последовательностей, отражающих оценку состояния объекта в заданный момент времени, см. например, [Hossain et al., 2021], [Ambadekar et al., 2020]. Поэтому в структуре разрабатываемой модели не рассматриваются особенности указанных преобразований.

При выборе базовой архитектуры применяемых для построения интеллектуальной модели ISPFOW нейронных сетей было учтено, что ее основное назначение заключается в прогнозировании переменных технологического процесса на заданный промежуток времени t_d , так как точное и эффективное прогнозирование множественных нагрузок и диагностика неисправностей имеет большое значение для управления работой ISPFOW.

В последние годы различные типы нейронных сетей широко использовались для предиктивной аналитики различных систем и обеспечивали хорошие результаты. В [Zhou et al., 2021] предложена модель прогнозирования электрической нагрузки энергетической системы на основе двусторонних генеративно-состязательных сетей и трансфертного обучения, отличающаяся высокой точностью прогноза. В [Shoumao et al., 2020] применена вейвлет-нейронная сеть для краткосрочного прогнозирования нагрузки. Модель прогнозирования на основе рекуррентной нейронной сети (recurrent neural network, RNN) с вейвлет-пакетной декомпозицией и анализом автокорреляции данных во временных рядах в различных частотных диапазонах разработана в [Zhu et al., 2020].

Особенностью ISPFOW является то, что в ней протекают сложные, многосвязные тепло-гидравлические, гидродинамические, физико-химические процессы, описываемые нелинейными зависимостями. Их аналитическое описание представляет собой сложные системы интегро-дифференциальных уравнений, полученных с учетом ряда упрощений и допущений, что сказывается на точности получаемых результатов. Кроме того, решение таких систем базируется на применении численных методов расчета (применении конечно-разностных схем), решение которых с взаимной увязкой граничных и начальных условий представляет собой отдельную трудную задачу. Изменение внешних условий, параметров технологической среды также требует перерасчет получаемых решений [Puchkov et al., 2022]. Указанные аспекты, усложняющие практическое использование аналитического подхода к расчету переменных технологического процесса, еще раз подтверждают целесообразность применения нейронных сетей, в частности, RNN с долгой кратковременной памятью (long short-term memory, LSTM), обладающих хорошей способностью прогнозирования нелинейные временные ряды [Wang et al., 2022].

ISPFOW представляет собой многоагрегатную систему, поэтому при формировании обучающих наборов данных для RNN необходимо обеспечить съем данных на входах и выходах агрегатов. Общая схема расположения точек съема информации и состава входных и выходных последовательностей отражена рис. 1 (физические агрегаты на рисунке имеют серую заливку, блоки, выполняющие обработку информации заливки не имеют).

ISPFOW содержит две технологические линии PL1 и PL2, поэтому создается две RNN, отличающихся размерностью входных и выходных последовательностей. Для сети RNN1 входная последовательность $RNN1_{out}(k-d)$ для момента дискретного времени t_k образуется конкатенацией векторов (для сокращения записи вместо t_k будем записывать k):

$$RNN1_{inp}(k-d)=[S_{inp}(k-d), S_{out}(k-d), FI_{inp}(k-d), FI_{out}(k-d)],$$

где $S_{inp}(k-d)$, $S_{out}(k-d)$ – входного и выходного векторов переменных технологической линии агломерационной машины; $FI_{inp}(k-d)$, $FI_{out}(k-d)$ – входного и выходного векторов переменных рудно-термической печи №1; d – интервал прогноза, формируемый блоками delay, показанными на рис. 1.

Выходная последовательность $RNN1_{out}$ сети RNN1 для момента времени t_k образуется конкатенацией векторов $S_{out}(k)$ и $FI_{out}(k)$. Аналогично формируются обучающие последовательности для PL2, состав и назначение элементов которых понятно из рис. 1:

$$RNN2_{inp}(k-d)=[G_{inp}(k-d), G_{out}(k-d), M_{inp}(k-d), M_{out}(k-d), F2_{inp}(k-d), F2_{out}(k-d)],$$

$$RNN2_{out}(k)=[G_{out}(k), M_{out}(k), F2_{out}(k)].$$

При формировании последовательностей учитывалось, что выходные переменные предыдущего агрегата являются входными для другого, поэтому они принимали участие при построении его входных последовательностей. Временное смещение входных и выходных последовательностей обучает нейронную сеть прогнозированию переменных на d интервалов дискретизации Δt .

В свою очередь для момента времени t_{k-d} каждая из составляющих в $RNN1_{inp}$ и $RNN2_{inp}$ содержит наборы компонент соответствующей переменной, взятые на историческую глубину от момента t_{k-d} еще на r отчетов. Таким образом, составляющая $S_{inp}(k-d)=[s(k-d-r), s(k-d-r+1), s(k-d-r+2), \dots, s(k-d)]$. Аналогично количество компонент задается для других составляющих в $RNN1_{inp}$ и $RNN2_{inp}$.

Последовательности $RNN1_{inp}$, $RNN1_{out}$, $RNN2_{inp}$ и $RNN2_{out}$ подвергаются нормализации в блоке Data Preprocessing для исключения влияния на процесс обучения нейронных сетей разных масштабов переменных и постоянных составляющих.

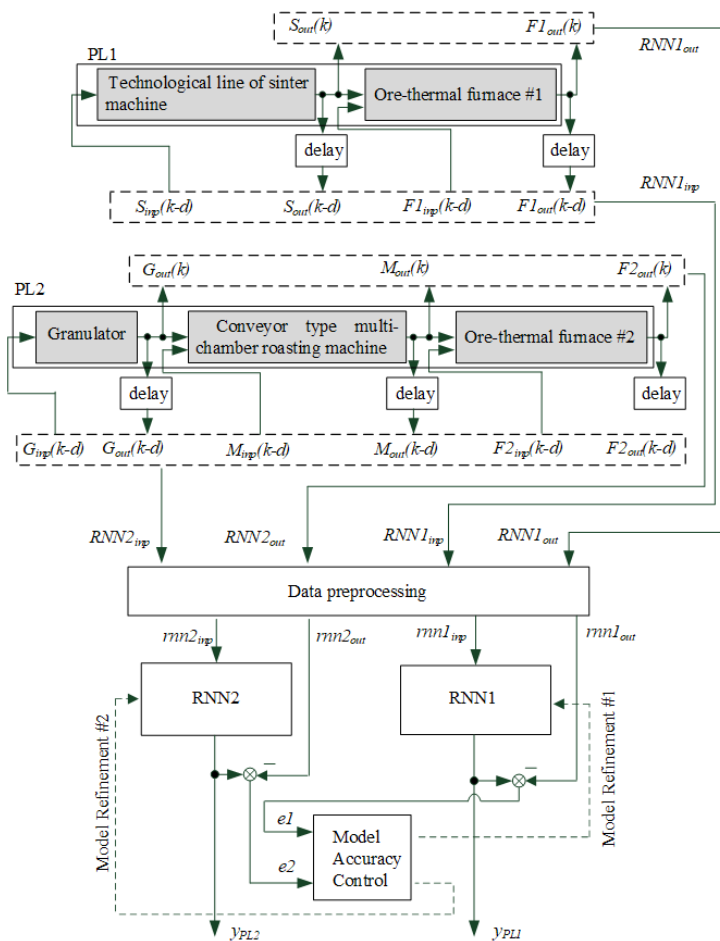


Рис. 1. Структура интеллектуальной модели ISPFOW

В результате работы интеллектуальной модели ISPFOW на ее выходе формируются вектора y_{PL1} и y_{PL2} , содержащие выходные переменные соответствующих технологических линий. При работе модели осуществляется постоянный контроль текущей величины усредненных, за определенный интервал времени, ошибок сетей $e1$ и $e2$, что осуществляется в блоке Model Accuracy Control и имеет вид:

$$e1_i = \frac{1}{K_a - 1} \sum_{i=0}^{K_a} (y_{PL1_i} - rnn1_{out_i})^2,$$

где K_d – количество дискретных значений t_k , за которое проводится временное усреднение ошибок; $i=1,2, \dots, i_{c1}$, i_{c1} – количество компонент вектора ошибки $e1$.

Аналогично вычисляются ошибки компонент вектора ошибки $e2$. Если какая-либо из ошибок компонент превышает предельную заданную величину, то соответствующая сеть обучается с учетом новых данных, поступающих в процессе функционирования ISPFOW (на рис. 1 эти процессы показаны пунктирными стрелками, отходящими от блока Model Accuracy Control) и так как сети фактически предобучены на старых данных, то новый процесс обучения будет требовать значительно меньших ресурсов. Такая организация подстройки нейронных сетей позволяет интеллектуальной модели быть постоянно актуальной, реализуя один из главных принципов построения цифрового двойника системы – обеспечивать адаптацию под новые условия ее функционирования [Dli et al., 2020].

2. Результаты и обсуждение

Проведение натурных экспериментов на ISPFOW на данном этапе исследований не проводилось ввиду необходимости разработки аппаратного сопряжения информационной и физической частей системы, что планируется на следующих этапах. Сейчас для генерации обучающих наборов данных выполнялись модельные эксперименты с использованием имитации ISPFOW в среде динамического моделирования Simulink. Передаточные функции агрегатов были аппроксимированы по разгонным характеристикам преобразования вход-выход (по каналу переменных, характеризующих преобразование массы) в форме апериодических звеньев первого порядка с запаздыванием [Dli et al., 2021]. Было проведено моделирование в течении 30000 отсчетов времени. Для синхронизации отсчетов был взят одинаковый для всех переменных ISPFOW интервал Δt , равный одной десятой от минимальной постоянной времени агрегатов, найденной по их разгонным характеристикам.

Для каждой из технологической линии из общего количества отсчетов по 25000 было выделено под обучающую и по 5000 – под тестирующую выборки. Программа разрабатывалась в среде MatLAB, обладающей набором инструментов глубокого обучения и удобным средством визуального проектирования нейронных сетей Deep Network Designer. Точность работы нейронных сетей зависит от их архитектуры, поэтому в модельном примере для сравнения использовались два популярных вида сетей: для PL1 применялась RNN с 128 рекуррентными блоками GRU (Gated Recurrent Units – упрощенный вариант сетей LSTM), а для PL2 – RNN с 128 блоками LSTM (архитектуры сетей отражены на рис. 2).

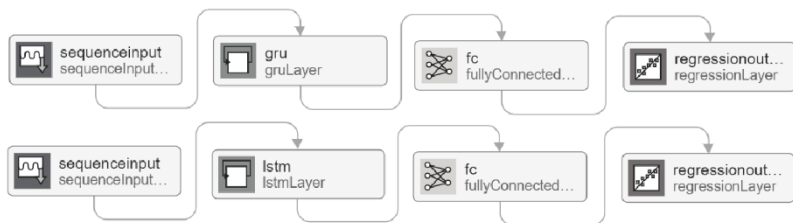


Рис. 2. Архитектура применяемых RNN

Обучение нейронных сетей проводилось в течении 50 эпох, использовалась метрика RMSE (Root Mean Square Error – среднеквадратичная ошибка). В табл. 1 представлены значения RMSE, достигаемые в процессе обучения при разных интервалах прогноза.

Таблица 1

d	ARIMA PL1/PL2	GRU	LSTM
1	0.040/0.052	0.007	0.006
2	0.087/0.099	0.007	0.006
5	2.651/3.512	0.013	0.009
10	--	0.015	0.012
20	--	0.024	0.018
50	--	0.089	0.045

Для сравнения в табл. 1 приведены ошибки, получаемые при применении для прогноза выходных переменных технологических линий на заданный интервал d простой (по сравнению с нейросетевыми) интегрированной модели авторегрессии–скользящего среднего ARIMA(2,1,1), где 2, 1 и 1 – порядки авторегрессионной, интегрированной и скользящего среднего частей модели соответственно. Прочерки в табл.1 для модели ARIMA обозначают то, что величина ошибки очень большая и применение модели не имеет смысла.

Анализ результатов моделирования позволяет сделать вывод, что применение сети LSTM дает лучшую точность на больших интервалах прогноза, что делает ее предпочтительной для применения в алгоритмическом обеспечении интеллектуальной модели ISPFOW для каждой технологической линии.

Заключение

В результате проведенных исследований предложена структура интеллектуальной модели комплексной многостадийной химико-энерготехнологической системы переработки мелкодисперсных рудных отходов, в

основе которой лежит применение глубоких рекуррентных нейронных сетей. Также создано программное обеспечение, реализующее эту модель, представлены результаты ее применения на синтетическом наборе данных, сгенерированных с помощью динамической модели химико-энерготехнологической системы в среде Simulink. Результаты работы программы показали высокую точность прогноза переменных модели Simulink химико-энерготехнологической системы на больших интервалах, что может свидетельствовать о целесообразности дальнейшего совершенствования программного обеспечения на основе выбранной нейросетевой концепции с целью ее внедрения на реальном прототипе системы.

Список литературы

- [Ambadekar et al., 2020] Ambadekar P.K., Choudhari C.M. CNN based tool monitoring system to predict life of cutting tool // SN Appl. Sci. – 2020. – No. 2, 860. – <https://doi.org/10.1007/s42452-020-2598-2>.
- [Arnold et al., 2022] Arnold L., Jöhnk J., Vogt F. et al. IIoT platforms' architectural features – a taxonomy and five prevalent archetypes // Electron Markets. – 2022. – 32.
- [Boyes et al., 2018] Boyes H., Hallaq B., Cunningham J., Watson, T. The Industrial Internet of Things (IIoT): An analysis framework // Computers in Industry. – 2018. – 101.
- [Chen et al., 2022] Chen J., Meng C., Gao Y. et al. Multi-fidelity neural optimization machine for Digital Twins // Struct Multidisc Optim. – 2022. – 65, 340. – <https://doi.org/10.1007/s00158-022-03443-2>.
- [Chernova et al., 2019] Chernova G., Khalin V., Yurkov A. Digitalization and its impact on the development of Russia // Prikladnaya Informatika=Journal of Applied Informatics. – 2019. – Vol. 14, No. 5 (83). – DOI: 10.24411/1993-8314-2019-10035.
- [Dli et al., 2020] Dli M., Puchkov A., Meshalkin V., Abdeev I., Saitov R., Abdeev R. Energy and Resource Efficiency in Apatite-Nepheline Ore Waste Processing Using the Digital Twin Approach // Energies. – 2020. – 13, 5829. – <https://doi.org/10.3390/en13215829>.
- [Dli et al., 2021] Dli M., Puchkov A., Vasiliev A., Kirillova E., Selyavskiy Y., Kulyasov N. Intelligent Control System Architecture for Phosphorus Production from Apatite-Nepheline Ore Waste // Energies. – 2021. – 14 (20): 6469. – <https://doi.org/10.3390/en14206469>.
- [Glukhikh et al., 2021] Glukhikh I., Karyakin Yu., Glukhikh D. Neural network architecture for outputting solutions in dangerous situations at a complex technological facility // Prikladnaya informatika=Journal of Applied Informatics. – 2021. – Vol.16, No. 5. – DOI: 10.37791/2687-0649-2021-16-5-99-107.
- [Gosavi et al., 2022] Gosavi A., Le V.K. Maintenance optimization in a digital twin for Industry 4.0 // Ann Oper Res. – 2022. – <https://doi.org/10.1007/s10479-022-05089-1>.
- [Hickey et al., 2023] Hickey B., Gachon C., Cosgrove J. Digital Twin – A Tool for Project Management in Manufacturing // Procedia Computer Science. – 2023. – Vol. 217.
- [Hossain et al., 2021] Hossain M.S., Taheri H. In-situ process monitoring for metal additive manufacturing through acoustic techniques using wavelet and convolutional neural network (CNN) // Int J Adv Manuf Technol. – 2021. – No. 116. – <https://doi.org/10.1007/s00170-021-07721-z>.

- [Shoumao et al., 2020] Shoumao L., Jiaying Q., Xingzhen B. et al. IPSO-WNN based short-term load prediction for integrated power systems // Electric Measurement and Instrumentation. – 2020. – 57 (09). – DOI: 10.19753/j.issn1001-1390.2020.09.016.
- [Zhou et al., 2021] Zhou D., Ma S., Hao J., et al. An electricity load forecasting model for Integrated Energy System based on BiGAN and transfer learning // Energy Reports. – 2020. – Vol. 6.
- [Meshalkin et al., 2017] Meshalkin V.P., Puchkov A.Yu., Dli M.I., Bobkov V.I. Generalized Model for Engineering and Controlling a Complex Multistage Chemical Energotechnological System for Processing Apatite-Nepheline Ore Wastes // Theoretical Foundations of Chemical Engineering. – 2019. – No. 53.
- [Puchkov et al., 2021] Puchkov A., Dli M., Lobaneva E., Fedulov Y. Monitoring the Granulometric Composition on the Basis of Deep Neural Networks // Theory and Engineering of Dependable Computer Systems and Networks. DepCoS-RELCOMEX 2021. Advances in Intelligent Systems and Computing. – Vol. 1389. – Springer, Cham, 2021. – https://doi.org/10.1007/978-3-030-76773-0_34.
- [Puchkov et al., 2022] Puchkov A., Dli M., Prokimmov N., Shutova D. Multilevel algorithms for evaluating and making decisions on the optimal control of an integrated system for processing fine ore raw materials // Prikladnaya informatika=Journal of Applied Informatics. – 2022. – Vol. 17 (6). – DOI: 10.37791/2687-0649-2022-17-6-102-121.
- [VanDerHorn et al., 2021] VanDerHorn E., Mahadevan S., Digital Twin: Generalization, characterization and implementation // Decision Support Systems. – 2021. – Vol. 145.
- [Wang, et al., 2022] Wang S., Takyi-Aninakwa P., Jin S., Yu C., Fernandez C., Stroe D., An improved feedforward-long short-term memory modeling method for the whole-life-cycle state of charge prediction of lithium-ion batteries considering current-voltage-temperature variation. // Energy. – 2022. – Vol. 254, Part A. – 124224. – <https://doi.org/10.1016/j.energy.2022.124224>.
- [Zhang et al., 2020] Zhang L., Shi J., Wang L., Xu C. Electricity, Heat, and Gas Load Forecasting Based on Deep Multitask Learning in Industrial-Park Integrated Energy System // Entropy. – 2020. – 22, 1355.
- [Zhu et al., 2020] Zhu L., Wang X., Ma J., Chen Q., & Qi X. Short-term load forecast of integrated energy system based on wavelet packet decomposition and recurrent neural network // Electr. Power Constr. – 2020. – Vol.4 1.
- [Поспелов, 1986] Поспелов Д.А. Ситуационное управление: теория и практика. – М.: Наука, 1986.

УДК 004.838.2

ПОДХОД К ОБНАРУЖЕНИЮ АТАК НА СИСТЕМЫ МАШИННОГО ОБУЧЕНИЯ С ИСПОЛЬЗОВАНИЕМ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНОЙ СЕТИ

И.В. Котенко (*ivkote@comsec.spb.ru*)

И.Б. Саенко (*ibsaen@comsec.spb.ru*)

Санкт-Петербургский Федеральный исследовательский центр РАН,
Санкт-Петербург

О.С. Лауга (*laos-82@yandex.ru*)

Государственный университет морского и речного флота
им. адмирала Макарова, Санкт-Петербург

Н.А. Васильев (*vasn2020@mail.ru*)

В.Е. Садовников (*bladimir1998@mail.ru*)

Военная академия связи, Санкт-Петербург

В работе проведена оценка имеющихся атак на системы машинного обучения (МО) с учетом стадий обучения и эксплуатации информационной инфраструктуры. Предложена систематизация атак на системы МО, выделены классы атак, которые в особенности актуальны и опасны для систем МО. Также проведено исследование потенциальных мер противодействия атакам на системы МО, позволившее представить подход к их защите, базирующийся на одновременном применении технологий Neural-Cleance и Jpeg-сжатия, а также генеративно-сопоставительной сети. Проведена экспериментальная оценка предложенного подхода, подтвердившая его результативность.

Ключевые слова: фреймворки для защиты систем машинного обучения, машинное обучение, нейронные сети, искусственный интеллект, генеративно-сопоставительная сеть.

Введение

В настоящее время технология машинного обучения (МО) находит достаточно широкое распространение во многих областях, связанных с решением сложных задач, таких как обнаружение вредоносных программ, распознавание объектов, самоуправляемые транспортные средства и дру-

гие [Котенко и др., 2014], [Котенко и др., 2018], [Kotenko et al., 2018]. Однако, несмотря на очевидные достоинства, методы МО имеют один существенный недостаток. Они являются уязвимыми к преднамеренным атакам на обучающие и тестовые наборы данных, а также параметры конфигурирования моделей МО, что может привести к значительному искажению результатов применения систем МО [Саенко и др., 2022], [Иванов и др., 2023]. Таким образом, возникает острая необходимость в разработке эффективного подхода к обнаружению атак на системы МО, учитывающего состояние исследований в этой области.

Анализ известных работ, исследующих атаки на системы МО и способы защиты от них, показывает, что наибольший интерес в этой области вызывают генеративно-состязательные сети (ГСС), технология нейронного очищения (Neural-Cleanse) и технология Jpeg-сжатия (Jpeg-Compression). Каждый из этих способов показал свою эффективность для отдельных типов атак на МО. При этом наибольший интерес вызывает применение ГСС, так как сети такого типа появились сравнительно недавно [Белокопытов, 2022], [Abdel-Basset et al., 2023]. ГСС представляют собой нейросети, состоящие из двух частей. Первой частью является генератор, который формирует новые примеры данных. Вторая часть – дискриминатор, который определяет, насколько эти примеры данных реалистичны. Преимущество ГСС перед другими видами нейронных сетей проявляется в возможности генерировать новые данные на основе существующих, расширяя набор данных, а также учитывать структуры и корреляции между различными признаками данных при генерации новых примеров, что позволяет создавать более реалистичные примеры, чем другие методы генерации данных [Wang X. et al., 2022].

Интеграция этих способов, как предполагается, должна существенно повысить эффективность обнаружения атак на системы МО. Реализация этого подхода является основной целью настоящей работы.

Вклад и новизна статьи заключаются в следующем. Она приводит результаты анализа достаточно большого количества мер противодействия атакам на системы МО. Эти результаты позволили предложить подход к защите систем МО, основанный на одновременном использовании технологий Neural-Cleanse и Jpeg-сжатия, а также ГСС, в которой присутствуют классификатор и автокодировщик. Практическая реализация этого подхода обеспечила проведение его экспериментальной оценки, которая продемонстрировала более высокую эффективность предлагаемого подхода по сравнению с традиционными.

1. Анализ работ в области защиты от атак на системы МО

Известны различные подходы к классификации атак на системы МО и методов защиты от них. Так, в работе [Vorobeychik et al., 2018] предлагается строить эту классификацию на основе следующих трех параметров: времени проведения атаки, информации, доступной для злоумышленника, и цели злоумышленника. В работе [Kantarcioglu et al., 2016] полагается, что все наиболее распространенные стратегии защиты от атак на МО заключаются в удалении вредоносных данных из обучающего набора и в улучшении алгоритма обучения для защиты от вредоносных обучающих данных. Анализируя эти работы, по-нашему мнению, можно выделить следующие основные методы защиты.

Первый метод защиты – состязательная тренировка [Duddu et al., 2018]. Данный метод защиты допускает три подхода к защите:

1. Ансамблевое состязательное обучение, которое включает в себя возмущенные входные данные, переданные из других предварительно обученных моделей. Данный подход эффективен против таких атак, как «Быстрый градиентный метод» (Fast Gradient Sign Method, FGSM), «Итеративный градиентный метод» (Iterative Gradient Sign Method, IGSM) и «Итеративный метод наименьших классов» (Iterative Least-Likely Class Method).

2. Расширенное состязательное и виртуальное состязательное обучение, являющееся средством упорядочивания текстового классификатора путем стабилизации функции классификации. Данный подход эффективен против атак типа «Замена, вставка и удаление слов» (Replacement, Insertion and Removal of Words).

3. Обучение современному распознаванию речевых эмоций на смеси чистых и состязательных примеров. Этот подход к защите эффективен против атак типа «Шум реального мира» (Real-World Noise).

Второй метод защиты – оборонительная дистилляция [Joseph et al., 2019]. Этот метод защиты подразумевает двукратное обучение модели. Первоначально она обучается с использованием меток истинности с одним основанием. Затем для обучения модели используются вероятностные данные. Данный метод эффективен против FGSM и IGSM.

Третий метод защиты – реконструкция входных данных (анализ многообразия) [Goodfellow et al., 2018]. В данном методе защиты используется модель MagNet, в которой состязательные примеры путем реконструкции преобразуются в чистые данные. Метод аппроксимирует многообразие нормальных примеров и приводит к этому многообразию состязательные примеры. Данный метод правильно классифицирует состязательные примеры, имеющие небольшое возмущение. Он используется для отражения атак FGSM, IGSM, «Глубокий дурак» (DeepFool, DF) и «Атака Карлини и Вагнера» (Carlini and Wagner’s Attack, C&W).

Четвертый метод – применение фреймворка Defense-GAN [Li et al., 2018]. Этот фреймворк использует выразительные возможности генеративных моделей для защиты глубоких нейронных сетей от враждебных атак типа FGSM, IGSM, DF, C&W и «Атака с использованием якобианной карты значимости» (Jacobian Saliency Map Attack, JSMA).

Пятый метод защиты основан на модели обучения с подкреплением [Yuan et al., 2019]. Этот метод использует три подхода к защите:

1. Установка автокодировщика перед входным слоем модели, а затем обучение модели для устранения конфликтных возмущений. Данный подход эффективен против атак типа «Алгоритм жадного поиска» (Greedy Search Algorithm), «Усовершенствованный генетический алгоритм» (Improved Genetic Algorithm) и «Вероятностно-взвешенная значимость слова» (Probability Weighted Word Saliency).

2. Применение архитектуры, использующей для построения надежных нейронных сетей гауссовы байесовские классификаторы. Эта защита используется против таких атак, как FGSM и C&W.

3. Применение совокупности классификаторов со взвешенным или средневзвешенным значением прогноза для повышения устойчивости к атакам. Такой подход к защите используется против атак типа FGSM и «Основной итерационный метод» (Basic Iterative Method, BIM).

Шестой метод – защита от предварительной обработки [Shi et al., 2019]. В данном методе применяются следующие подходы:

1. Использование методов анализа главных компонент, фильтрации нижних частот, сжатия JPEG, определения пороговых значений в целях повышения устойчивости моделей машинного обучения. Этот подход используется для защиты от атак FGSM, IGSM и C&W.

2. Использование операций рандомизации, в частности:

- случайного изменения размера входных изображений;
- случайного заполнения нулями вокруг входных изображений.

Данный подход эффективен для защиты от атак типа FGSM, DF и C&W.

Седьмой метод защиты – обнаружение примеров состязательности [Zhang et al., 2020]. Данный метод имеет три подхода к защите:

1. Сжатие объектов путем уменьшения разрядности цвета каждого пикселя или путем сглаживания выборки с помощью пространственного фильтра. Затем создается двоичный классификатор, в котором признаками являются предсказания целевой модели до и после сжатия входной выборки. Этот подход используется против таких атак, как FGSM, JSMA, BIM и C&W.

2. Использование фреймворка, применяющего признаки качества изображения для различения образцов легитимных и состязательных атак. Такой подход эффективен при защите от атак FGSM, IGSM, JSMA и DF.

3. Многоверсионное программирование, основанное на использовании готовых систем автоматического распознавания речи. Данный подход к защите используется при атаках типа «Целевые аудио-состязательные примеры» (Targeted Audio Adversarial Examples), «Генетические алгоритмы и градиентная оценка» (Genetic Algorithms and Gradient Estimation).

Во многих перечисленных подходах используется технология Neural-Cleanse, которая позволяет нейросети удалять шум из зашумленных сигналов, восстанавливая оригинальный чистый сигнал [Цибулис и др., 2022]. Зашумленный сигнал передается через нейросеть, которая обучается на наборе зашумленных сигналов и соответствующих им чистых сигналов. Нейросеть пытается изучить структуру шума и на основе этого восстанавливает чистый сигнал [Kotenko et al., 2023]. Технология Neural-Cleanse успешно применяется для защиты от бэкдоров [Wang et al., 2022], в ходе которой производится:

1) обнаружение бэкдора (принятие решения о том, что глубокая нейронная сеть заражена бэкдором; если такое случилось, то обнаружение, какой триггер использует бэкдор);

2) выявление бэкдора (определение ожидаемой работы бэкдора и перепроектирование триггера, используемого для атаки);

3) устранение бэкдора (визуализация бэкдора и снижение его эффективности).

Во многих подходах используется JPEG-сжатие, уменьшающая размеры файлов изображений и позволяющая удалять высокочастотные компоненты сигнала внутри квадратных блоков изображения [Shen et al., 2022]. Такая операция эквивалентна выборочному размытию изображения, что помогает устранить аддитивные возмущения. Это положительно влияет на скорость обучения и время работы модели, а также позволяет уменьшить шум, так как сжатие может сократить шумы и артефакты на изображении.

Используя JPEG-сжатие, возможно защитить модель МО от нескольких типов атак, не требуя знаний о модели. Данный подход имеет несколько преимуществ. Во-первых, JPEG – это широко используемый метод кодирования, причем большинство изображений хранятся в этом формате. Большинство операционных систем также имеют встроенную поддержку кодирования и декодирования JPEG-изображений. Поэтому этот этап предварительной обработки может быть применен даже неопытными пользователями. Во-вторых, этот подход не требует знаний о модели МО и об атаке. Поэтому он может применяться в широком диапазоне наборов данных об изображениях.

2. Реализация подхода

Наиболее изученными методами состязательных атак являются FGSM и JSMA, причем FGSM является наиболее распространенным методом [Ren et al., 2019], [McCarthy et al., 2022]. Основными способами защиты от них являются состязательная тренировка и модель обучения с подкреплением.

Предлагаемый подход предполагает два уровня обнаружения – на этапе разработки и на этапе эксплуатации обученной системы. Обнаружение атак на системы МО осуществляется в 7 этапов:

- 1) накопление обучающей выборки и предобработка;
- 2) обучение ГСС;
- 3) проверка обученной сети на наличие уязвимостей с применением Neural-Cleance и Jpeg-сжатия;
- 4) сохранение данных в обученной ГСС при отсутствии уязвимостей или удаление данных из выборки в случае их обнаружения;
- 5) обучение необходимой модели нейронной сети;
- 6) использование обученной модели нейронной сети;
- 7) отображение результатов.

Для обучения модели нейронной сети ГСС использует классификатор (генератор) и автокодировщик (дискриминатор) (рис. 1).

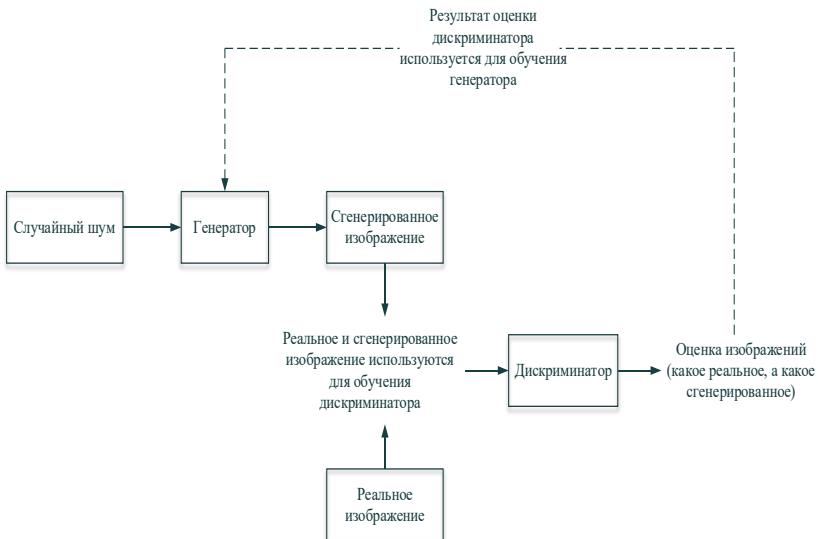


Рис. 1. Блок-схема генеративно-состязательной сети

На вход автокодировщика подаются только достоверные данные. На вход классификатора подаются нормальные данные, данные с уязвимостями и скрытые латентные представления, полученные с автокодировщика после кодирования информации.

Для обучения ГСС была сформирована вычислительная среда, ориентированная на использование графического процессора под операционной системой Ubuntu 20.04. В этой среде были установлены следующие фреймворки и библиотеки: Python 3.8, CUDA Toolkit 10.0.130, cuDNN 7.6.0, TensorFlow 2.1 и Keras.

Обучающая выборка для ГСС состояла из изображений, представляющих дорожные знаки. Набор данных содержал 500 цветных изображений дорожных знаков различных типов, в частности, предупреждающих знаков, знаков ограничения скорости, знаков приоритета, указательных знаков и других (рис. 2). Каждое изображение имело размер 256x256 пикселя и было представлено в формате RGB с 8-битными значениями для каждого цветового канала. Кроме того, обучающая выборка содержала несколько примеров каждого типа знака в разных перспективах, освещении и условиях окружающей среды. Внутри каждой категории присутствовали как достоверные данные (60% выборки), так и данные с уязвимостями (40% выборки).



Рис. 2. Примеры изображений дорожных знаков из обучающей выборки

Подбор гиперпараметров ГСС осуществляется таким образом, чтобы функция потерь при обучении автокодировщика уменьшалась, а точность классификатора росла.

3. Экспериментальная оценка подхода

Экспериментальная оценка предложенного подхода к обнаружению атак на системы МО была проведена в два этапа. Сначала модель МО обучалась с использованием данных с известными уязвимостями. После этого была произведена оценка точности и полноты модели, которая представлена на рис. 3. Точность обнаружения атак составила 0,97, а функция потерь принимает значение 0,03.

```
scores = encoder.evaluate(X_pad, data_last['anomaly'].values, verbose=1)
print('Точность: {}% \nLoss: {}'.format(scores[1]*100, 1 - scores[1]))
```

```
1799/1799 [=====] - 52s 29ms/step - loss: 0.0478 - binary_accuracy: 0.9850
Точность: 96.90268635749817%
Loss: 0.03097313642501831
```

Рис. 3. Оценка точности и полноты модели на известных уязвимостях

Затем был сформирован новый набор данных с ранее не известными уязвимостями. При этом, как оказалось, система допускает ложные срабатывания, однако их количество было небольшим. Так, в эксперименте нейронная сеть отбросила всего 10 запросов.

Обучение нейронной сети осуществлялось на основе алгоритма «обратного распространения», который позволил рассчитать значения весовых коэффициентов таким образом, чтобы ошибка сети была минимальна (табл. 1).

Таблица 1

Номер эпохи	Потери обучения	Потери валидации
0	0.16	0.16
1	0.12	0.10
2	0.09	0.045
3	0.08	0.045
4	0.072	0.062
5	0.078	0.031
6	0.076	0.03
7	0.075	0.06
8	0.077	0.053
9	0.078	0.03
10	0.074	0.059
12	0.076	0.027
14	0.074	0.045
16	0.074	0.073
17	0.078	0.037

Из таблицы видно, что потери при обучении больше, чем потери при проверке производительности (валидации). Это указывает на то, что предложенная модель хорошо работает с обучающими данными и с новыми данными в тестовой выборке.

Проведено сравнение предложенного подхода с методами k -ближайших соседей (kNN) и k -средних частот (kMC), а также подходом MAGAN, взятым из [Wang X. et al., 2022] (рис. 4,а). Автокодировщик ГСС был реализован на основе рекуррентных сетей типа LSTM (Long Short-Term Memory) с различным количеством нейронов в скрытом слое, равным 200, 400 и 800 для различных вариантов сети. На рис. 4,б показаны результаты сравнительной оценки различных вариантов ансамблевой реализации автокодировщика. Ensemble-averaging – это метод, при котором ответы нескольких моделей объединяются путем вычисления среднего значения. Этот метод широко используется для задач регрессии в машинном обучении. Ensemble-voting – это метод, при котором ответы нескольких моделей объединяются путем голосования. Он считается традиционным. Ensemble-proposed – это улучшенный метод ансамблирования, который используется в предложенном подходе. В нем используются дополнительные методы для управления смещениями и дисперсиями моделей, входящих в ансамбль.

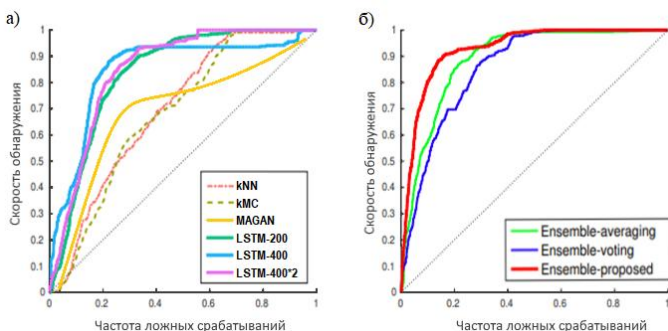


Рис. 4. Результаты сравнения предложенного подхода с другими известными методами

Таким образом, результаты проведенной сравнительной оценки показали, что предложенный подход демонстрирует более высокую скорость обнаружения атак на МО, чем другие схожие подходы, при минимальном количестве ложных срабатываний. Кроме того, предложенный подход более эффективно справляется со сложными задачами классификации атак и дает более лучшие результаты, чем традиционные подходы, основанные на Ensemble-Voting и Ensemble-Averaging.

Заключение

В статье представлены результаты анализа методов защиты от атак на системы МО, на основе которого предложен подход к обнаружению такого типа атак, демонстрирующий высокую скорость обнаружения атак при минимальном количестве ложных срабатываний. Для обучения моделей в предложенном подходе применена ГСС, в которой нейронные сети учатся друг у друга, а в качестве одной из сетей использовался автокодировщик, основанный на LSTM-сетях. Для усиления способности обнаруживать вредоносные данные в предложенном подходе использованы технологии Neural-Cleanse и Jpeg-сжатия.

Результаты исследования показали, что для защиты от атак на МО необходимо определить параметры, используемые в методах МО и нейронных сетях, с наибольшей вероятностью подверженных воздействию. Классификаторы МО и лежащие в их основе алгоритмы должны быть усилены способностью обнаруживать вредоносные данные, не смешивая их с текущими достоверными обучающими данными или не искажая результаты. Необходима база данных с обученными и проверенными моделями, в которых отсутствуют выявленные уязвимости, позволяющие проверить правильность работы системы МО и возможность прогнозировать реализацию атак.

Дальнейшее направление исследований связывается интеграцией предложенного и известных методов защиты от компьютерных атак.

Список литературы

- [Белокопытов, 2022] Белокопытов М.Л. Использование генеративно-состязательных нейронных сетей для расширения обучающих (тестовых) выборок // Журнал радиоэлектроники. – 2023. – № 2.
- [Иванов и др., 2023] Иванов Ю.С., Горькавый М.А., Грабарь Д.М. Анализ устойчивости предиктивных моделей к состязательным атакам в робототехнических комплексах // Ученые записки Комсомольского-на-Амуре государственного технического университета. – 2023. – № 1 (65).
- [Котенко и др., 2014] Котенко И.В., Саенко И.Б. Создание новых систем мониторинга и управления кибербезопасностью // Вестник Российской академии наук. – 2014. – Т. 84, № 11.
- [Котенко и др., 2018] Котенко И.В., Саенко И.Б., Кушнеревич А.Г. Архитектура системы параллельной обработки больших данных для мониторинга безопасности сетей Интернета вещей // Труды СПИИРАН. – 2018. – Вып. 4 (59).
- [Саенко и др., 2022] Саенко И.Б., Лаута О.С., Васильев Н.А., Крибель К.В. Модель угроз систем поддержки принятия решений, функционирующих на основе элементов искусственного интеллекта // Электросвязь. – 2022. – № 05.
- [Цибулис и др., 2022] Цибулис Д.Э., Рагозин А.Н., Даровских С.Н., Кулганатов А.З. Исследование нелинейной цифровой фильтрации сигналов с использованием генеративно-состязательной нейронной сети // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. – 2022. – Т. 22, № 2.

- [**Abdel-Basset et al., 2023**] Abdel-Basset M., Moustafa N., and Hawash H. Generative Adversarial Networks (GANs) // *Deep Learning Approaches for Security Threats in IoT Environments*. – 2023.
- [**Duddu et al., 2018**] Duddu V. A Survey of Adversarial Machine Learning in Cyber Warfare // *Defence Science Journal*. – 2018. – Vol. 68, No. 4.
- [**Goodfellow et al., 2018**] Goodfellow I., McDaniel P., and Papernot N. Making machine learning robust against adversarial inputs // *Communications of the ACM*. – 2018. – Vol. 61.
- [**Joseph et al., 2019**] Joseph A., Nelson B., Rubinstein B., and Tygar J. Overview of Adversarial Machine Learning. – Cambridge University Press, 2019.
- [**Kantarcioglu et al., 2016**] Kantarcioglu M. and Xi B. Adversarial Data Mining: Big Data Meets Cyber Security // *Proc. ACM SIGSAC Conference on Computer and Communications Security, USA*. 2016.
- [**Kotenko et al., 2018**] Kotenko I., Saenko I., Branitskiy A. Framework for Mobile Internet of Things Security Monitoring based on Big Data Processing and Machine Learning // *IEEE Access*. – 2018. – Vol. 6. – P. 72714-72723.
- [**Kotenko et al., 2023**] Kotenko I., Saenko I., Lauta O., Kribel K., and Vasiliev N. Attacks on artificial intelligence systems: classification, the threat model and the approach to protection // *Proc. Sixth International Scientific Conference “Intelligent Information Technologies for Industry”*, Turkey. – 2023.
- [**Li et al., 2018**] Li Y. and Wang Y. Defense Against Adversarial Attacks in Deep Learning // *Applied sciences*. – 2019. – Vol. 9, No. 1.
- [**McCarthy et al., 2022**] McCarthy A., Ghadafi E., Andriotis P., and Legg P. Functionality-Preserving Adversarial Machine Learning for Robust Classification in Cybersecurity and Intrusion Detection Domains: A Survey // *Journal of Cybersecurity and Privacy*. – 2022. – Vol. 2, No. 1.
- [**Ren et al., 2019**] Ren K., Zheng T., Qin Z., and Liu X. Adversarial Attacks and Defenses in Deep Learning // *Engineering*. – 2019. – Vol. 6, No. 3.
- [**Shi et al., 2019**] Shi Y., Zeng H., and Nguyen T.T. Adversarial Machine Learning for Network Security // *Proc. IEEE International Symposium on Technologies for Homeland Security, USA*. – 2019.
- [**Vorobeychik et al., 2018**] Vorobeychik Y., Kantarcioglu M. Categories of Attacks on Machine Learning // *Adversarial Machine Learning. Synthesis Lectures on Artificial Intelligence and Machine Learning*. – Springer, Cham, 2018.
- [**Wang X. et al., 2022**] Wang X., Guo H., Hu Sh., Chang M.-Ch., and Lyu S. GAN-generated Faces Detection: A Survey and New Perspectives. – 2022.
- [**Yuan et al., 2019**] Yuan X., He P., Zhu Q., and Li X. Adversarial Examples: Attacks and Defenses for Deep Learning // *IEEE transactions on neural networks and learning systems*. – 2019. – Vol. 30, No. 9.
- [**Zhang et al., 2020**] Zhang W., Alhazmi Q., and Li C. Adversarial Attacks on Deep-learning Models in Natural Language Processing: A Survey // *ACM Transactions on Intelligent Systems and Technology*. – 2020. – Vol. 11, No. 3.

УДК 004.9

КОГНИТИВНЫЕ ИССЛЕДОВАНИЯ В ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ: НЕЙРОСЕТЕВОЙ И ГЕНЕРАТИВНЫЙ ИНСТРУМЕНТАРИЙ

А.А. Харламов (*kharlamov@analyst.ru*)

Институт высшей нервной деятельности и нейрофизиологии РАН,
Москва

Московский государственный лингвистический университет, Москва

Национальный исследовательский университет

«Высшая школа экономики», Москва

Московский физико-технический институт

(национальный исследовательский университет), Долгопрудный

М.А. Пильгун (*pilgunm@yandex.ru*)

Российский государственный социальный университет, Москва

В статье представлены некоторые результаты сопоставительного исследования преимуществ и ограничений экспертного анализа с использованием нейросетевой технологии TextAnalyst 2.3 и анализа, проведённого с помощью ChatGPT Plus на материале данных социальных медиа, генерированных акторами с различным генезисом, для выявления социального напряжения. Сбор данных проводился в социальных медиа ВКонтакте, Телеграм, Ютуб, Дзен.

Ключевые слова: нейросетевые технологии, ChatGPT Plus, когнитивные исследования, социальные медиа, текстовый анализ.

Введение

Когнитивные исследования в сфере искусственного интеллекта стали особенно актуальными после трансформации медиапространства в связи с распространением искусственных сущностей, производительность которых соответствует или превышает возможности человека. В частности, коммуникативные задачи, которые решают чат-боты типа Playground от OpenAI, Bing, Claude от Anthropic AI, NeevaAI, YouChat, Perplexity Chat, Чат CohereAI, Poe от Quora, HuggingChat и др., свидетельствуют о формировании новых условий взаимодействия агентов с различным генезисом и необходимости расширения когнитивной парадигмы.

Число пользователей ChatGpt, который компания OpenAI представила в ноябре 2022 года уже в марте 2023 года превысил 1 млрд, по данным на 28 апреля 2023 года имел 1,16 млрд пользователей. С февраля 2023 года по март 2023 года сервис увеличил показатели почти на 55%. Запуск ChatGPT привел к увеличению количества посетителей веб-сайта OpenAI на 304 миллиона. Самое большое количество пользователей (11,72%) ChatGPT в США, далее следует Индия (6%) (<https://www.demandsage.com/>). GPT-4, появившаяся в 2023 г. превосходит ChatGPT по показателям и возможностям реализовывать задачи в сфере математики, кодирования, зрения, медицины, юриспруденции, психологии и др. [Bubeck S., et al., 2023].

Показательно, что 53% пользователей не смогли идентифицировать контент, написанный ChatGPT, как контент ИИ. Кроме того, среди многочисленного инструментария для создания контента выделяют в топе Jasper (для написания ИИ), Grammarly (для корректуры и проверки грамматики), CopySmith (для электронной коммерции), INK For All (для SEO). Контент, сгенерированный ИИ, вызывает потребность в инструментах его обнаружения, которые также стремительно развиваются [Ruby, 2023], Chua et al., 2023].

Генеративные инструменты успешно используются в разных типах исследований [Tang, 2023]. Значимость роли ИИ в коммуникативных ситуациях при взаимодействии между людьми уже сегодня очень высока, и, очевидно, еще возрастёт на порядок после того, как он станет бесшовным. Уже после появления ChatGPT стала очевидной необходимость трансформации коммуникативной культуры, в частности, понимание принципа “AI first culture” в организациях и в виртуальном пространстве. Корпоративная культура компании должна измениться, чтобы эффективно использовать потенциал ИИ, который очевидно изменит методы работы и ситуацию на рынке труда.

В аналитике социальных медиа после того, как акторы с искусственным генезисом стали полноправными участниками коммуникативного процесса, потеряло актуальность разделение контента по признаку природы актора, поскольку результаты анализа контента с разделением акторов с различным генезисом и без него, совпадают. Между тем, разработка алгоритмов использования разных типов ИИ как инструментария для проведения когнитивных исследований становится важной задачей.

Целью данной работы было выявление преимуществ и ограничений методов исследования социального напряжения с помощью экспертного анализа с использованием нейросетевой технологии TextAnalyst 2.3. и анализа, проведённого с помощью ChatGPT Plus,

Данные

Эмпирическим материалом для анализа послужил контент, посвящённый строительству Троицкой линии метро (ТЛМ) в г. Москве. Сбор данных проводился 01.10.2022 - 31.03.2023 в социальных медиа Вконтакте, Телеграм, Ютуб, Дзен. Датасет содержит 22 199 085 токенов.

Метод

Экспертный анализ для интерпретации данных проводился с помощью программы TextAnalyst 2.3. Дизайн исследования, представленный на блок-схеме 1, апробирован в ходе многочисленных проектов, которые нашли отображение в публикациях [Kharlamov et al., 2021], [Kharlamov et al., 2023].



Блок-схема. 1. Дизайн исследования с помощью нейросетевой технологии TextAnalyst 2.3

Алгоритм выделения индексов социального стресса и социального благополучия представленный на блок-схеме 2.



Блок-схема 2. Алгоритм выделения индексов социального стресса и социального благополучия

Подробно описание см. [Kharlamov et al. 2023].

Результаты

Аудитория, которая генерировала и обсуждала пользовательский контент (публикации и комментарии), относящийся к строительству ТЛМ предпочитала социальные сети (189 122 465), а также блоги (75 976 537) и каналы мессенджеров (72 374 692) (рис. 1).

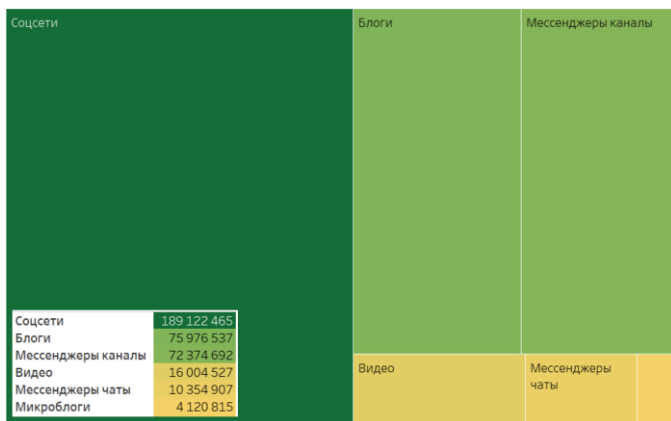


Рис. 1. Аудитория в разных типах источников

Наиболее активно акторы использовали социальную сеть ВКонтакте (144 585 557), а также мессенджер Телеграм (82 729 599) и платформу Дзен (75 952 157) (рис. 2).

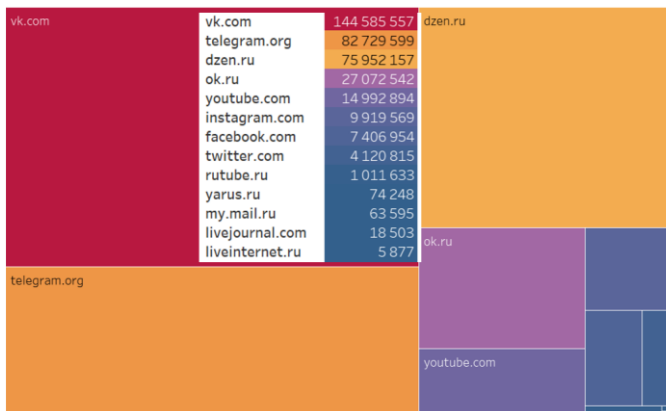


Рис. 2. Аудитория в разных источниках

Анализ тональность в разных типах источников показывает, что в социальных сетях, мессенджерах (каналах и чатах), на видеохостингах и в микроблогах преобладает нейтральный контент, негативный кластер уступает по объёму позитивному. Только в блогах позитивный кластер превышает нейтральный и негативный (рис. 3):

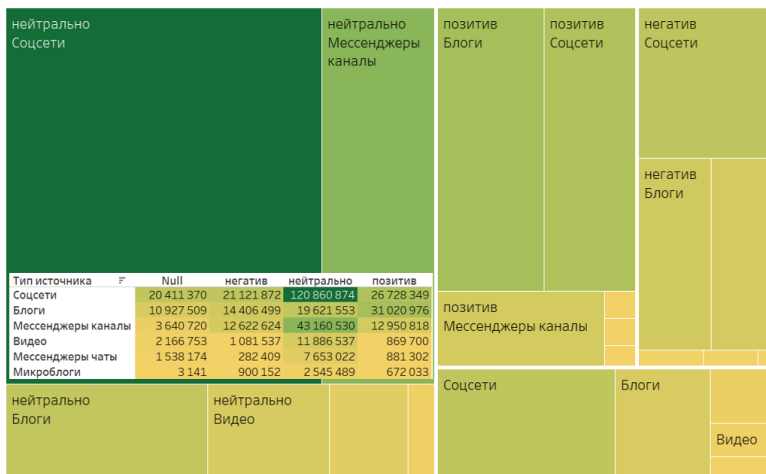


Рис. 3. Тональность в разных типах источников

Результат анализа социального напряжения вокруг строительства первого участка Троицкой линии метро в период с 01 октября 2022 года по 31 марта 2023 года, показал наличие социального стресса с низким индексом – 5,3 и высокий индекс социального благополучия – 31,2 (рис. 4).

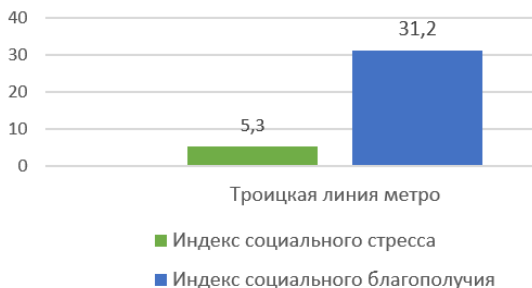


Рис. 4. Индексы социального стресса и благополучия

Проведение контент-анализа, определение тематической структуры, саммаризации, выделение и анализ семантической и ассоциативной сетей позволили определить ключевые темы и сформировать рейтинг контента, а также определить отношение пользователей (одобрения/не одобрения жителей) к строительству первого участка ТЛМ.

При экспертном выделении рейтинг ключевых негативных акцентов составлялся по двум параметрам: по агрессии и негативной тональности.

Рейтинг сообщений с негативной реакцией по количеству аудитории в ВКонтакте:

- закрытие на 4 дня оранжевой линии метро в связи с прокладкой ТЛМ.
- Заявления С. Собянина о строительстве новых линий метро, в частности ТЛМ, вызвали негативную реакцию горожан, поскольку ассоциируются с закрытием старых веток метро, что приводит к транспортным проблемам в активные периоды реализации проекта.
- Сроки завершения строительства, которые, по мнению москвичей затянуты, поскольку во время реализации проекта горожане вынуждены терпеть транспортные проблемы.
- Временное закрытие участка метро от Октябрьской до Новых Черемушек. Попытки московских властей облегчить ситуацию (бесплатные автобусы, чай с печеньем) не спасают ситуацию, которую москвичи переживали крайне болезненно. Все сообщения в негативном кластере социальной сети ВКонтакте, которые собрали максимальное количество аудитории, связаны именно с данным сюжетом.

Рейтинг сообщений с негативными реакциями по количеству аудитории в Телеграме также возглавляют сообщения о закрытии участка оранжевой линии из-за строительства ТЛМ:

- Закрытие части оранжевой ветки, по мнению горожан, осложняет их положение из-за отсутствия четкого информирования жителей о графике закрытия конкретных линий метро в связи со строительством ТЛМ и некорректной коммуникацией городских властей и строителей с москвичами. Кроме того, транспортный коллапс, который возник из-за строительства Троицкой линии, по мнению жителей, может привести к вспышкам ОРВИ в южных районах Москвы.

Результаты исследования показали, что наиболее точным показателем выражения претензий и недовольства горожан является анализ агрессии, поскольку позволяет фиксировать эмоциональное отношение к тем или иным явлениям, событиям и персонам. Несмотря на незначительные по сравнению с негативным кластером количественные показатели, оценка агрессии крайне важна для анализа имплицитных оценок и сформировавшихся стереотипов. Анализ данных позволил выявить:

- недифференцированную агрессию в отношении строительства ТЛМ, которая выражает крайне резкое эмоциональное негативное отношение и характеризует оценку части горожан всей деятельности Московского правительства.
- Резкую реакцию на возможные разрушения старых районов Москвы при строительстве ТЛМ.
- ТЛМ ставится в один ряд с другими новыми линиями метро, которые, по мнению акторов, располагаются по плохо продуманным маршрутам и не будут улучшать транспортную ситуацию в городе.

Рейтинг ключевых негативных акцентов, которые были получены в результате анализа данных негативного кластера с использованием ChatGPT:

1. Движение транспорта и пробки: строительство линии метро может привести к временным перебоям в движении и увеличению заторов в районах, расположенных рядом с проектом.

2. Шум и вибрация: горожане, проживающие вблизи строительных площадок, могут жаловаться на шум и вибрацию, вызванные строительством.

3. Задержки сроков строительства: если строительство линии метро будет отложено, это может вызвать разочарование у граждан, которые надеются использовать новую инфраструктуру.

4. Воздействие на местный бизнес: строительство может повлиять на местный бизнес из-за прекращения пешеходного и автомобильного движения, а также снижения видимости магазинов и заведений.

5. Отвлечение государственных ресурсов: некоторые граждане могут подумать, что государственные ресурсы, вложенные в строительство линии метро, могут быть использованы для других проектов или более неотложных нужд.

6. Возможное повышение налогов: если для финансирования строительства метро потребуется повышение налогов, это может вызвать недовольство граждан.

7. Изменения в городском ландшафте: строительство линии метро может изменить внешний вид и динамику развития отдельных районов города, что может вызвать опасения у местных жителей.

Заключение

В статье представлены некоторые результаты сопоставительного исследования, в результате которого были выявлены преимущества и ограничения для экспертного анализа с использованием нейросетевой технологии TextAnalyst 2.3 и анализа, проведённого с помощью ChatGPT Plus на материале данных социальных медиа, генерированных акторами с различным генезисом.

Анализ данных показал, что наиболее эффективно использовать комбинированный подход с учётом результатов обоих указанных подходов.

В докладе будут подробно изложены полученные результаты, а также сформулированы перспективы развития когнитивных исследований с использованием подобных сервисов.

Список литературы

- [Bubeck et al., 2023] Bubeck S., Chandrasekaran V., Eldan R., Gehrke J., Horvitz E., Kamar E., Lee P., Lee Y.T., Li Y., Lundberg S., Nori H., Palangi H., Ribeiro M.T., & Zhang Y. Sparks of artificial general intelligence: Early experiments with gpt-4 // arXiv preprint arXiv:2303.12712. – 2023.
- [Chua et al., 2023] Chua A.E., Isom M., Jarosova R., & Hua D. Distinguishing academic science writing from humans or ChatGPT with over 99% accuracy using off-the-shelf machine learning tools // Cell Reports Physical Science. – 2023. – No. 4 (6). – <https://doi.org/10.1016/j.xcrp.2023.101426>.
- [Kharlamov et al., 2023] Kharlamov A.A., Pilgun M. Perception of the Situation: Social Stress and Well-Being Indices // Lecture Notes in Networks and Systems. – 2023. – No. 544 LNNS.
- [Kharlamov et al., 2021] Kharlamov A.A., Raskhodchikov A. N., Pilgun M. Smart City Data Sensing during COVID-19: Public Reaction to Accelerating Digital Transformation // Sensors. – 2021. – No. 21 (12). – <https://doi.org/10.3390/s21123965>.
- [Ruby, 2023] Ruby D. How To Identify AI Written Content In 2023 (Legit Way).2023. February 6. – <https://www.demandsage.com/how-to-identify-ai-written-content/>.
- [Tang, 2023] Tang J., et al. Semantic reconstruction of continuous language from non-invasive brain recordings // Nat Neurosci. – 2023. – No. 26. – P. 858-866.

Abstracts of Plenary Reports

DESIGN ONTOLOGY: GENESIS AND DEVELOPMENT

N.M. Borgest (*borgest@yandex.ru*)

Samara National University named after S.P. Korolev

The report examines the genesis and development of the ontology of design, which includes: the study of the conceptual apparatus of the subject area, the development of a thesaurus, the analysis of criteria and models of the projected object, the choice of methods and design scenarios, the processing of information about the object. The ontology of design explores and formalizes knowledge about the process of designing new and modernized artifacts.

Keywords: design ontology, subject area, information system.

FROM KNOWLEDGE ENGINEERING TO KNOWLEDGE SCIENCE

V.I. Gorodetsky (*vladim.gorodetsky@gmail.com*)

Joint Stock Company EUREKA, St. Petersburg

The objective of the paper is to briefly analyze modern sources of knowledge for artificial intelligence applications and to analyze the prospects of their further development. It is shown that currently quite mature and widely demanded are methods, algorithms and technologies of knowledge acquisition from data based on machine learning, as well as various variants of knowledge acquisition using digital twin technology. However, as far as expert knowledge is concerned, this source is currently rarely used in practice, although it is in demand in a number of critical, topical and often unique classes of next-generation applications for which experts are the only available source of knowledge. To address the problem of effective access to expert knowledge, intensive research and development in knowledge engineering is needed to raise it to the level of Knowledge Science, which will be able to solve knowledge processing problems of the same scale and complexity that Data Science is currently able to do with big data. The analysis of potential directions of research and development in hypothetical Knowledge Science, as well as some methods, models, algorithms and technologies for processing large volumes of raw fragments of expert knowledge to serve new generation intelligent applications is presented.

Keywords: knowledge, data-driven machine learning, digital twins, expert knowledge, expert knowledge-based machine learning, expert knowledge consistency.

**CONCEPTUAL APPARATUS OF THE DSM-METHOD
FOR AUTOMATED RESEARCH SUPPORT**

V.K. Finn (*ira.finn@gmail.com*)

Federal Research Center “Computer Science and Control”
of the Russian Academy of Sciences, Moscow

The report examines the Conceptual Apparatus of the DSM-method for automated research support that implement a special class of plausible reasoning such that they generate hypotheses about the cause, hypotheses about the prediction of the studied effects and empirical patterns.

Keywords: DSM-method, conceptual apparatus, automated research support.

Abstracts of Sectional Reports

Section 1

Knowledge Engineering

DEVELOPMENT OF AN ONTOLOGY OF THE SUBJECT AREA OF SCIENTIFIC AND EDUCATIONAL ORGANIZATION

N.K. Andrievskaya (*nataandr@yandex.ru*)

T.A. Vasyaeva (*vasyaeva@gmail.com*)

T.V. Martynenko (*umbralada@gmail.com*)

Donetsk National Technical University, Donetsk

The paper describes the approach to the development of the ontology model, in which different methods of its formation were used at different stages of the existence of the ontology. At the initial stage the "basic" ontology model is implemented in expert mode. Further expansion of the ontology was carried out in semi-automatic mode. The following semi-automatic methods have been developed for further development of the ontology: using a corpus of documents, based on electronic dictionaries, using NLP processing of text sources, as well as using the universal cross-domain ontology DBpedia.

Keywords: Ontology, Ontology Learning, SOAT, WordNet, Wiktionary, DBpedia.

ONTOLOGY-DRIVEN KNOWLEDGE GRAPH CONSTRUCTION IN SEMANTIC LIBRARY

O.M. Ataeva (*oli.ataeva@gmail.com*)

V.A. Serebryakov (*serebrvas@gmail.com*)

N.P. Tuchkova (*natalia_tuchkova@mail.ru*)

FRC CSC RAS, Moscow

The paper considers an approach to building a knowledge graph based on an ontological representation of scientific subject areas within a semantic library. The ontology of a digital semantic library defines the data structure of the content of the library. Each data element loaded into the library can be associated with an ontology vertex that determines the position of the data element in the ontology. Based on the ontology links and the links defined at the design stage, you can build a data graph. On the example of the ontology of the LibMeta semantic library, the technology of forming the knowledge graph of modern applications in mathematics is discussed. The problems of filling a graph, embedding in a graph, extracting links and nodes of a graph are discussed.

Keywords: ontology, semantic relations, subject area, knowledge graph.

DEVELOPMENT OF ONTOLOGIES OF SCIENTIFIC AND EDUCATIONAL WORK

T.A. Gavrilova (*gavrilova@gsom.spbu.ru*)
A.V. Kuznetsova (*menshikova@gsom.spbu.ru*)
I.A. Lescheva (*leshcheva@gsom.spbu.ru*)
St. Petersburg State University, St. Petersburg

The paper is devoted to the ontological engineering educational and scientific work of university teachers. The developed ontologies formed the basis of knowledge maps, which serve to simplify the processes of knowledge search and processing and allow you to get a visual "portrait" of a particular teacher or university department.

Keywords: ontologies, knowledge maps, knowledge management.

DESCRIPTION OF THE BASIC ELEMENTS OF THE METAGRAPH MODEL USING THE GRANULAR-PROTOGRAPH APPROACH

Yu.E. Gapanyuk (*gapyu@bmstu.ru*)
Bauman Moscow State Technical University, Moscow

The problem, which is solved in this paper, is connected with the fact that previous variants of formalization of metagraph model did not consider a unified basis, on the basis of which such different in nature elements of the model as vertices, metavertices and edges could be described. In this paper the granular-protograph model is proposed as such a basis. The prerequisites for combining the granular approach and the protograph model are considered. The basic elements of the granular-protograph model are proposed: facets of granules and granular protograph elements. The basic operations on the elements of the granular-protograph model are also proposed: the operation of signature fit and the operation of granule organization. The main provisions of the annotated metagraph model are considered. The representation of basic elements of metagraph model (vertices, metavertices, edges) by means of elements of granular-protograph model and operations introduced over them are proposed. A complex example of metagraph description with the help of granular-protograph model elements is considered.

Keywords: metagraph, metavertex, edge, protograph, granular-protograph model, facet of granule, protograph granular element.

LINGUISTIC KNOWLEDGE GRAPHS OF THE “TURKIC MORPHEME” WEB-PORTAL

A.R. Gatiatullin (*ayrat.gatiatullin@gmail.com*)

N.A. Prokop'ev (*nikolai.prokopyev@gmail.com*)

D.Sh. Suleymanov (*dvdt.slt@gmail.com*)

Institute of Applied Semiotics Academy of Sciences
of the Republic of Tatarstan, Kazan

The paper describes the knowledge graphs presented in the “Turkic Morpheme” web-portal, technologies for their creation and their use for natural language processing. Peculiarity of these knowledge graphs is that, on the one hand, they contain linguistic units of different linguistic levels, and, on the other hand, concepts denoting the meanings of these linguistic units that are built into the general system of concepts. These knowledge graphs are used for semantic annotation of electronic corpora presented within the portal, and are themselves updated with the information provided by these electronic corpora.

Keywords: knowledge graph, web-portal, linguistic unit.

ONTOLOGICAL PATTERNS OF KNOWLEDGE ON THE DIAGNOSIS OF PROCESSES

V.V. Gribova (*gribova@iacp.dvo.ru*)

E.A. Shalfeeva (*shalf@iacp.dvo.ru*)

Institute of Automation and Control Processes FEB RAS, Vladivostok, Russia;
Far Eastern Federal University, Vladivostok

Ontological patterns are a tool to reduce the complexity of developing intelligent systems. The paper presents a pattern for a class of diagnostic problems. It contains a complex of subject-independent semantic structures, a universal diagnostic solver and means of adapting a subject-independent pattern to specialized ones, so that expert knowledge is formed in subject terms.

Keywords: ontological pattern, diagnostic knowledge, domain-specific ontology, ontology adaptation.

A QUERY LANGUAGE FOR EXTENDED SEMANTIC NETS

D.V. Demidov (*dydemidov@mephi.ru*)

NRNU MEPhI, Moscow

The paper describes a query language for extended semantic nets that may have relations of different arity. Some algorithms of an interpreter of queries are given.

Keywords: extended semantic nets, query language, query interpreter.

KNOWLEDGE ENGINEERING IN RESILIENCE RESEARCH OF ENERGY AND ECOLOGICAL SYSTEM

L.V. Massel (*massel@isem.irk.ru*)

A.G. Massel (*amassel@gmail.com*)

Melentiev Energy Systems Institute of SB RAS, Irkutsk

The article describes the application of knowledge engineering in the course of research on the "elastic stability" of energy and ecological systems: ontological and cognitive engineering. A system of ontologies is shown that integrates ontologies of energy facilities and ontologies of the impact of energy facilities on the environment, examples of ontologies and cognitive maps are given. Computational experiments are described to assess the sustainability of energy and environmental systems, including those using machine learning methods.

Keywords: knowledge engineering, ontological and cognitive engineering, ontologies, cognitive maps, computational experiment.

OBJECT-ORIENTED ONTOLOGICAL ENGINEERING BASED ON META-ASSOCIATIVE GRAPHS

A.E. Misnik (*anton@misnik.by*)

Inter-state educational institution of higher education

“Belarusian-Russian University”, Mogilev

V.V. Borisov (*vbor67@mail.ru*)

The Branch of National Research University

“Moscow Power Engineering Institute” in Smolensk, Smolensk

The object-oriented approach is one of the most common approaches to the development of complex cyber-physical systems. It is based on the representation of the system as a set of interacting objects that work together to achieve the goals of the system. In the context of ontological engineering, this approach can be especially useful, since it simplifies the description and modeling of complex cyber-physical systems that include many interrelated objects and concepts. Also, the object-oriented approach simplifies the development and maintenance of ontologies that can be used to describe complex cyber-physical systems. The ontology of a cyber-physical system should be a convenient and flexible basis for modeling processes and ensuring the functioning of information-analytical and system processes. A metagraph can be used to build system ontologies and extract process ontologies. A modification of the metagraph (meta-associative graph) is proposed for inclusion in the ontology of events and system methods. Such an approach to constructing a metagraph makes it possible to include a process component in the ontological model of the system as an integral part of it. This allows you to flexibly and with less time to design, update and maintain process models based on segments of the general ontological model.

Keywords: object-oriented approach, ontological engineering, metagraphs, meta-associative graphs.

IMPLEMENTATION OF REQUESTS TO HYERARCHICAL GRAPH KNOWLEDGE AND DATA BASES ON THE IACPAAS PLATFORM

Ph.M. Moskalenko (*philipmm@iacp.dvo.ru*)

Institute of Automation and Control Processes FEB RAS, Vladivostok
Far-Eastern Federal University, Vladivostok

The paper presents mechanisms for support of formation and execution of ontological queries to data units (knowledge bases) built within the digraph connected two-level model of information representation. The use of the proposed approaches in knowledge and data processing should allow one to create such queries without intermediaries (programmers). The implementation of the proposed solutions on the IACPaaS platform, designed for the creation and use of cloud-based artificial intelligence systems, is also described.

Keywords: ontology, graph database, knowledge base, query editor, query language, IACPaaS.

GENERATION OF EVENT DATA BASED ON THE ONTOLOGICAL APPROACH

A.M. Namestnikov (*am.namestnikov@gmail.com*)

Ulyanovsk State Technical University, Ulyanovsk

Currently, the scientific direction of process mining (intelligent analysis of processes) is actively developing, which allows solving machine learning problems in relation to event data. The article presents a formal model of applied ontology, which is used to generate an event log based on the results of simulation experiments with a discrete event model. The description of the ontological system architecture, the structure of the main concepts with the corresponding semantic relations and sets of instances is given.

Keywords: ontology, event data, process, simulation model.

ORGANIZATION OF SECURE DATA ROUTING IN ELECTRIC POWER COMPLEXES USING ONTOLOGICAL AND DEEP MODELS

A.S. Fedulov (*director@sbmpei.ru*)

A.I. Lazarev (*anonymous.project@gmail.com*)

Branch of the National Research University «MPEI» in Smolensk, Smolensk

The existing distribution systems and traffic accounting of electric and thermal power structures are aimed at intelligent package delivery without data loss. One of the implementing these topologies problems is the presence of multiple backbone and managed switches, including mesh coverage. Side problems of these tools may be possible data leaks. The main solution is an approach to flexible management of the network electric power plants segments with the possibility of centralized changes analysis based on ontological variables.

Keywords: electric power complexes, ontological models, data classification, decision-making, fuzzy logic.

Section 2

Data Mining

ON THE FORMATION OF EMPIRICAL THEORIES IN POORLY FORMALIZED DOMAINS

S.M. Gusakova (*svem45@yandex.ru*)
FRC “Computer Science and Control” RAS, Moscow

M.A. Mikheyenkova (*m.mikheyenkova@yandex.ru*)
FRC “Computer Science and Control” RAS, Moscow
RSUH, Moscow

The paper presents the problems of the formation of new knowledge by means of artificial intelligence in open empirical domain, where there is largely no formal apparatus and the procedures for the formation of theories are heuristic in nature. The issues of data and knowledge representation, the formation of fact bases, and the choice of procedural solutions are considered. Approaches to specifying the algebraic operation of similarity, a key element of plausible reasoning by means of the JSM method of automated research support, are characterized for various structures. It is emphasized that effective solutions to these problems can be proposed only in conditions of interdisciplinary interaction.

Keywords: data and knowledge representation, fact base, JSM method, data structure, similarity operation.

TO THE PROBLEM OF EXPLANATION GENERATION OF THE INTELLIGENT DATA ANALYSIS RESULTS

M.I. Zabezhailo (*m.zabezhailo@yandex.ru*)
FRC “Computer Science and Control” RAS, Moscow
MIPT, Dolgoprudny

The possibilities to evaluate the acceptability of the results generated by artificial intelligence (AI) systems in the process of intelligent data analysis (IDA) are discussed. The generation of explanations is one of the convenient "tools" of such an assessment. Some well-known approaches to the formation of explanations are considered. Particular attention is paid to causal explanations, which are based on cause-effect dependencies extracted from the data in the process of IDA. Some algorithmic characteristics of causal explanations are given.

Keywords: artificial intelligence (AI), intelligent data analysis, empirical dependencies, causality, explanation in AI systems.

ABOUT SOME CURRENT MYTHS OF MODERN ARTIFICIAL

M.I. Zabezhailo (*m.zabezhailo@yandex.ru*)
FRC "Computer Science and Control" RAS, Moscow
MIPT, Dolgoprudny

M.A. Mikheyenkova (*m.mikheyenkova@yandex.ru*)
V.K. Finn (*v.k.finn@yandex.ru*)
FRC "Computer Science and Control" RAS, Moscow
RSUH, Moscow

Several widespread myths accompanying modern research and development in the field of artificial intelligence (AI) are considered. Some judgments are discussed regarding the properties and capabilities of the so-called Large Language Models, strong AI (artificial general intelligence), as well as the universal nature of approaches based on artificial neural networks. Three principles of the organization of scientific expertise are formulated, allowing to separate the actual AI research and development from the accompanying structures of mythological character.

Keywords: artificial intelligence, research and development, intelligent data analysis, scientific expertise.

INVESTIGATION OF THE WOS SAMPLES USING NEURAL NETWORK CLASSIFIERS

N.A. Nazarov (*straider105@gmail.com*)
V.O. Tolcheev (*tolcheevvo@mail.ru*)
National Research University "MPEI", Moscow

The study conducts an analysis of a text collection compiled from the Web of Science (WOS) database. The research is aimed at evaluating the quality of scientific article classification using modern neural network classifiers (all articles are provided in the form of bibliographic descriptions, which include titles and abstracts). Despite BERT's leadership in many areas of natural language processing, its performance on the WOS collection significantly depends on the "complexity" of the sample (the number of classes and their balance). For instance, on a "simple" sample (7 classes), the F1 score was 0.908, while on a "complex" sample (134 classes), it was 0.815. The study also explored less resource-intensive neural network approaches. In particular, high-quality indicators on the WOS collection were achieved using the FastText model for document representation and the Convolutional Neural Network (CNN) for classification (for the "simple" sample F1=0.888, for the "complex" sample F1=0.766). The combined application of FastText and CNN neural networks allows for a significant reduction in training time and classification speed compared to BERT.

Keywords: text classification, FastText, BERT, linear classifier, CNN, LSTM, machine learning.

APPROACH TO BUILDING A DATASET FOR THE PROBLEM OF EXTRACTING ARGUMENTATIVE RELATIONS

E.A. Sidorova (*lena@iis.nsk.su*)

Yu.A. Zagorulko (*zagor@iis.nsk.su*)

I.S. Kononenko (*irina_k@cn.ru*)

A.S. Sery (*alexey.seryj@iis.nsk.su*)

P.M. Chagina (*p.chagina@gmail.com*)

A.P. Ershov Institute of Informatics Systems of SB RAS, Novosibirsk

The paper considers data generation methods for training neural network models and solving problems of argumentation analysis. Data preparation includes the creation of an annotated corpus and the generation of a dataset for a specific task and the selected training model. Corpora manually marked up by experts can be supplemented with synthetic text data obtained by paraphrasing or back-translation of the texts of the base corpus. A feature of the task of generating a dataset is the need to recognize the components in the text and indicate their role in the argument (premise, thesis, counterargument). The role of argumentation indicators in the problem of data preparation is considered and a statistical study of the information content of indicators on the prepared ArgNet corpus is carried out.

Keywords: argumentation analysis, argumentative relations, annotated data creation, dataset, argumentation indicator, text segmentation.

INTELLIGENT SYSTEM FOR GENERATING MEANINGFUL POETIC TEXT

S.D. Kharitonov (*pintastd@gmail.com*)

D.G. Bukhanov (*dgbukh@gmail.com*)

Belgorod National Technological University named after S.D. Shukhov,
Belgorod

The paper considers various modern approaches to the generation of a poetic text. A comparative analysis of various language models is carried out, and the NLP pipeline architecture for automatic generation of meaningful poetic text based on a language model, a genetic algorithm, and a classifier of semantic coherence of the text is proposed and implemented.

Keywords: natural language processing, poetic text generation, language model, genetic algorithm, convolutional neural networks.

Section 3

Intelligent Agents, Robots, Intelligent Control, Computer Vision

OPTIMIZATION OF THE TEMPERATURE REGIME OF PHOSPHORITE PELLET ROASTING USING THE COEVOLUTIONAL BEE COLONY ALGORITHM

V.V. Bobkov (*vovabobkoff@mail.ru*)

O.V. Bulygina (*baguzova_ov@mail.ru*)

Branch of National Research University MPEI in Smolensk, Smolensk

The article deals with the scientific and practical problem of increasing the energy and resource efficiency of chemical-technological processes of the phosphorite pellet production from the apatite-nepheline ore waste. One of the ways to solve it is the search for the optimal temperature regime for heating the charge in a roasting conveyor machine. The specifics of the optimization task (nonlinearity of the objective function, large dimension of the search space, high computational complexity) determines the expediency of using stochastic methods, in particular, the bee colony algorithm. To select the values of "free" parameters that affect the algorithm convergence, it was proposed to use the idea of coevolution, which consists in running several ABC-algorithms in parallel with different "settings" for each subpopulation.

Keywords: energy and resource efficiency, phosphorite pellets, roasting conveyor machine, temperature regime, coevolution, bee colony algorithm, fuzzy logic.

ABOUT ROBOTS ORIENTATION BASED ON VISUAL LANDMARKS AND COMPASS

I.P. Karpova (*karpova_ip@mail.ru*)

National Research University Higher School of Economics, Moscow

The paper considers a method of using visual landmarks to memorize the path traveled by a mobile robot (animat), based on the mechanism of ant navigation. A route representation model and an algorithm for converting a direct route into a reverse route are described. The results of simulation experiments and experiments with real robots confirm the efficiency of the method.

Keywords: autonomous mobile robot, ant navigation, reverse route, foraging task.

DYNAMIC FORMATION OF ROBOT MOVEMENT ROUTE IN NON-DETERMINISTIC ENVIRONMENT WITH BYPASSING STATIONARY AND NON-STATIONARY OBSTACLES

D.V. Kotov (*dim.kot2009@yandex.ru*)
Military Academy of the GS of the RF AF, Moscow

O.B. Lebedev (*lebedev.ob@mail.ru*)
Southern Federal University, Taganrog

D.D. Fugarov (*ddf_1@mail.ru*)
Don State Technical University, Rostov-on-Don

The paper describes a hybrid algorithm for the dynamic formation of a robot's path in non-deterministic environments with bypassing stationary and non-stationary obstacles for two-dimensional space, based on the integration of wave and ant algorithms, which allows building trajectories of minimum length in real time with simultaneous optimization of a number of quality criteria for the constructed path. Identification of restrictions that prevent the laying of the trajectory from the current position is carried out in the process of constructing the trajectory. The process of laying the trajectory is carried out step by step. The entire trajectory connecting the initial position of the robot with the target position is a collection of separate sections. The time complexity of the algorithm depends on the lifetime of the colony l (number of iterations), the number of graph vertices n , the number of ants m , and is defined as $O(l \cdot n^2 \cdot m)$.

Keywords: trajectory planning, partial uncertainty, two-dimensional space, stationary and non-stationary obstacles, wave algorithm, ant colony algorithm, hybridization.

SKILL LEARNING WITH EMPOWERMENT IN REINFORCEMENT LEARNING

A.K. Latyshev (*latyshev.ak@phystech.edu*)
Moscow Institute of Physics and Technology, Moscow

A.I. Panov (*panov@airi.net*)
AIRI, Moscow

Federal Research Center "Computer Science and Control" RAS, Moscow

Unsupervised skill learning is one of the problems that intrinsically motivated reinforcement learning works on. In this article we propose the algorithm of forming skills based on empowerment evaluation that is the calculation of mutual information between the agent actions and states. This approach defines skills that lead to the near uniform distribution of achieved states. We examined the skills learned by the algorithm in the continuous control domain. We analyzed the performance of the reinforcement learning agent using empowerment-based skills.

Keywords: skill learning, mutual information, empowerment, intrinsic motivation, reinforcement learning.

THE SYNTHESIS OF HEURISTIC DECISIONS FOR ADAPTIVE CONTROL OF A MOBILE ROBOT IN A DYNAMIC ENVIRONMENT

M.V. Makarov (*nauka-murom@yandex.ru*)

I.A. Semenov (*79601712352@yandex.ru*)

N.S. Pidzhakova (*trantinanatalya@yandex.ru*)

Murom Institute of Vladimir State University, Murom

The paper presents theoretical information that reveals the essence of the phenomenon of synthesis of a new type of heuristic decision in adaptive control of a mobile robot. An experimental study justifying the technical feasibility of such a process is described, as well as indicating that its implementation leads to an increase in the adaptive abilities of the robot in a dynamic environment of existence.

Keywords: artificial intelligence, mobile robots, adaptive control, decision-making.

MULTI-AGENT TECHNOLOGY FOR ORGANIZING MANUFACTURING AND BUSINESS PROCESSES OF DIGITAL ENTERPRISES

Yu.F. Telnov (*Telnov.YUF@rea.ru*)

V.A. Kazakov (*Kazakov.VA @rea.ru*)

A.V. Danilov (*Danilov.AV @rea.ru*)

Plekhanov's Russian University of Economics, Moscow

The use of Industrie 4.0 platforms makes it possible to integrate the manufacturing and business processes of digital enterprises. To increase the flexibility of forming of value-added chains of interacting enterprises based on digital platforms, the use of multi-agent technology is proposed. To organize a multi-agent production system that ensures the implementation of the component structure of products and services of a digital enterprise, a set of microservices is substantiated in the active part of the assets administrative shells. The article proposes a role model for the interaction of agents (administrative shells) of a digital enterprise based on the ontology of the domain and the business model, which is determined depending on the stage of the products and services life cycle.

Keywords: digital platform, digital twin, digital enterprise components, asset administrative shell, multi-agent system, manufacturing and business processes, value-added chain.

AUTOMATIC REGISTRATION OF STUDENTS IN THE CLASSROOM USING COMPUTER VISION

O.I. Fedyaev (*olegfyayev@yahoo.com*)

I.A. Kolomoitseva (*bolatiger@mail.ru*)

Donetsk National Technical University, Donetsk

The purpose of the work is to eliminate the teacher's time losses that occur when manually registering the presence of a large number of students in the classroom. Automatic registration is carried out using computer vision in real time. The signs of a person's face are formed by a convolutional neural network. The control system performs the following functions: captures the faces of students on a video camera at the entrance to the auditorium, compares the faces with the database of students of the group, notes the presence at the lecture (or lateness) in case of successful identification, saves the data in an electronic journal. The system can be recommended to teachers of educational institutions.

Keywords: computer vision, student registration, convolutional neural network.

SCENARIO-INFORMATION ANALYSIS METHOD FOR INTELLIGENT MANAGEMENT OF ADAPTIVE TRAINING OF GROUPS OF SPECIALISTS

I.A. Frolov (*igor-frolov-81@mail.ru*)

Russian Federation Armed Forces Army Air Defense Military Academy,
Smolensk

The paper substantiates the need to intellectualize the management of adaptive training of specialists groups to perform complex tasks that are coordinated in time, place and goals in organizational and technical systems for various purposes. The approach to adaptive training is shown, as well as the tasks whose solution is in demand for managing this process. A method of scenario-information analysis for intelligent management of adaptive training of groups of specialists is proposed. The results of experimental studies on improving the level of preparedness of group specialists through intelligent management of their adaptive training are shown.

Keywords: adaptive training, scenario and information analysis method, groups of specialists, intelligent management, unified virtual space.

Section 4

Machine Learning, Neural Networks

ENHANCING IOT SECURITY THROUGH CONVOLUTIONAL VARIATIONAL AUTOENCODERS AND RESAMPLING TECHNIQUES WITH GAN

Huiyao Dong (*hydong@itmo.ru*)
ITMO University, Saint Petersburg

Igor Kotenko (*ivkote@comsec.spb.ru*)
St. Petersburg Federal Research Center of the Russian Academy
of Sciences, Saint Petersburg

The Internet of Things (IoT) is a pivotal constituent of the contemporary technological revolution and has experienced expeditious expansion in recent times. The proliferation of IoT devices has led to enhanced convenience and automation. However, the extensive deployment of IoT devices has also engendered concerns regarding data privacy and security. Among various detection and prevention methodologies, deep learning is emerging as a prominent trend. This paper utilises convolutional variational autoencoders (CNN-VAE) and resampling techniques for network attacks detection. The proposed methodology employs a hybrid data resampling technique to tackle the issue of imbalanced classes, followed by the implementation of a CNN-VAE classification model with a weighted loss function. The experiments demonstrate that the light-weighted CNN-VAE outperforms the baseline models. Therefore, it possesses the capability to effectively detect intrusive activities in real-world settings and strengthen the IoT security.

Keywords: Internet of Things, cyber attacks, variational autoencoder.

DEVELOPMENT OF NEURAL NETWORK MODULE FOR FORECASTING MASS BUILDING OBJECTS' ENERGY CONSUMPTION DEMAND

M.S. Bashlykov (*BashlykovMatS@mpei.ru*)
P.R. Varshavskii (*VarshavskyPR@mpei.ru*)
S.V. Guzhov (*GuzhovSV@mpei.ru*)
A.A. Sesin (*SesinAA@mpei.ru*)

National Research University "Moscow Power Engineering Institute", Moscow

Solution to the relevant task of forecasting mass building objects' energy consumption demand by utilizing artificial neural networks (ANN). Information about ANN models used for such task was analyzed. ANN model was built for a dataset, containing real data of schools built by the same plan. A neural network software prototype was developed for ANN model development by users without coding experience.

Keywords: machine learning, artificial neural network, demand forecasting, energy consumption.

NON-POLYNOMIAL NUMBER OF SIMILARITIES FOR RANDOM TRAINING SAMPLES

D.V. Vinogradov (*krrguest@yandex.ru*)
Federal Research Center “Computer Science and Control”
of the Russian Academy of Sciences, Moscow

In this paper we provide an additional argument to the probabilistic approach to knowledge acquisition using the similarity operation (VKF method) through the nonpolynomial number of all candidates. Concretely, we prove two results on such nonpolynomiality for random training samples generated by Bernoulli sequences. In the case of a dense lattice, it will be proved that the probability of occurrence of a large sublattice, isomorphic to Boolean algebra, in the lattice of candidates will tend to unity as the sample size increases. For a medium-density lattice, we reproduce a slightly modified argument of T. Sakurai about the nonpolynomiality of the average number of candidates.

Keywords: similarity, Bernoulli trials, candidates lattice, Boolean algebra.

REMOTE RESEARCH SYSTEM OF ARCHAEOLOGICAL SITES USING DEEP LEARNING

A.V. Vokhmintcev (*vav@csu.ru*)
Chelyabinsk State University, Chelyabinsk

A.V. Melnikov (*mav@csu.ru*)
Yugra Research Institute of Information Technologies, Khanty-Mansiysk

M.A. Romanov (*std.romanov.ma@gmail.com*)
Ural State University, Chelyabinsk

A.V. Shitselov (*a_shicelov@ugrasu.ru*)
Yugra State University, Khanty-Mansiysk

This work is devoted to the problem of the detection and research of archaeological sites based on machine learning methods, mapping and geophysics methods. In this paper a methodology was considered for mapping and surveying archaeological objects in the form of 2D and 3D models for objects with a large area, and machine learning methods based on CNN for detecting and classifying archaeological objects were suggested.

Keywords: residual neural networks, classification and segmentation of data, archaeological site.

INTELLIGENT MODEL OF INTEGRATED TECHNOLOGICAL SYSTEM FOR ORE WASTE PROCESSING

M.I. Dli (*midli@mail.ru*)

A.Yu. Puchkov (*putchkov63@mail.ru*)

National Research University «Moscow Power Engineering Institute» (Branch)
in Smolensk; Smolensk

The paper proposes a digital intellectual model of a complex technological system for processing ore waste, consisting of two technological lines. The model is based on the use of two deep recurrent neural networks, each of which serves as a digital copy of the corresponding line that is part of the system. The results of a simulation experiment are presented, which showed a high accuracy of the forecast generated by an intelligent model for variables characterizing the technological process of processing ore waste.

Keywords: deep neural networks, regression models.

APPROACH TO DETECTING ATTACKS AGAINST MACHINE LEARNING SYSTEMS USING A GENERATIVE ADVERSARIAL NETWORK

I.V. Kotenko (*ivkote@comsec.spb.ru*)

I.B. Saenko (*ibsaen@comsec.spb.ru*)

St. Petersburg Federal Research Center of the Russian Academy of Sciences
(SPC RAS), Saint-Petersburg

O.S. Lauta (*laos-82@yandex.ru*)

Admiral Makarov State University of Maritime and Inland Shipping,
Saint-Petersburg

N.A. Vasiliev (*vasn2020@mail.ru*)

V.E. Sadovnikov (*bladimir1998@mail.ru*)

Military Telecommunication Academy, Saint-Petersburg

The paper assesses the existing attacks on machine learning (ML) systems, taking into account the stages of training and operating of the information infrastructure. A systematization of attacks on ML systems is proposed, classes of attacks that are especially relevant and dangerous for ML systems are identified. Also, a study was made of potential measures to counter attacks on ML systems, which made it possible to present an approach to their protection based on the simultaneous use of Neural-Cleance and Jpeg-Compression technologies, as well as a generative-adversarial network. An experimental evaluation of the proposed approach.

Keywords: frameworks for protecting machine learning systems, machine learning, neural networks, artificial intelligence, computer security.

COGNITIVE STUDIES IN ARTIFICIAL INTELLIGENCE: NEURAL NETWORK APPROACH

A.A. Kharlamov (*kharlamov@analyst.ru*)

Institute of Higher Nervous Activity and Neurophysiology RAS, Moscow

Moscow State Linguistic University, Moscow

Higher School of Economics, Moscow,

Moscow Institute of Physics and Technology, Dolgoprudny

M.A. Pilgun (*pilgunm@yandex.ru*)

Russian State Social University, Moscow

The paper represents some results of a comparative study of the advantages and limitations of expert analysis using the TextAnalyst 2.3 neural network technology, and analysis performed with ChatGPT Plus on social media data generated by actors with different genesis to identify social tension. Data collection was carried out on 01.10.2022-31.03.2023 in social media Vkontakte, Telegram, YouTube, Zen. The dataset contains 22,199,085 tokens.

Keywords: neural network technologies, ChatGPT, cognitive research, social media, text analysis.

АВТОРСКИЙ УКАЗАТЕЛЬ

Андриевская Н.К.	33	Котенко И.В.	316, 366
Атаева О.М.	46	Котов Д.В.	254
Башлыков М.С.	329	Кузнецова А.В.	60
Бобков В.И.	234	Лазарев А.И.	157
Боргест Н.М.	6	Латышев А.К.	264
Борисов В.В.	123	Лаута О.С.	366
Бульгина О.В.	234	Лебедев О.Б.	254
Буханов Д.Г.	223	Лещева И.А.	60
Варшавский П.Р.	329	Макаров М.В.	275
Васильев Н.А.	366	Мартыненко Т.В.	33
Васяева Т.А.	33	Массель А.Г.	113
Виноградов Д.В.	337	Массель Л.В.	113
Вохминцев А.В.	344	Мельников А.В.	344
Гаврилова Т.А.	60	Мисник А.Е.	123
Гапанюк Ю.Е.	68	Михеенкова М.А.	169, 190
Гатиатуллин А.Р.	81	Москаленко Ф.М.	134
Городецкий В.И.	14	Назаров Н.А.	201
Грибова В.В.	93	Наместников А.М.	146
Гужов С.В.	329	Панов А.И.	264
Гусакова С.М.	169	Пиджакова Н.С.	275
Данилов А.В.	282	Пильгун М.А.	377
Демидов Д.В.	101	Прокопьев Н.А.	81
Дли М.И.	356	Пучков А.Ю.	356
Донг Х.	316	Романов М.А.	344
Забежайло М.И.	180, 190	Садовников В.Е.	366
Загорулько Ю.А.	211	Саенко И.Б.	366
Казаков В.А.	282	Семенов И.А.	275
Карпова И.П.	242	Серебряков С.А.	46
Коломойцева И.А.	294	Серый А.С.	211
Кононенко И.С.	211	Сесин А.А.	329

Сидорова Е.А.	211	Фролов И.А.	304
Сулейманов Д.Ш.	81	Фугаров Д.Д.	254
Тельнов Ю.Ф.	282	Харитонов С.Д.	223
Толчеев В.О.	201	Харламов А.А.	377
Тучкова Н.П.	46	Чагина П.М.	211
Федулов А.С.	157	Шалфеева Е.А.	93
Федяев О.И.	294	Шицелов А.В.	344
Финн В.К.	30, 190		

СОДЕРЖАНИЕ

ПЛЕНАРНЫЕ ДОКЛАДЫ

Н.М. Боргест ОНТОЛОГИЯ ПРОЕКТИРОВАНИЯ: ГЕНЕЗИС И РАЗВИТИЕ	6
В.И. Городецкий ОТ ИНЖЕНЕРИИ ЗНАНИЙ К НАУКЕ О ЗНАНИЯХ	14
В.К. Финн АППАРАТ ПОНЯТИЙ ДСМ-МЕТОДА АВТОМАТИЗИРОВАННОЙ ПОДДЕРЖКИ ИССЛЕДОВАНИЙ	30
Секция 1. ИНЖЕНЕРИЯ ЗНАНИЙ	
Н.К. Андриевская, Т.А. Васяева, Т.В. Маргынченко РАЗРАБОТКА ОНТОЛОГИИ ПРЕДМЕТНОЙ ОБЛАСТИ ДЕЯТЕЛЬНОСТИ НАУЧНО-ОБРАЗОВАТЕЛЬНОЙ ОРГАНИЗАЦИИ	33
О.М. Атаева, С.А. Серебряков, Н.П. Тучкова СОЗДАНИЕ ГРАФА ЗНАНИЙ НА ОСНОВЕ ОНТОЛОГИИ СЕМАНТИЧЕСКОЙ БИБЛИОТЕКИ	46
Т.А. Гаврилова, А.В. Кузнецова, И.А. Лещева К ВОПРОСУ О РАЗРАБОТКЕ ОНТОЛОГИЙ НАУЧНОЙ И УЧЕБНОЙ РАБОТЫ.....	60
Ю.Е. Гапанюк ОПИСАНИЕ БАЗОВЫХ ЭЛЕМЕНТОВ МЕТАГРАФОВОЙ МОДЕЛИ С ИСПОЛЬЗОВАНИЕМ ГРАНУЛЯРНО-ПРОТОГРАФОВОГО ПОДХОДА	68
А.Р. Гатиатуллин, Н.А. Прокопьев, Д.Ш. Сулейманов ЛИНГВИСТИЧЕСКИЕ ГРАФЫ ЗНАНИЙ ПОРТАЛА «ТЮРКСКАЯ МОРФЕМА»	81
В.В. Грибова, Е.А. Шалфеева ОНТОЛОГИЧЕСКИЕ ПАТТЕРНЫ ЗНАНИЙ ПО ДИАГНОСТИКЕ ПРОЦЕССОВ.....	93
Д.В. Демидов ЯЗЫК ЗАПРОСОВ К РАСШИРЕННЫМ СЕМАНТИЧЕСКИМ СЕТЯМ...	101
Л.В. Массель, А.Г. Массель ИНЖЕНЕРИЯ ЗНАНИЙ В ИССЛЕДОВАНИЯХ УСТОЙЧИВОСТИ ЭНЕРГЕТИЧЕСКИХ И ЭКОЛОГИЧЕСКИХ СИСТЕМ.....	113

А.Е. Мисник, В.В. Борисов	
ОБЪЕКТНО-ОРИЕНТИРОВАННЫЙ ОНТОЛОГИЧЕСКИЙ ИНЖИНИРИНГ НА ОСНОВЕ МЕТА-АССОЦИАТИВНЫХ ГРАФОВ.....	123
Ф.М. Москаленко	
РЕАЛИЗАЦИЯ ЗАПРОСОВ К ИЕРАРХИЧЕСКИМ ГРАФОВЫМ БАЗАМ ЗНАНИЙ И ДАННЫХ НА ПЛАТФОРМЕ IASRAAS	134
А.М. Наместников	
ГЕНЕРАЦИЯ СОБЫТИЙНЫХ ДАННЫХ НА ОСНОВЕ ОНТОЛОГИЧЕСКОГО ПОДХОДА	146
А.С. Федулов, А.И. Лазарев	
ОРГАНИЗАЦИЯ БЕЗОПАСНОЙ МАРШРУТИЗАЦИИ ДАННЫХ В ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ КОМПЛЕКСАХ С ИСПОЛЬЗОВАНИЕМ ОНТОЛОГИЧЕСКИХ И ГЛУБОКИХ МОДЕЛЕЙ	157
Секция 2. ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ	
С.М. Гусакова, М.А. Михеенкова	
О ФОРМИРОВАНИИ ЭМПИРИЧЕСКИХ ТЕОРИЙ В ПЛОХО ФОРМАЛИЗОВАННЫХ ОБЛАСТЯХ	169
М.И. Забейло	
К ПРОБЛЕМЕ ОБЪЯСНЕНИЯ РЕЗУЛЬТАТОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ	180
М.И. Забейло, М.А. Михеенкова, В.К. Финн	
О НЕКОТОРЫХ АКТУАЛЬНЫХ МИФАХ СОВРЕМЕННОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА	190
Н.А. Назаров, В.О. Толчеев	
ИССЛЕДОВАНИЕ ВЫБОРОК WOS С ПОМОЩЬЮ НЕЙРОСЕТЕВЫХ КЛАССИФИКАТОРОВ	201
Е.А. Сидорова, Ю.А. Загоруйко, И.С. Кононенко, А.С. Серый	
П.М. Чагина	
ПОДХОД К ПОСТРОЕНИЮ ДАТАСЕТА ДЛЯ ЗАДАЧИ ИЗВЛЕЧЕНИЯ АРГУМЕНТАТИВНЫХ ОТНОШЕНИЙ.....	211
С.Д. Харитонов, Д.Г. Буханов	
ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ГЕНЕРАЦИИ ОСМЫСЛЕННОГО СТИХОТВОРНОГО ТЕКСТА	223

Секция 3. ИНТЕЛЛЕКТУАЛЬНЫЕ АГЕНТЫ, РОБОТЫ, ИНТЕЛЛЕКТУАЛЬНОЕ УПРАВЛЕНИЕ, КОМПЬЮТЕРНОЕ ЗРЕНИЕ

В.И. Бобков, О.В. Булыгина ОПТИМИЗАЦИЯ ТЕМПЕРАТУРНОГО РЕЖИМА ОБЖИГА ФОСФОРИТОВЫХ ОКАТЫШЕЙ С ИСПОЛЬЗОВАНИЕМ КОЭВОЛЮЦИОННОГО АЛГОРИТМА ПЧЕЛИНЫХ КОЛОНИЙ	234
И.П. Карпова К ВОПРОСУ ОБ ОРИЕНТАЦИИ РОБОТОВ НА ОСНОВЕ ВИЗУАЛЬНЫХ ОРИЕНТИРОВ И КОМПАСА	242
Д.В. Котов, О.Б. Лебедев, Д.Д. Фугаров ДИНАМИЧЕСКОЕ ФОРМИРОВАНИЕ МАРШРУТА ПЕРЕМЕЩЕНИЯ РОБОТА В НЕДЕТЕРМИНИРОВАННЫХ СРЕДАХ С ОБХОДОМ СТАЦИОНАРНЫХ И НЕСТАЦИОНАРНЫХ ПРЕПЯТСТВИЙ.....	254
А.К. Латышев, А.И. Панов ФОРМИРОВАНИЕ УМЕНИЙ АГЕНТА ПО ПРИНЦИПУ ДОСТИЖИМОСТИ В ОБУЧЕНИИ С ПОДКРЕПЛЕНИЕМ.....	264
М.В. Макаров, И.А. Семенов, Н.С. Пиджакова СИНТЕЗ ЭВРИСТИЧЕСКИХ РЕШЕНИЙ В РАМКАХ АДАПТИВНОГО УПРАВЛЕНИЯ МОБИЛЬНЫМ РОБОТОМ В ДИНАМИЧЕСКОЙ СРЕДЕ	275
Ю.Ф. Тельнов, В.А. Казаков, А.В. Данилов МНОГОАГЕНТНАЯ ТЕХНОЛОГИЯ ОРГАНИЗАЦИИ ПРОИЗВОДСТВЕННЫХ И БИЗНЕС-ПРОЦЕССОВ ЦИФРОВЫХ ПРЕДПРИЯТИЙ.....	282
О.И. Федяев, И.А. Коломойцева АВТОМАТИЧЕСКАЯ РЕГИСТРАЦИЯ ПРИСУТСТВИЯ СТУДЕНТОВ НА УЧЕБНОМ ЗАНЯТИИ С ПОМОЩЬЮ КОМПЬЮТЕРНОГО ЗРЕНИЯ	294
И.А. Фролов МЕТОД СЦЕНАРНО-ИНФОРМАЦИОННОГО АНАЛИЗА ДЛЯ ИНТЕЛЛЕКТУАЛЬНОГО УПРАВЛЕНИЯ АДАПТИВНОЙ ПОДГОТОВКОЙ ГРУПП СПЕЦИАЛИСТОВ	304

Секция 4. МАШИННОЕ ОБУЧЕНИЕ, НЕЙРОСЕТЕВЫЕ МЕТОДЫ

Huiyao Dong, Igor Kotenko ENHANCING IOT SECURITY THROUGH CONVOLUTIONAL VARIATIONAL AUTOENCODERS AND RESAMPLING TECHNIQUES WITH GAN.....	316
М.С. Башлыков, П.Р. Варшавский, С.В. Гужов, А.А. Сесин РАЗРАБОТКА НЕЙРОСЕТЕВОГО МОДУЛЯ ДЛЯ ПРОГНОЗИРОВАНИЯ СПРОСА НА ЭНЕРГОПОТРЕБЛЕНИЕ ОБЪЕКТАМИ МАССОВОГО СТРОИТЕЛЬСТВА.....	329
Д.В. Виноградов НЕПОЛИНОМИАЛЬНОСТЬ ЧИСЛА СХОДСТВ ДЛЯ СЛУЧАЙНЫХ ОБУЧАЮЩИХ ВЫБОРОК.....	337
А.В. Вохминцев, А.В. Мельников, М.А. Романов, А.В. Шицелов СИСТЕМА ДИСТАНЦИОННОГО ИССЛЕДОВАНИЯ АРХЕОЛОГИЧЕСКИХ ПАМЯТНИКОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ГЛУБОКОГО МАШИННОГО ОБУЧЕНИЯ	344
М.И. Дли, А.Ю. Пучков ИНТЕЛЛЕКТУАЛЬНАЯ МОДЕЛЬ КОМПЛЕКСНОЙ ТЕХНОЛОГИЧЕСКОЙ СИСТЕМЫ ПЕРЕРАБОТКИ РУДНЫХ ОТХОДОВ	356
И.В. Котенко, И.Б. Саенко, О.С. Лаута, Н.А. Васильев, В.Е. Садовников ПОДХОД К ОБНАРУЖЕНИЮ АТАК НА СИСТЕМЫ МАШИННОГО ОБУЧЕНИЯ С ИСПОЛЬЗОВАНИЕМ ГЕНЕРАТИВНО- СОСТЯЗАТЕЛЬНОЙ СЕТИ	366
А.А. Харламов, М.А. Пильгун КОГНИТИВНЫЕ ИССЛЕДОВАНИЯ В ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ: НЕЙРОСЕТЕВОЙ И ГЕНЕРАТИВНЫЙ ИНСТРУМЕНТАРИЙ.....	377
Abstracts of Plenary Reports	386
Abstracts of Sectional Reports.....	388
АВТОРСКИЙ УКАЗАТЕЛЬ	404

Научное издание

Двадцать первая Национальная конференция
по искусственному интеллекту с международным участием
КИИ-2023
Труды конференции в 2-х томах
Том 1

Подписано в печать 19.09.2023 г.
Формат 60x84¹/₁₆. Тираж 300 экз. Усл. печ. л. 23,8.

Типография ООО «Принт-Экспресс»



ISBN 978-5-91812-231-0